

Memory for Time

Heather Van Volkinburg

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in the  
Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY  
2012

© 2012

Heather Van Volkinburg

All rights reserved

## ABSTRACT

### Memory for Time

Heather Van Volkinburg

Research on the perception of time has focused on isolating an internal time keeping mechanism. This focus has caused research in this domain to stay focused on perceptual mechanisms. However, information that has been perceived must also be stored into memory. The dominant model of time perception, SET, specifies a clock stage, a memory stage, and a decision stage (Gibbon, 1977), but there has been comparatively more research focusing on the internal clock than on memory mechanisms. This dissertation focuses on the memory for time by incorporating manipulations used in non-temporal memory research into timing tasks. Chapter 1 targets working memory for time and shows that brief delays between learning and recall cause temporal estimates to lengthen. Chapter 2 targets retention and storage of multiple time intervals over a delay of a few minutes and shows that the estimates of target times learned together will migrate towards each other in memory. We also showed that scalar variability arises at retrieval. Chapter 3 attempts to reconfigure a categorization task to target implicit memory for time. Overall, the research demonstrates that exploring memory mechanisms for time will increase our understanding of time perception and provide us with information that focusing on the internal clock will not.

## Table of Contents

General Introduction .....	1
Chapter 1: Working Memory for Time	
Introduction .....	7
Experiment 1	
Methods .....	10
Results .....	13
Discussion .....	15
Experiment 2	
Methods .....	17
Results.....	18
Discussion .....	19
General Discussion .....	19
Chapter 2: Memory for Multiple Time Intervals	
Introduction .....	24
Experiment 1	
Methods .....	27
Results .....	29
Discussion .....	32
Experiment 2	
Methods .....	37
Results .....	38
Discussion .....	42

Experiment 3	
Methods .....	44
Results .....	45
Discussion .....	48
General Discussion .....	50
Chapter 3: Implicit and Explicit Timing	
Introduction .....	54
Experiment 1	
Methods.....	58
Results .....	63
Discussion .....	68
Experiment 2	
Methods .....	71
Results .....	72
Discussion .....	76
Experiment 3	
Methods .....	77
Results .....	78
Discussion .....	80
General Discussion .....	82
General Conclusions .....	86
References .....	95

## List of Figures

Figure 1: Model of Scalar Expectancy Theory (SET) .....	3
Figure 2: Proportional errors (PEs) for Experiment 1, Chapter 1 .....	13
Figure 3: Proportional errors (PEs) for Experiment 2, Chapter 1 .....	18
Figure 4: Proportional errors (PEs) for Experiment 1, Chapter 2 .....	31
Figure 5: Proportional errors (PEs) for Experiment 2, Chapter 2 .....	41
Figure 6: Proportional errors (PEs) for Experiment 3, Chapter 2 .....	46
Figure 7: Coefficient of variation (CVs) for Experiment 3, Chapter 2 .....	47
Figure 8: Example trials for training in Experiments 1, 2, & 3, Chapter 3 .....	60
Figure 9: Accuracy for training blocks in Experiment 1, Chapter 3 .....	64
Figure 10: Accuracy for testing trials in Experiment 1, Chapter 3 .....	65
Figure 11: Probability of responding long in Experiment 1, Chapter 3 .....	67
Figure 12: Accuracy for training blocks in Experiment 2, Chapter 3 .....	72
Figure 13: Accuracy for testing trials in Experiment 2, Chapter 3 .....	74
Figure 14: Probability of responding long in Experiment 2, Chapter 3 .....	75
Figure 15: Accuracy for training blocks in Experiment 3, Chapter 3 .....	78
Figure 16: Probability of responding long in Experiment 3, Chapter 3 .....	79

## List of Tables

Table 1: List of target times used in Experiment 1, Chapter 1 .....	11
Table 2: Coefficient of variation (CVs) for Experiments 1 & 2, Chapter 1 .....	14
Table 3: Coefficient of variation (CVs) for Experiments 1 & 2, Chapter 2 .....	30
Table 4: Pearson correlation coefficients for Experiment 1, Chapter 2 .....	31
Table 5: Mean estimates for retraining and testing phases in Experiment 2, Chapter 2 ..	40
Table 6: Pearson correlation coefficients for Experiment 2, Chapter 2 .....	41
Table 7: Pearson correlation coefficients for Experiment 3, Chapter 2 .....	47
Table 8: Verbal estimates given in Experiment 1, Chapter 3 .....	67
Table 9: Reaction times from training in Experiment 2, Chapter 3 .....	73
Table 10: Verbal estimates given in Experiment 2, Chapter 3 .....	76
Table 11: Verbal estimates given in Experiment 3, Chapter 3 .....	80

## Acknowledgements

First of all I would like to thank all of the research assistants that worked with me for all their hard work and who, with everyone else in the lab, allowed me to yell at the computer uninhibited when things weren't going my way and who laughed with me when it all made sense again. I am grateful to the friends and colleagues who helped me work through the little things so that I was better equipped to tackle the big things. I'm also very grateful to my mother who listened patiently to all the ups and downs and consistently reminded me of why I was doing this.

I am truly indebted and thankful to Peter Balsam for his support and guidance throughout graduate school and for having the confidence in me to allow me to set up my own space to explore our shared interest in time perception from a completely different perspective than his. I learned a lot about who I am and what I'm made of that I wouldn't have learned working with anyone else.

Without David's unconditional love and support, I can't imagine where I would be now, and quite frankly, I don't want to.



*Think of time, not as something that is lost or spent, but as a companion that reminds us to cherish every moment, because it will never come again.*

Interval timing is the branch of psychology that focuses on the study of the perception and estimation of arbitrary intervals of time. It speaks to our experiences of time flying while having fun, or why a watched pot never boils. When our impatience increases as a webpage doesn't load or a subway train doesn't arrive, we use our ability to time arbitrary intervals to know that we've waited long enough.

Interval timing in human participants can be assessed using a variety of tasks, which break down into three major task types, discrimination, reproduction, and verbal estimation. Discrimination tasks present a standard interval with a visual, auditory, or tactile cue. A second interval is presented (typically in the same modality) and participants must decide if the second interval is the same, longer, or shorter than the standard. In a reproduction task, an interval is presented to subjects via a visual, auditory, or tactile cue and participants are required to reproduce the interval by marking it with a key press. Sometimes participants are required to produce an interval presented via a verbal label (e.g. 3 seconds) and this is typically called production instead of reproduction. In the last major task type, verbal estimation, participants are presented with a duration using visual, auditory, or tactile cues and must provide a verbal estimate of the duration. The first two chapters use a reproduction task, although much of the relevant literature uses discrimination tasks. The final chapter uses a discrimination task.

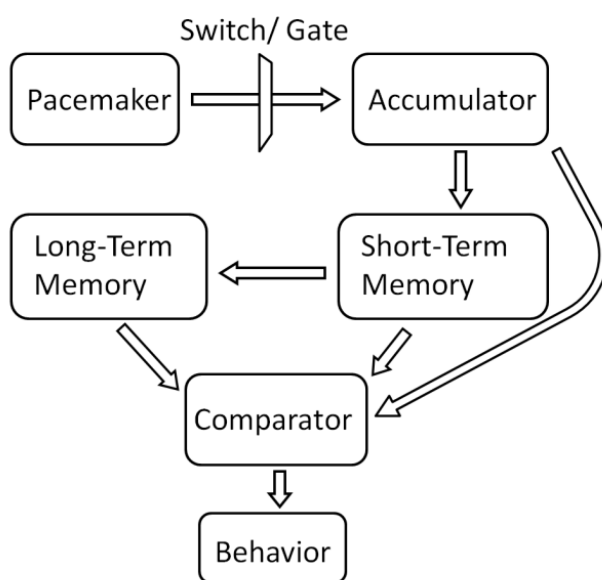
The distinction between temporal reproduction and discrimination can be thought of analogously to the distinction between recall and recognition memory in non-temporal domains. In memory tasks that focus on recall, participants must retrieve memory representations of the

target. For example, after being presented with a list of words, participants must generate a list containing the words they remember from the list. Recognition tasks, in contrast, will require participants to view an alternate list of words and choose the words that they remember from the first list. These distinctions assume that recall tasks require more effort. Since a temporal reproduction task requires participants to retrieve a memory of the target time for reproduction, we assumed that it was analogous to a memory recall task. In contrast, in temporal discrimination tasks participants are provided with alternate durations and they must remember if it was the one they had seen before.

The core concepts for a model of temporal information processing are a clock, a memory store, and a decision processor (Church, 2006; Grondin, 2005, Nobre, O'Reilly, 2004; Treisman, 1963). There are other proposed models of time perception, which may better reflect neurological processes involved in time perception, such as the oscillator model in which an oscillation of a mean period represents the time duration in the memory store and may account for the internal clock mechanism (Matell & Meck, 2004; Treisman, et al., 1990) or the multiple time scale model in which the decay of a memory trace represents the time duration (Staddon & Higa, 1999). However, throughout this paper, I will focus on the internal clock model and specifically scalar expectancy theory (SET; Gibbon, 1977; Gibbon, Church, & Meck, 1984) because it provides a clear metaphor that facilitates the explanation of data (Macar & Vidal, 2009).

The internal clock, which has also been called a counter or a pacemaker, is an internal timekeeping device that generates a countable rate of pulses. Time is measured by the number of pulses that accumulate in the accumulator during a specific interval (Fig. 1). The number of pulses that accumulate for a specific interval can be affected by the rate of the clock such that an

increase in clock rate would increase pulse count. For example, increases in arousal are known to increase estimates, likely through the increase of clock rate (Fetterman & Killeen, 1995; Meck, 1996; Penton-Voak, et al., 1996; Treisman, et al., 1990; Wearden & Penton-Voak, 1995). Likewise, boredom has been suggested as a cause for decreases in estimates, likely because of lower levels of arousal, and therefore, decreased clock rate (Wearden, et al., 1999).



*Figure 1: The pacemaker (clock/timer) emits pulses at a given rate which pass through an attention modulated gate/switch and accumulate in an accumulator. The memory stage is represented by a short-term memory store and a long-term memory store. At the decision stage, accumulated information is compared to representation(s) in memory leading to a behavior.*

The number of accumulated pulses can also be affected by the amount of attention directed towards time such that an increase in attention increases pulse count (Brown, 1997, 2008; Grondin, 2005; Macar, et al., 1994). Between the clock and the accumulator is an attention

modulated gate (Zakay & Block, 1996; Zakay, 1998), or switch (Lejeune, 1998), or both (Zakay & Block, 1997). The gate and switch predict separate effects of attention on time estimation. Thinking of attention as a spotlight, the gate is the size of the beam; the larger the beam, the more area covered in light, and more pulses accumulate. In contrast, the switch in SET is the on/off switch of the spotlight; the latency for turning it on will affect the number of pulses that accumulate. The behavior of the gate makes it very difficult to distinguish it from clock rate, since both an increase in clock rate and an increase in attention yield an increase in pulses to the accumulator. Changes in the switch can be differentiated from clock rate if more than one interval is being assessed and switch latency is identical for the different durations because the relative effect of switch latency will decrease with increasing target time. The clock, attentional mechanisms, and the accumulator could be thought of as the perceptual (or encoding) stage of temporal information processing. A lot of research has focused on this level of interval timing, in part motivated by the search for the internal clock in the brain.

After the clock stage of SET (which includes the clock, attentional mechanisms, and the accumulator) temporal information is shifted to the memory stage that includes both a short-term and long-term memory store (Gibbon, et al., 1984; Meck, 1996). The next three chapters focus on different aspects of memory for time that will be unpacked in the introductory sections that begin each chapter.

The final stage outlined in the SET model is the comparison stage, in which accumulated temporal information is compared to representations in memory. When a specific threshold is met, a response is made, but the threshold may be some amount that is near the target time, and not specifically the exact target time. The threshold can be manipulated with payoff schemes, in which money is used as a motivator for increased perceptual distinction, as is found with other

perceptual stimuli (Wearden & Culpin, 1998). All measurable behavior is the cumulated outcome of these three stages.

SET also proposes that learned intervals obey the scalar property, which is a domain specific form of Weber's law (Gibbon, et al, 1984). The variability of estimates increase proportionally such that the coefficient of variation, or the standard deviation divided by the mean, remains constant. It is unclear where scalar variance arises in SET and it could be from the internal clock, from storage or retrieval processes, or from the decision stage (Gibbon, et al., 1984; Meck, 1996; Wearden, 1999). The internal clock has been assumed to emit pulses according to a Poisson distribution (Gibbon, 1992; Meck, 1996), however, this implies that relative variability would decrease with increasing target time, a behavior that would not demonstrate the scalar property (Gibbon, 1992; Staddon & Higa, 1999). While it is plausible that the pacemaker rate varies from trial to trial due to a normally distributed random variable, or bias, which causes temporal estimates to remain scalar, it is also plausible that scalar variability arises from the memory stage of SET (Gibbon, et al. 1984; Gibbon, 1992; Wearden, 1999). Since the following chapters focus on memory for time, we can assess adherence to the scalar property in memory.

The three following chapters each represent separate research initiatives. The first, 'Working Memory for Time' consists of two experiments that show the effects of brief delays inserted between learning and immediate recall of target times. This was designed with the intention of targeting active working memory processes. The second chapter, 'Memory of Multiple Time Intervals', consists of three experiments that show the effects of learning multiple intervals together and recalling them after a longer delay. This was designed to assess the storage

and retrieval mechanisms when multiple target times are learned together. The final chapter, 'Implicit and Explicit Timing' consists of three experiments designed to show the implicit learning and usage of temporal information.

## Chapter 1

### Working Memory for Time

There is mounting evidence that millisecond (ms) durations are processed differently than supra-second (ss) durations. Lewis & Miall (2003a, 2006) theorized that there are dichotomous timing mechanisms, automatic timing and cognitive timing. Automatic timing occurs when estimating brief intervals (ms) especially when repetitive or skilled movement is used to measure timing and when the task presents target times in a repeated, continuous, or predictable fashion. In contrast, cognitive timing occurs when measuring longer intervals (ss), in tasks that do not use temporally skilled motor movement, and in which the target time is presented using a discontinuous and unpredictable method. Additionally, there is evidence that when ms target times and ss target times are learned together the ms target times are overestimated while the ss target times are underestimated with respect to the target time (Jones, et al., 2004; Koch, et al., 2007). This behavioral difference may be due to the automatic and cognitive timing mechanisms associated with ms and ss target times respectively. Estimates for the ms target times could be overestimated because the response time is added to the estimate. In contrast, the cognitive timing involved in the ss target time could cause estimates to be shorter because the cognitive processes detract from the accumulation of temporal information. Cognitive processes, especially those associated with working memory, are thought to be important for timing (Fortin & Breton, 1995; Fortin et al., 1993; Gibbon, 1977; Wearden, 1999, Miall & Lewis, 2003a). Working memory includes a variety of cognitive components, such as maintenance of information, inhibiting the influence of irrelevant information, and task switching. Working memory processes, including maintenance, are thought to be mediated by frontal regions (Smith &

Jonides, 1991). Since frontal regions are also implicated in processing ss target times (Harrington, et al., 1998; Jones, et al., 2004; Lewis & Miall, 2003; Mangels, et al. 1998), we proposed that maintenance processes, specifically, are required during the accumulation of longer intervals and may contribute to the underestimation of long durations.

Scalar expectancy theory (SET) can be used to explain this. As stated in the introduction, in SET, there are three stages of temporal information processing (Gibbon, 1977; Gibbon, Church, & Meck, 1984; Treisman, 1963). The clock stage consists of an internal clock (pacemaker/timer), a gate (switch), and an accumulator. The memory stage consists of a short-term memory store and a long-term memory store. The decision or comparator stage consists of a 'comparator' and represents the decision made when comparing temporal information in the accumulator and information in the relevant memory store. When a specific criterion is met, a behavioral response is made (Fig. 1). Supra-second target times are typically underestimated because as the length of the target duration increases, maintenance processes become increasingly necessary in order to keep temporal information active and available for use with limited decay. We propose that the underestimation of the ss target times is caused by decay of the temporal information during the accumulation process.

Traditionally, the length of time that information can remain active in working memory has been tested using delays in recall tasks (Jonides et al, 2008). We decided to incorporate a delay into a temporal reproduction task. Since temporal information is dynamic and must accumulate over time we propose that longer target times are underestimated because the earlier temporal information decays while the ongoing temporal information accumulates. We propose that delays following ms target times would also lead to underestimation because the temporal information would decay over the delay. Delays have been used in previous studies in



discrimination timing tasks and have demonstrated a phenomenon called subjective shortening (Wearden & Ferrara, 1993, Wearden, et al, 2002). Using temporal generalization, in which participants are required to specify if a comparator time is longer, shorter, or the same as a previously presented standard target time, Wearden & colleagues found that accuracy increased when the comparator was longer, decreased when the comparator was the same, and remained the same when the comparator was shorter as the length of the delay between the standard target time and comparison increased. This response pattern suggested that the memory representation of the standard target time decreased with increasing delay implying that temporal information decayed. Based on this and similar findings we expected that increasing delays would lead to shortened reproductions and that this would occur with the ss and ms target times.

Our study, therefore, includes both ms and ss target times. In addition, we wanted to test our hypothesis that temporal information began decaying even while still accumulating. Therefore, we made the lengths of the delay, when summed with the ms target times, equal the length of the ss target times. We anticipated that the amount of underestimation of the ms target times coupled with a delay would be similar to the amount of underestimation of the matched ss target times.

We incorporated multiple ms and ss target times into a roving referent paradigm in which the specific duration presented in a trial was pulled from a range of possible times (Allan, 1977). We wanted to minimize long-term encoding of the target time so that we could assess how the short-term memory representation was affected by the delay. Long-term encoding of a target time would likely minimize the effect of the delay because participants could rely on long-term representations as a basis for reproductions instead of having to maintain the information over the delay. Changing the specific target time from trial to trial has been the preferred procedure in

previous studies to minimize long-term encoding of target times (Wearden & Ferrara, 1993, Wearden, et al., 2002; Koch et al., 2007; Jones, et al., 2004). We matched delay lengths of the ms target times so that the sum of the target and delay times equaled the duration of the supra-second target times in order to assess if underestimations of the ms and matched ss target times were similar.

## Experiment 1

### Methods

*Participants:* Thirty-one participants (29 female and 2 male) were recruited from Barnard College. Their ages ranged from 18 to 35 years with a mean age of 20. Participants were naïve to the purpose of the experiment and were compensated for their time with course credit.

*Materials:* Experimental events were presented using the psychophysical toolbox (Brainard, 1997; Pelli, 1997) compiled into a Matlab® (2007a, The MathWorks, Natick, MA) executable program presented on Microsoft XP® OS PC's. Keyboard responses were recorded to the nearest millisecond.

*Procedure:* We used a roving referent paradigm to prevent long-term encoding of the target times. Millisecond targets, averaging 0.58s, consisted of five durations in logarithmic steps ranging from 0.40s to 0.80s (0.40s, 0.48s, 0.57s, 0.67s, and 0.80s). Supra-second targets, averaging 2.91s, consisted of five durations in log steps ranging from 2.00s to 4.00s (2.00s, 2.38s, 2.83s, 3.36s, 4.00s). Delay one was designed as a control with a short fixed inter-stimulus interval (ISI) of 0.75 seconds. Pilot studies suggested that using a delay period shorter than this caused participants to feel that they did not have enough time to process the duration, especially the short durations, making the condition seem more like a reaction time experiment -

qualitatively different than when the delay period was 0.75s or larger. The second and third delays roved with the target times. Delay two, averaging 2.33s, was composed of five intervals which ranged from 1.60s to 3.20s (1.60s, 1.90s, 2.26s, 2.69s, and 3.20s). When the five delay periods were summed with their corresponding millisecond target the sum corresponded to one of the supra-second intervals, i.e. 0.40s target summed with 1.60s delay equaled 2.00s; 0.48 summed with 1.90 equaled 2.38, etc (Table 1). This allowed us to compare the estimation in ms target times with delay to matched ss target times. Delay three was twice the length of delay two (3.20s, 3.80s, 4.52s, 5.38s, 6.40s) and was included to obtain a fuller mapping of how delays altered temporal reproduction.

Average						
0.58s	Roving Referents (ms)	0.40s	0.48s	0.57s	0.67s	0.80s
2.33s	Delay 2	1.60s	1.90s	2.26s	2.69s	3.20s
2.91s	Roving Referents (ss)	2.00s	2.38s	2.83s	3.36s	4.00s

*Table 1: Each target time category is represented by five intervals, or referents. The delay periods were chosen to ‘rove’ with the roving referents such that the sum of the delay period and the ms target time equal the ss target time. The second row shows the ms target referents used (Roving Referents (ms)), the third row shows the delay period (Delay 2), and the bottom row shows the ss target referents used (Roving Referents (ss)).*

Participants were run individually in isolated testing rooms. Instructions were read aloud by an administrator in which participants were informed that the goal of the task was to learn time intervals. They were instructed not to count during the trial and no feedback was given. A trial consisted of the presentation of the target duration with a blue circle 1.5 inches in diameter

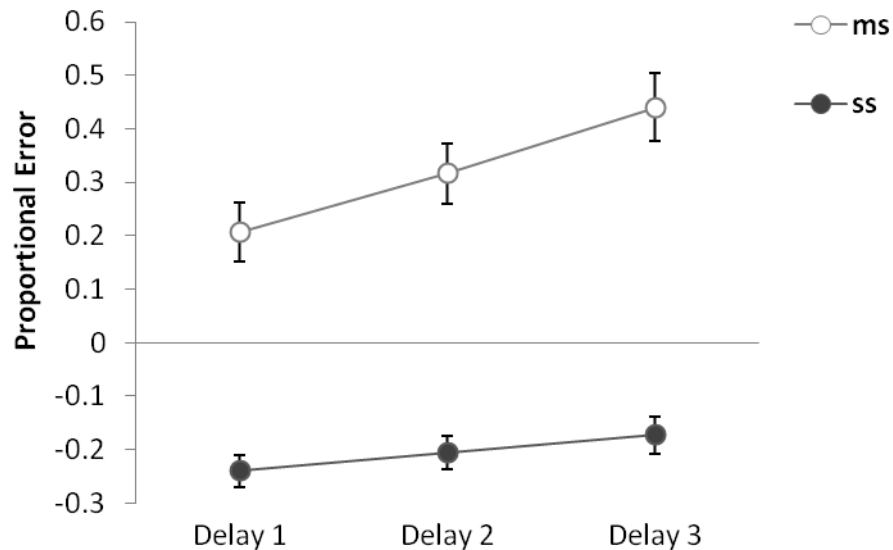
on a black background. This was followed by a delay period during which the video screen was black. The reproduction phase was initiated by the appearance of a red square 1.5 inches wide on a black background. A different shape was chosen for presentation and reproduction trials to prevent participants from becoming confused about which phase of the trial they were in. Any effects this may have on temporal estimates would be consistent across all trials for all participants. Participants were required to press the space bar to mark the end of the target duration and terminate the trial causing the red square to disappear. A variable inter-trial interval (ITI) ranging from 3s to 5 s with a mean of 4s separated trials.

There were 10 trials for each delay period (delays one, two, and three) and average target duration (0.58s and 2.91s) for a total of 60 trials which took participants approximately 15 minutes to complete. There were two presentations of each of the five roving referents for both average target durations. The trials were intermixed randomly by the computer for each participant such that no more than 2 trials of the same type (a single referent-delay combination) would be displayed consecutively.

*Data Analysis:* The proportional error (PE) was calculated by subtracting the presented target time from the estimate and dividing this value by the presented target time ((estimate – target) / target) for each trial for each participant. Data points that fell beyond two standard deviations of the mean PE for each condition for each participant were removed with the assumption that estimates extending beyond this distribution represented non-task related distraction. This resulted in removing 3.4% of the data. The PE is a normalized measure of accuracy that allowed us to assess changes in estimates across delay periods and the roving target times. There were not enough trials for each roving referent target time to analyze them separately. The coefficient of variation (CV) was calculated by dividing the standard deviation

by the mean for each condition using the trimmed data set. The CV is a normalized measure of temporal variability. PEs and CVs were averaged for each condition for each participant and submitted to separate 3 (delay period) x 2 (target time) repeated measures ANOVAs.

## Results



*Figure 2: Proportional errors (PEs) for each delay length for Experiment 1. PEs greater than zero signify that estimates were larger than the target time, PEs less than zero signify that estimates were shorter than the target time, and PEs of zero signify that estimates were the exact target time. The millisecond target (ms) was overestimated with respect to the target time while the supra-second target (ss) was underestimated. Error bars are standard error.*

As seen in Figure 2, PEs for the supra-second targets were significantly shorter than PEs for the millisecond targets,  $F(1,30) = 100.093$ ,  $p < .001$ . As predicted by previous studies the supra-second targets were estimated as shorter than the target time (negative PEs), but the

millisecond targets were estimated as larger than the target time (positive PEs). Unexpectedly, PEs increased with increasing delay,  $F(2,60) = 17.900$ ,  $p < .001$ . The increase in estimates with increasing delay is opposite to our expectation. There was also a significant target time by delay period interaction,  $F(2,60) = 6.433$ ,  $p < .01$ . We had expected that this interaction would be caused by a decrease in the PE for the ms time when coupled with delay periods as opposed to the typical overestimation found for ms target times when there is no delay. However, as Figure 2 demonstrates PEs of the millisecond target increased more steeply with increasing delay than the PEs of the supra-second target. The delay did not cause estimates to shorten. With CVs as the dependent measure, there were no significant effects, Target:  $F(1,30) < 1$ ; Delay:  $F(2,60) = 1.569$ ,  $p > 0.2$ . The scalar property was not violated across target times. Table 2 has mean CVs for the 3 delay lengths and two target times.

	Experiment 1		Experiment 2	
	ms Target	ss Target	ms Target	ss Target
Delay 1	0.29 (0.11)	0.31 (0.08)	0.27 (0.06)	0.30 (0.11)
Delay 2	0.32 (0.08)	0.32 (0.14)	0.29 (0.08)	0.29 (0.09)
Delay 3	0.32 (0.11)	0.31 (0.09)	0.33 (0.13)	0.33 (0.09)

*Table 2: CVs (standard deviations) for the two target times across the three delays for Experiments 1 and 2. For Experiment 1: delay 1 is 0.75s, the delay 2 average is 2.33s, and the delay 3 average is 4.66s. For Experiment 2: delay 1 is the same (0.75s). Delays 2 and 3 differ depending on the target time. For the millisecond target the delay 2 average is 1.17s and the delay 3 average is 2.33s. For the supra-second target the delay 2 average is 5.83s and the delay 3 average is 11.66s. There were no significant effects of CV in Experiment 1. There was a main effect of delay in Experiment 2.*

## Discussion

We expected that reproductions would decrease with increasing delays. Based on the results of prior studies that demonstrated subjective shortening we had proposed that the inclusion of the delay would cause temporal information to decay and that the decay would increase (shortening target times) with increasing delay. However, we found that reproductions increased with increasing delays.

We hypothesized that ss target times are typically underestimated because early temporal information decayed while the duration still accumulated. Using delays we created a situation in which ms target times needed to be maintained over a duration similar to the early information of the ss target time. The fact that estimates of the ms target time did not shift to underestimation with the delay could be for one of two reasons. If the underestimation found with ss timing was not due to memory processes, then we would not expect the delay to affect the ms target. For example, participants might be motivated to leave the experiment and provide estimates that are slightly shorter than the actual target time when the targets are longer (i.e. ss).

The second reason could be that the maintenance memory process that occurs during accumulation is not the same as the maintenance process that occurs after temporal information is stored in memory. It is well documented that temporal information must be attended to in order to accumulate (Brown, 1997, 2008; Macar, et al., 1994). It is logical to assume that the attended to accumulated information must also be maintained as additional temporal information continues to accumulate. Temporal stimuli differ from most other types of stimuli studied in psychology because temporal stimuli are dynamic in the sense that they change over time (Spetch & Rusack, 1992). It is possible to know how long a duration is only after it has elapsed. The process by which temporal information is maintained during accumulation may be different

from the process of maintaining information about the memory representation of an elapsed duration.

The other unexpected result from our data was the fact that estimates increased with increasing delay instead of the predicted decrease. The phenomenon of subjective shortening was interpreted as a decay of temporal information over time. Decay is just one possibility for the degradation of a memory representation. Interference from other information can also degrade memory representations (Jonides, et al, 2008). The delay period, being an interval of time, could be interfering with the memory representation of the target times. This distortion could be conceived of within SET. If the accumulation of temporal information continues during the delay period, the result would be a lengthening in the memory of the target time. We assumed that participants understand that the delay period is not a continuation of the target time nor is it a relevant target time for the task and this would explain why estimates are shorter than the sum of the delay and the target time. This implies that temporal information that accumulates during the delay does so at a reduced rate and this smaller accumulation is added to the memory representation. Instead of finding that temporal information decays from the memory store, it appears to *leak* into it during the delay period.

To account for the target time by delay interaction we propose that the *leakage* from the accumulator to the memory store during the delay period occurs with a constant probability. If this were true the amount of interference from the delay would be a constant proportion of the delay length. As the delay increases, the interference would also increase. However, the increase in estimates will represent a larger percentage of the ms target time than of the ss target time. The relative size of the effect would be greater for a shorter target time, as we find in our data.

To test this assumption further we designed a second experiment with the goal of making



the effect of the delay the same across target times. If the leakage is a constant proportion of the delay length, and if the delay length were proportionally related to the target time, the effect size should be the same for all target durations. In Experiment 1 the delay periods were the same for both ms and supra-second target times. For Experiment 2 we made the proportional relationship between the delay period and the target time the same. Effectively, the second delay is twice the target time and the third delay is four times the target time.

## Experiment 2

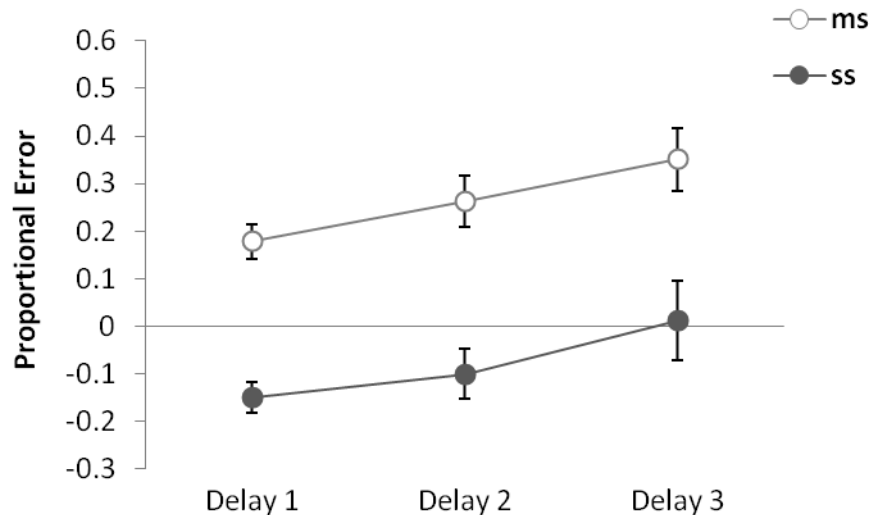
### Methods

*Participants:* Thirty participants (26 female and 4 male) were recruited from Barnard College. Their ages ranged from 18 – 29 years with a mean age of 20. Participants were naïve to the purpose of the experiment and were compensated with course credit.

*Materials and Procedure:* The presentation software and hardware and the target durations were identical to Experiment 1. However, the lengths of the delay conditions differed. Delay 1 was identical for both targets (0.75 s) but delays 2 and 3 increased with the target duration. Delay 2 was twice the target duration. For the millisecond targets the delays were 0.80s, 0.95s, 1.13s, 1.35s, 1.60s (1.17s average) and for the supra-second targets they were 4.0s, 4.76s, 5.66s, 6.73s, 8.00s (5.83s average). Delay 3 was four times the target duration. For the for millisecond targets they were 1.60s, 1.89s, 2.26s, 2.69s, 3.20s (2.33s average) and for the supra-second targets they were 8.00s, 9.51s, 11.31s, 13.45s, 16.00s (11.66s average). All other elements of the procedure were the same as in Experiment 1.

*Data Analysis:* Data points that fell more than two standard deviations from the average PE for each condition were assumed to represent non-timing based distractions that occurred

during the task and were removed. This resulted in removing 3.2% of the data. The same measures (PEs and CVs) were calculated identically to Experiment 1 using the trimmed data set and submitted to separate 3 (delay length) x 2 (target time) repeated measures ANOVAs.



*Figure 3: Proportional errors for each delay length in Experiment 2. The millisecond target (ms) was overestimated and the supra-second (ss) was underestimated.*

## Results

PEs increased significantly with increasing delay,  $F(2,58) = 7.958$ ,  $p < .01$ , but as can be seen in Figure 3, there was not a significant target by delay interaction,  $F(2,58) < 1$ . PEs for the supra-second targets were significantly lower than PEs for the millisecond targets,  $F(1,29) = 44.344$ ,  $p < .001$ . Participants underestimated the supra-second target time and overestimated the millisecond target time. There was a main effect of variability,  $F(2,58) = 5.830$ ,  $p < .01$ .

However, variability was significantly larger for the third delay (delays 1 & 2:  $F(1,29) < 1$ ; delays

2 & 3:  $F(1,29) = 4.78, p < .05$ ; delays 1 & 3:  $F(1,29) = 10.06, p < .01$ .. CVs did not differ across target times,  $F(1,29) < 1$ , manifesting the scalar property.

## Discussion

We hypothesized that the increase in estimates with increasing delay period was due to interference of the delay period on the memory representation of the target time. Specifically, we speculated that a constant proportion of the delay length leaked into the memory store. In Experiment 1, the delay periods were the same length for the ms and ss target times leading to a larger effect in the ms targets than in the ss targets. By making the lengths of the delay periods proportional to the target time in Experiment 2, the effect of the delay was identical for the two target ranges.

## General Discussion

Delays have often been used in non-timing studies to assess how memories change over time (Jonides, et al, 2008). Alterations in memory over time can occur due to decay of the relevant information, but also due to interference by related, but irrelevant information. Our data demonstrates that delays, because they are a duration of time, interfere with existing memories of duration. Reproductions increased systematically with increasing delay period such that the effect size was directly related to the proportion of delay length to the target time.

Estimates did not increase by the entire length of the delay period. Participants did not add the entire length of the delay period to the length of the target time. Instead estimates increased by a smaller amount that was directly related to the length of the delay period and independent of the target time. Our paradigm required participants to attend to time, and this

attention continued with a constant probability during the delay period causing temporal information to accumulate. Interference occurred because the additional accumulation effectively *leaked* into the memory store and merged with the existing memory representation creating a longer memory and a longer estimate. This implies that temporal information from a non-relevant duration (i.e. the delay period) can interfere with representations in the memory store.

Our results corroborate previous studies demonstrating that ms target times tend to be overestimated with respect to the target and ss target times tend to be underestimated (Jones, et al., 2004; Koch et al., 2007). We had originally proposed that these differences derive from differences in how these target times are processed, specifically, that more cognitive processes are necessary for accurate ss timing. However, since the delay did not cause ms target times to be underestimated similarly to ss target times, we can assume that the maintenance that occurs during a delay is not a cause of the underestimation of ss target times. The fact that both target times were affected similarly by the delay suggests that working memory processes affect the two target ranges similarly.

An alternative explanation for the overestimation of ms target times and the underestimation of ss target times in our study comes from Vierordt's law (Lejeune & Wearden, 2009). The original studies done by Vierordt and his students demonstrated that when a group of target times are learned, the shorter ones will be overestimated and the longer ones will be underestimated (Lejeune & Wearden, 2009). These studies tested durations from a few hundred ms to 60 minutes, suggesting that these behaviors were not dependent on range effects or automatic and cognitive timing, but instead were a result of learning multiple times in a single task.

Our results are the opposite of studies that demonstrated subjective shortening with delays. Subjective shortening implies that the memory representation of the standard decayed over time. Evidence for subjective shortening originally derived from work with pigeons (Spetch & Wilkie, 1983; Spetch & Rusack, 1992). In the delayed match-to-sample task (non-human) subjects learned to associate a short target with one lit key and a long target with another lit key and in testing trials they had to choose the correct lit key associated with the presented target time. The delay period between target and key light presentation was fixed during training, but changed during testing (Spetch & Wilkie, 1983). Subjects were prone to respond by pecking at the lit key associated with the short target time as the delay between interval presentation and key light presentation increased. An alternative interpretation of these results is that the subject was confused by the differences between training and testing (e.g. Zentall, 2006). This claim is supported by the fact that when the inter-trial interval (ITI) is altered instead of the delay period, subjects exhibit behavior that supports subjective shortening (Spetch & Rusak, 1992). Response accuracy increases when the ITI is shorter and decreases when the ITI is longer, even though the ITI occurs prior to presentation of the duration. Zentall (2006) argues that changes in the delay length make it difficult for the subject to differentiate when a new trial has begun. In fact, Dorrance, et al. (2000) completely obliterated the effects of subjective shortening when the ITI and delay were sufficiently distinct.

However, we assumed that human participants were able to understand the task and, as discussed previously, studies using human participants were consistent with the phenomenon of subjective shortening (Wearden & Ferrara, 1993; Wearden, et al, 2002). Using discrimination timing tasks Wearden & colleagues demonstrated that participants were more accurate at responding long when a comparator duration was longer than the presented standard (Wearden &

Ferrara, 1993; Wearden, et al, 2002). While the specific target times and delay periods used differed from ours, the ranges of the durations used do overlap.

A major difference between the studies is the task type used. Wearden and colleagues used a discrimination task, whereas we used a reproduction task. It has been argued that the tasks measure similar processes leaving no clear theoretical reason why the two tasks would yield directly conflicting results (Ivry & Hazeltine, 1995). However, the difference between these two tasks is similar to the difference in non-temporal memory tasks that investigate recall and recognition processes. Typically, recall tasks require participants to generate representations of the target from memory, such as the words from a word list. In contrast, recognition tasks present cues and the participant must decide if they are the same or different from a previous list. Reproduction tasks require the generation of the memory representation for recall, whereas discrimination tasks require a recognition to be made.

Additionally, Bangert, et al. (2011) suggested that different timing tasks may be subserved by different mechanisms. Jones, et al. (2008) also demonstrated differences in processing different types of timing tasks. Furthermore, Lewis & Miall's (2003) distinction between automatic and cognitive timing depends not only on the length of the target time being measured, but also on whether movement is used to measure the duration and whether the presentation occurred in a repeated and continuous fashion. Task type appears to be very important in understanding different cognitive and memory processes involved in timing.

In conclusion, we demonstrated that the time period between duration presentation and reproduction affects the memory representation of the target time. When this time period was increased it systematically distorted reproductions by adding a proportion of the delay period to the memory representation. The potential interference of inter-stimulus intervals on time

perception has not previously been addressed. While this interference effect differed from previous work demonstrating that delay periods between duration presentation and reproduction would lead to decay of the memory representation, our results suggest that recognition and recall of temporal information may involve different memory processes as is similar with non-temporal information.

## Chapter 2

### Memory for Multiple Time Intervals

In the previous studies we used a delay embedded in each trial to target the working memory system. In the following set of experiments we shift focus to a typical, albeit inconsistent, finding in temporal memory, Vierordt's law.

Vierordt's law was derived by Karl von Vierordt in 1868 (Lejeune & Wearden, 2009) and simply states that when multiple target times are learned together estimates of the shortest time will be overestimated and estimates of the longest target time will be underestimated. This is essentially what we found with estimates of ms and ss target times in the previous studies. Lejeune & Wearden (2009) in their comprehensive review of Vierordt's law report that Vierordt verified this pattern across a range of target times from ms to minutes and used a variety of timing tasks including reproduction, discrimination, and verbal estimation. While comprehensive in these respects, a major drawback to the studies is that they were completed by Vierordt and his students, instead of a random sample of the population.

Despite Vierordt's studies spanning target times and timing methods, it remains unclear if Vierordt's law is a robust effect. Some published studies have explicitly discussed the existence of Vierordt's-like patterns using target times ranging from 4s to 80min (Grondin & Plourde, 2007; Yarmey, 2000). Others have described the Vierordt-like effects with target times in the supra-second range in studies with patient populations such as Parkinson's disease (Malapani, Deweer, Gibbon, 2002) or Schizophrenia (Van Volkinburg, et al., unpublished). However, there is also data in which Vierordt's law does not occur. For example, Rakitin, Stern, & Malapani (2005) tested participants on 6s and 17s target times which did not show overestimation of the 6s



target time or underestimation of the 17s target time. It becomes difficult assessing when Vierordt's law does and does not occur in the literature because it is not always noted, even when it does occur. For example, in their discussion of Vierordt's law, Lejeune & Wearden (2009) report that some of their previously published data demonstrated Vierordt-like patterns with target times ranging from ms to supra-seconds, but the original publications did not discuss Vierordt's law (Penton-Voak, et al., 1996; Wearden et al, 1998).

One reason Vierordt's law isn't always discussed is that it contradicts evidence that estimates of a duration are veridical with real time (Wearden, 2003). Evidence demonstrating the veridical nature of estimates is in direct conflict with evidence of Vierordt's law since estimates of the shortest target time are underestimated and estimates of the longest target time are overestimated (Jones & Wearden, 2003; Wearden & Lejeune, 2008; Wearden & Jones, 2007).

Unpublished data collected in our lab demonstrated a Vierordt-like pattern when target times 1.1s, 2.2s, and 3.3s were initially learned (Gooch et al, submitted). However, at recall one hour later, estimates for the larger two target times (2.2s & 3.3s) increased so much that the Vierordt-like effect was eliminated and the larger target times were overestimated, instead of underestimated.

The current study had two purposes. The first phase, training, consisted of the learning and reproduction of multiple time intervals which should be sufficient to observe Vierordt's law. Since Vierordt's law was not observed with target times of 6s & 17s (Rakitin, et al., 2005) but was present in Gooch's experiment when intervals were less than 3.3 s we decided to use target times that spanned the two studies. Therefore we used 3s, 6.5s, and 10s.

Another unique element we incorporated into this study was intermixing the learning and subsequent recall of target times. Previous research typically presents a block of target time *a*

and then a block of target time  $b$ , and so on. In later recall, participants are required to reproduce the target time presented in block 1, then the target time in block 2, etc. By associating distinct shapes with specific target times, we were able to elicit the memory representations of different durations by presenting the associated shape. Therefore, both learning and subsequent recall for different durations could be intermixed.

The second purpose of the current study was to assess the reliability of the increase in estimates, especially of the longer two target times, with a delay between initial learning and a second reproduction. Instead of an hour delay, participants were given a delay period of a few minutes. During the delay, participants performed a task designed to control behavior with the intention of reducing intersubject variability. Different states of arousal are known to affect timing performance (Fetterman & Killeen, 1995; Meck, 1996; Penton-Voak, et al., 1996; Treisman, et al., 1990; Wearden & Penton-Voak, 1995) and we didn't want participants distracted by potentially arousing events in their lives (e.g. upcoming exams) or by sleeping (during an hour break). Therefore we used a listening response task shown to not disrupt timing behaviors (Wearden, 1999, and in pilot data of our own). After the delay task we elicited estimates again by presenting shapes for participants to reproduce the associated durations. We expected either that the brief delay would replicate the increase in estimates found in Gooch's study or that the brief delay would not be long enough to remove the Vierordt-like effects found at training.

## Experiment 1

### Methods

*Participants:* Twenty-seven participants (26 females, 1 male) contributed to the final data for analysis. Their average age (and standard deviation) was 20.85 (4.68). Participants were naïve to the purpose of the experiment and were compensated for their time with course credit.

*Procedure and Materials:* Experimental events were presented using the psychophysical toolbox (Brainard, 1997; Pelli, 1997) compiled into a Matlab® (2007a, The MathWorks, Natick, MA) executable program presented on Microsoft XP® OS PC's. Responses were recorded by key press on a standard keyboard.

There were three separate blocks of trials. The first block was a timing reproduction task (training phase), the second block was a non-temporal listening task (distracter), and the third block was a timing reproduction test task (testing phase), which contained no stimulus presentation. The target times for the two timing phases were 3s, 6.5s, and 10s. Three different shapes corresponded to the three different target times. A blue circle, 1.5 inches in diameter, corresponded with 3s. A green square, 1.5 inches wide, corresponded with 6.5s. A red isosceles triangle, 1.5 inches high corresponded with 10s. Participants were run individually in isolated testing rooms. Instructions were read aloud by an administrator in which participants were informed that the goal of the task was to learn time intervals that corresponded with a specific shape. They were instructed not to count and no feedback was given.

A training trial began with the presentation of one of the colored shapes on a black background for the corresponding target time. After the target time elapsed, the shape was followed by a black screen for a variable inter-stimulus interval ranging from 0.3 to 0.7 seconds with a mean of 0.5s. Then the same shape reappeared, but in white and remained on the screen

indefinitely until participants recorded their duration estimate by pressing the space bar to make the white shape disappear when they believed that the white shape had been present for the same duration as the preceding colored shape (Fig. A). The lack of color for the reproduction phase helped make it clear to the participant which phase of the trial they were in. A variable inter-trial interval (ITI) ranging from 0.5s to 1.5s with a mean of 1s separated trials. There were ten trials for each target time for a total of 30 trials which were intermixed pseudo-randomly by the computer for each participant such that no more than 2 trials of the same type could be displayed consecutively. Reproductions were recorded to the nearest millisecond.

In the distracter task participants were instructed to press the space bar every time they heard the word “bat”. A vocal recording of the word “bat” was presented 40 times with a variable inter-trial interval ranging from 0.5s to 5s with a mean of 2.75s. The length of the task depended on the speed of participants reaction times. We anticipated that it would take subjects approximately 5 minutes assuming they completed the task quickly and accurately.

In the testing phase, participants reproduced the target times that they had learned during the training phase. Participants were shown white shapes and required to press the spacebar to terminate the duration based on their knowledge of the target times presented during the training phase. A variable ITI of 0.5s to 1.5s with a mean of 0.75s separated the presentation of the white shapes. As before there were ten trials for each target time and trials were pseudo-randomized by the computer for each participant such that no more than 2 trials of the same type would be displayed consecutively.

*Data Analysis:* Two measures of temporal perception were calculated. Proportional error (PE) was calculated as a measure of temporal accuracy by subtracting the target time from the estimate and dividing this difference by the target time. Since this provides a measure of

accuracy with respect to real time it allows for the easy assessment of Vierordt's law. Negative PEs show underestimation, positive PEs show overestimation, and PEs at zero show estimates veridical with real time. The coefficient of variation (CV) of estimates was used to assess the precision of temporal estimates. CVs were calculated by dividing the standard deviation by the mean of the estimates for each duration in each condition for every participant. Data points that fell beyond two standard deviations of the mean for each condition for each participant were removed with the assumption that estimates extending beyond this distribution did not represent temporal processes, but instead represented non-task related distraction. This resulted in removing 3.89% of the data. PEs and CVs were submitted to separate 2 (phase) x 3 (target duration) repeated measures ANOVAs.

## Results

As would be expected from Figure 4, there was a target time by phase interaction for PEs ( $F(2,52)=8.01, p < .01$ ). PEs decreased significantly with increasing target time at training ( $F(2,52)=4.15, p < .05$ ). There was also a main effect of target time at testing ( $F(2,52)=5.224, p < .05$ ). There was no significant difference between the PEs of the 6.5s and 10s targets, but the PE for the 3s target was significantly smaller than the PEs for the other two targets (3s, 6.5s:  $F(1,26)=8.215, p < .01$ ; 3s, 10s:  $F(1,26)=5.975, p < .05$ ; 6.5s, 10s:  $F(1,26) < 1$ ). Additionally, the PEs for the 3s target did not differ across timing phases ( $F(1,26) < 1$ ). PEs for the testing phase were larger than PEs for the training phase ( $F(1,26)=6.904, p < .05$ ). There were no other significant effects.

For precision, as measured by the CV, there was a main effect of target time (main effect:  $F(2,52)=7.12, p < .01$ ). The CV was larger for the 3s target than for the other two targets (; 3s,

6.5s:  $F(1,26)=19.73$ ,  $p < .001$ ; 3s, 10s:  $F(1,26)=6.12$ ,  $p < .05$ ; 6.5s, 10s:  $F(1,26)=1.44$ ,  $p > .2$ ).

However, there was no effect of phase ( $F(1,26) < 1$ ; Table 3). Variability increased proportionally with increasing estimates at testing.

On average, the distracter task lasted 6.82 minutes with a median time of 5.47 minutes (standard deviation of 4.74 minutes). Some participants did not respond immediately after hearing the word “bat” so the distracter task ranged from 2.18 minutes to 18.60 minutes. Since the length of the distracter task varied greatly, we wanted to see if changes in estimates were due to the increasing length of the distracter period. We correlated PEs for each target time and phase with the total length of the distracter task. We correlated CVs for each target time collapsed across phases since there were no significant differences of timing phase. PEs and CVs did not correlate with the length of the distracter task (Table 4).

Experiment 1	Training	Testing
3s	0.23 (0.10)	0.22 (0.18)
6.5s	0.16 (0.06)	0.16 (0.07)
10s	0.17 (0.07)	0.20 (0.22)
Experiment 2	Training	Testing
3s	0.24(0.10)	0.25(0.15)
6.5s	0.19(0.09)	0.20(0.09)
10s	0.18(0.07)	0.19(0.12)

*Table 3: The coefficient of variation (CVs) with standard deviations in parentheses for each of the three target times at the two timing phases for Experiments 1 and 2.*

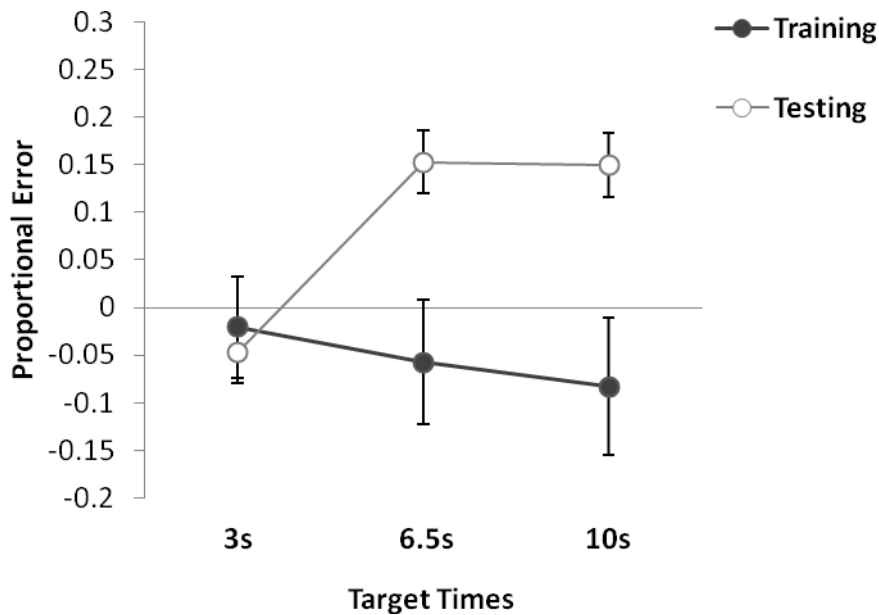


Figure 4: Proportional errors for each of the three the target times for training and testing in Experiment 1. Error bars are standard error.

	PE		CV	
	r	Significance	r	Significance
3s	-0.11	0.57	-0.09	0.65
6.5s	-0.16	0.43	0.15	0.45
10s	-0.08	0.69	0.00	0.99
Ave.	0.05	0.83	0.17	0.41

Table 4: Pearson correlation coefficients for each target time at testing for PEs and CVs for correlations conducted with the length of the distracter task (Experiment 1). 'Ave.' is the average of the 3 target times correlated with the length of the distracter task. There were no significant correlations.

## Discussion

The primary goal of the training phase was to explore whether Vierordt's law would be observed across a duration range that had previously produced ambiguous results (Gooch, et al., unpublished; Rakitin, et al, 2005). The PEs for all three target times were negative at training, meaning that all three were underestimated with respect to the target time. This is commonly found with target times in the supra-second range and is consistent with previous studies (Jones, et al., 2004; Koch, et al., 2007; Lewis & Miall, 2003a; see also chapter 2).

However, this does not demonstrate the archetypical pattern of Vierordt's law in which estimates of the shortest target time are overestimated. That said, if the shortest target time is overestimated and the longest target time is underestimated, we would expect a negative slope when plotted against increasing target times, which we see in Figure 4. Our data demonstrate a 'migration' effect in which the memory representations of the three target times tend towards each other, even though it does not demonstrate a true form of Vierordt's law. Interestingly, other work in which Vierordt's law is discussed does not always show it in the *true* form either. For example, Yarmey's (2000) study included target times ranging from 4s to 80min, all of which were found to be underestimated with respect to the target time. But because less underestimation was found for the shortest target time than the longest target time, it was presented as an example of Vierordt's law. Similar claims were made by Lejeune & Wearden, (2009) with target times in the supra-second range.

Migration effects are also discussed in work with patient populations such as Parkinson's (Malapani, Deweer, Gibbon, 2002) and Schizophrenia (Van Volkinburg, et al., unpublished). The migratory effect is discussed in comparison to controls, rather than in comparison to real time, with patients more prone to overestimating the shortest time and underestimating the



longest time than controls. Shea-Brown, et al. (2006) argue that this effect is caused by an internal clock that starts slow and increases speed as the duration progresses causing shorter target times to be overestimated and longer target times to be underestimated. This effect is diminished when Parkinson's patients are on medication suggesting that it is a disease related impairment (Malapani, et al., 2002; Shea-Brown, et al., 2006). The migration effect in our study and in other studies using healthy students as participants are likely caused by a different mechanism.

Both Goldstone, et al (1957) and Bobko, et al (1977) argue that the perception of stimuli, in general, is affected by the context in which it is presented. In our study, the *context* is defined by the presented target times. When finding the migration of estimates in their data, Bobko, et al, (1977) argued that the estimates tended toward a central tendency specified by the stimuli used. The 1977 study was a follow up to a previously published study and combined, the two demonstrated that the range of target times tested dictated the time order error in estimates (Schiffman & Bobko, 1974). The long target times used in Schiffman & Bobko (1974) were the short target times in Bobko, et al (1977). They were underestimated in the first study (being the long targets) and in the second study they were overestimated (being the short targets). The authors suggest that participants spontaneously grouped the stimuli into a category and the creation of the category caused estimates to drift towards the central representation of the category.

This sentiment is echoed in recent work in the visual domain. Brady & Alvarez (2011) demonstrate that when participants are presented with an array of visual stimuli that can be categorized via a perceptual feature (e.g. color), participants estimate the size of an individual stimulus as larger if the other stimuli presented in that category tend to be larger, regardless of

the size of the test item. Likewise, estimates tend to be smaller if the test stimulus is presented in a grouping in which the other stimuli tend to be smaller. Brady & Alvarez (2011) argue that the categorical grouping of perceptual items facilitates learning of the large volume of stimuli we are confronted with in our environment. Rather than remembering each detail of each item, we can remember a category of items and their relationships to each other. We propose that similar mechanisms are at work in our reproduction task, and that, participants spontaneously categorize the stimuli together. As in Bobko's studies (1974, 1977), estimates of the short and long target times tend towards each other because it is the combination of stimuli presented that define the parameters of the category.

Interestingly, this effect does not persist at later recall, even after only a few minutes. At testing, the PE for the 3s target time remained the same but the PEs for the 6.5s and 10s target times increased, essentially the opposite of Vierordt-like effects. This suggests that the migratory pattern of estimates is a relatively short-term effect. In our previous two studies we demonstrated that delays interfere with timing and cause estimates to increase. However, that occurred with delays immediately preceding each individual reproduction, whereas, the delay in this experiment followed a set of presentation and reproduction trials. Furthermore, there was large variability in the length of the distracter task (i.e. the delay period) across participants but the difference in the length did not correlate with increases in estimates. This suggests that the delay itself was not the cause of increased estimates.

The PEs for the 6.5 and 10s durations did not differ (see Fig. 4) which means that they increased by the same proportion of the target time. In contrast, the 3s target remained the same across timing phases. This effect replicated the results of Gooch, et al, (in revision), using a different range of target times. Additionally, the shortest time in our study (3s) is very similar to

the longest target time in their study (3.3s) suggesting that the result is not dependent on the absolute target time but on the relationship of the group of learned target times. Also, Gooch, et al., (in revision), showed that increases at testing were larger when a single target time was learned and recalled later than when multiple target times were learned and recalled later. Apparently, learning the target times together helps to attenuate the lengthening effect. This is likely because the target times were categorized as a group when learned.

A final unique and unexpected result of this data was the fact that the CV did not change between phases, even though estimates for 2 of the 3 target times did. One of the fundamental aspects of temporal perception, typically referred to as the scalar property (of precision), is that as estimates increase, the variance increases proportionally such that the standard deviation divided by the mean is constant (Gibbon, 1977; Gibbon, et al., 1984; Wearden & Lejeune, 2008). Even though estimates increased at testing, variability increased proportionally such that the scalar property was not violated. This implies that variability in timing is not something encoded when durations are learned. Gibbon (1992) had proposed that scalar variability arises from memory mechanisms and it had been assumed that it was most likely from encoding processes (Gibbon, et al., 1984; Meck, 1996). However, our data demonstrates that scalar variability arises from retrieval processes. Previous studies using a similar paradigm (e.g. Gooch, et al, revision; Rakitin, et al, 2005) found decreases in variability between training and subsequent testing. However, they also included feedback during training and it has been documented that temporal variability decreases with feedback (Wearden & Lejeune, 2008).

In contrast, Malapani, et al., (2002) demonstrated that patients with Parkinson's disease, when completing training and testing phases of a timing task, did not violate the scalar property during testing when on medication, but did violate it during testing when off medication. Of

course, the implication is that when patients are on medication, their timing performance is overall improved, however, the scalar property held at testing when the patients were on medication *regardless* of medicated state during training. This result combined with ours provides strong evidence the scalar variability arises from retrieval processes.

Typically, we anticipate that there will be a null result of the CV across target times demonstrating that the scalar property was not violated, however, our data showed a larger CV for the 3s target time than for the other two. Similar results have been found before and it has been argued that this is an effect of the reproduction task because the motor movement for the response adds non-scalar variability (Wearden, 2003; Wearden & Lejeune, 2008).

There is evidence in our data from both training and testing phases that when temporal stimuli are presented together they are spontaneously learned as a category. For the following experiment, we wanted to see if we could change the memory representations of the group of items simply by perturbing the memory representation of a single item. In order to do this we incorporated a retraining phase into the experiment. After participants learned the 3 target times and were given the delay, they would be presented with a retraining time using the shape associated with the middle cue. One group of participants would be retrained with a shorter time, another group with a longer time, and the control group with the same time presented during training (6.5s). The changes in the memory representation of the all the target times, dependent on retraining a single target time, would help us better understand how the categorical representation of the group of target times is stored in memory.

## Experiment 2

### Methods

*Participants:* Seventy-five undergraduate participants contributed to the final analysis. There were 31 in the control group (30 females, 1 male), 28 in the short retraining group (27 females, 1 male), and 16 (16 females) in the long retraining group. The average ages (and standard deviations) were 18.72 (0.99) for the control group, 18.74 (0.94) for the short retraining group, and 18.53 (0.83) for the long retraining group.

*Procedure:* As in Experiment 1, experimental events were presented using the psychophysical toolbox (Brainard, 1997; Pelli, 1997) compiled into a Matlab® (2007a, The MathWorks, Natick, MA) executable program presented on Microsoft XP® OS PC's. Responses were recorded by key press on a standard keyboard.

There were four separate blocks of trials in this experiment. The first block was a timing reproduction task (training phase), the second block was a non-temporal listening task (distracter), the third block a timing reproduction retraining task (retraining), and the fourth block was a timing reproduction test task (testing phase). The events of training (block 1), of the distracter (block 2), and of testing (block 4) were identical to Experiment 1. The target times for the training and testing phases were also identical to Experiment 1 (3s, 6.5s, 10s).

The events of retraining (block 3) were the same as training (block 1) except that a single target time was presented instead of three. Specifically, the task presented the shape that corresponded to the middle target time (the green square) for ten trials. In the control group, the single target time was identical to the middle target of the training task (6.5s). In the short retraining group, the single target time was 25% shorter than the middle target (4.875s) and in the long retraining group, the single target time was 25% larger than the middle target (8.125s).

The just noticeable difference (JND) of timing values tends to be 15% (Staddon & Higa, 1999). The 25% shift was selected with the intention that the change would be a small enough that participants would still consider it a new value of the original stimulus but large enough to produce a detectible shift in timing. Instructions to participants did not specify that this was a retraining phase and simply stated that they would need to reproduce a single target time during the block.

*Data Analysis:* The same two measures (PEs and CVs) calculated in Experiment 1 were calculated for Experiment 2. Data points that fell beyond two standard deviations of the mean for each condition for each participant were removed with the assumption that estimates extending beyond this distribution did not represent temporal processes, but instead represented non-task related distraction. This resulted in removing 3.75% of the data. PEs and CVs were submitted to separate 2 (timing phase) x 2 (target) repeated measures ANOVAs with group (control, short retraining, or long retraining) as a between subjects factor.

## Results

For PEs, there was no effect of retraining ( $F(2,72) < 1$ ). There was a significant interaction between phase and target ( $F(2,144) = 6.23, p < .01$ ; see Fig. 5). The main effect of target time at training was marginally significant ( $F(2,144) = 2.59, p = .079$ ). There was also a main effect of target time at testing ( $F(2,144) = 4.42, p < .05$ ). There was no significant difference between phases for the 3s target ( $F(1,72) = 1.37, p > .2$ ), but PEs were significantly larger at testing for the 6.5s and 10s targets (6.5:  $F(1,72) = 20.08, p < .001$ ; 10:  $F(1,72) = 15.92, p < .001$ ). Additionally the PEs for 6.5s and 10s targets at testing did not differ ( $F(1,72) < 1$ ).

There were also no effects of retraining on CVs (main effect:  $F(2,72) < 1$ ). As in Experiment 1, CVs did not differ across phases ( $F(1,72) = 2.21, p > .1$ ). However, there was a main effect of target time ( $F(2,144) = 18.56, p < .001$ ). As before, CVs were larger for the 3s target than for the other two target times (3s & 6.5s:  $F(1, 72) = 18.90, p < .001$ ; 3s & 10s:  $F(1,72) = 26.57, p < .001$ ; 6.5 & 10s:  $F(1,72) < 1$ ; Table 3).

As shown in Table 5, participants learned the new target times in retraining fairly accurately. A univariate analysis of the estimates at the 3 levels of retraining demonstrated that the increase in estimates with increasing retraining condition was significant ( $F(2,72) = 58.28, p < .001$ ). Even though participants learned the new intervals, this knowledge did not affect the intervals learned previously.

On average the distracter task lasted 7.96 minutes, with a median time of 5.63 minutes (standard deviation of 7.76 minutes). The length of the task ranged from 2.20 to 48.80 minutes because some participants did not respond immediately to the word 'bat'. We correlated PEs at each level of target time and timing phase with the total length of the distracter task. We correlated CVs at each level of target time collapsed across timing phases since there was not a significant difference between phases. CVs did not correlate with the length of the distracter task (Table 6). PEs for the 3s target times at training did not correlate with the length of the distracter task, however, PEs for the 3s target at testing negatively correlated with the length of the distracter task ( $r(73) = -0.266, p < .05$ ). With increasing length of the distracter task, PEs for the 3s target at testing decreased. PEs at the two other target times at testing did not correlate with the length of the distracter task.

On a post-test questionnaire we asked participants to report if they believed the retraining value was the same, shorter, or longer than the middle value of the previously learned three

target times. (Remember that the shapes were identical across phases which would help participants to relate the two). For the short retraining group (retrained on 4.875s), 24 of the 28 participants (86%) answered this question. Of these, 12 (50%) answered that it was shorter, 9 (37%) answered that it the same, and 3 (13%) replied that the retraining value was longer. For the control group (retrained on 6.5s), 28 of the 31 (90%) participants answered the question. Of these, 9 (32%) answered that it was shorter, 11 (39%) answered that it was the same, and 8 (29%) answered that the retraining value was longer. For the long retraining group (retrained on 8.125s), 10 of the 16 (63%) participants answered the question. It is not clear why this group had a lower response rate. Of these, 1 (10%) answered that the retraining value was shorter, 7 (70%) answered that it was the same, and 2 (20%) answered that it was longer. Overall, participants were not completely aware of any changes in duration between the middle target time when presented during training versus during retraining.

Group	Retraining Values	Estimates			
		Retraining	Test		
			3s	6.5s	10s
Short	4.875	4.80 (0.71)	2.97 (0.63)	6.92 (2.44)	10.51 (2.68)
Same	6.5	6.57 (0.99)	2.94 (1.03)	7.55 (2.58)	11.92 (5.39)
Long	8.125	7.82 (1.11)	3.12 (0.94)	7.99 (2.47)	11.23 (5.18)

*Table 5: Mean estimates (standard deviations) of the target time presented during retraining (Retraining) for the three groups in Experiment 2. Mean estimates (standard deviations) are also presented for each target time from the testing phase (3s Test, 6.5s Test, 10s Test) for each group.*



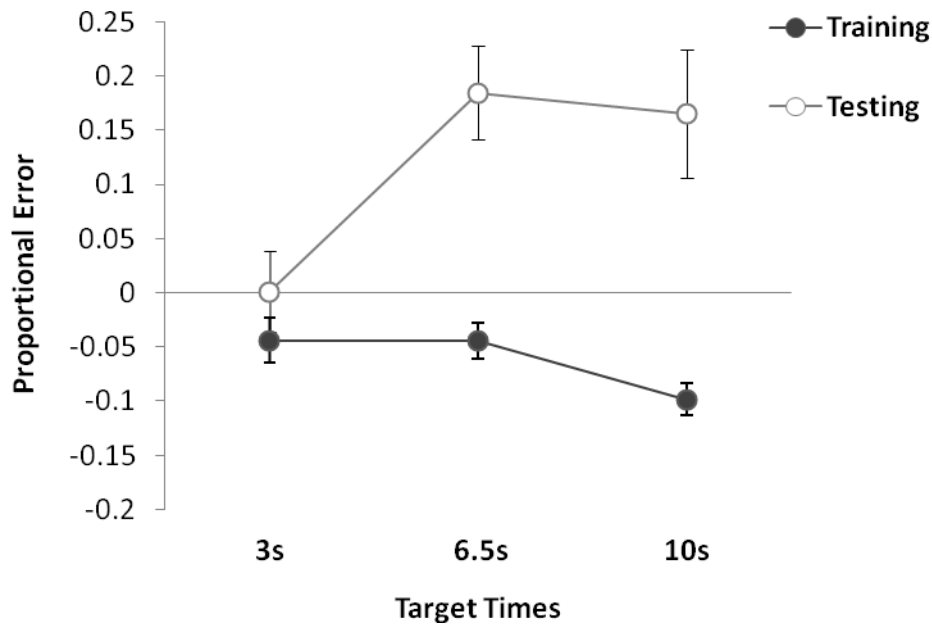


Figure 5: Proportional errors for each of the three target times for the two timing phases in Experiment 2. Error bars are standard error.

	PE		CV	
	r	Significance	r	Significance
3s	-0.27	0.02	0.10	0.4
6.5s	0.01	0.94	0.13	0.27
10s	0.00	0.98	0.13	0.25
Ave.	0.04	0.76	0.14	0.24

Table 6: Pearson correlation coefficients for each target time at testing for both PEs and CVs for correlations conducted with the length of the distracter task (Experiment 2). 'Ave.' is the average of the 3 target times correlated with the length of the distracter task. Only the 3s PE was found to significantly correlate with distracter task length.

## Discussion

The main goal of this study was to show how changing the memory representation of one of the three target times presented at training would change the representation of the other times. Instead, we found that retraining did not alter estimates at testing. The migratory effect replicated at training showed that participants created a spontaneous category of the presented target times. Table 5 shows that participants accurately learned the retrained duration, but this knowledge was not applied to later recall when all three target times had to be recalled. Even though the retraining was meant to alter the representation of one of the previously learned target times, by using the same presentation cue, it seems likely that participants kept the memory representations of the three target times separate from the representation of the retraining target time. This unanticipated result suggests that the multiple target times learned together were in fact separated from the retraining interval, even though the cue used to present the target times was identical. Apparently, when target times are categorized as a group, presentation context is very important, more so even, than the cue.

When the retraining item was presented, we simply told participants that they would be shown more iterations of one of the intervals. We kept the retraining values within 25% of the middle target time length because we didn't want the possible differences in the retraining duration to be too obvious. We wanted to prevent participants from recognizing the difference and changing their estimates. This goal was effective according to post-test self-reports. Participants were not completely aware of the change in the target time. The group retrained on the shorter retrain time was the only group in which at least 50% of the participants were correct in stating that the retrain was shorter. The other groups were worse at stating the difference between the retraining duration and the training duration. Even though the presentation cue was

the same, the retraining was presented within the same experimental context, and participants didn't generally notice a difference in duration length, they still kept the categorically learned intervals at training separate from the retrained duration.

This experiment replicated the previous effects on CV. Most importantly, there were no changes in CV between training and testing, even though estimates increased between training and testing, supporting the idea that the scalar property arises from retrieval. As in the previous experiment, all estimates at training were underestimated with respect to the target time. This is typical in estimates of supra-second (ss) target times. In contrast, ms target times are often overestimated (Jones, et al., 2004; Koch, et al., 2007; Lewis & Miall, 2003a; see also chapter 1).

As discussed in the previous chapter, it is likely that estimates of these duration ranges are mediated by different mechanisms. Lewis & Miall (2003a, 2006) theorized that automatic timing occurs when estimating short durations (e.g. ms), when measuring timing with tasks that require temporally based motor movement, and when the target times are presented predictably. In contrast, cognitive timing occurs when estimating longer durations (e.g. supra-seconds), when using tasks in which estimates are not dependent on motor movement, and when the target times are not presented predictably. Estimates are dependent on motor movement in the reproduction task, consistent with a situation that utilizes automatic timing, but with the variable ISI and ITI, target times are not presented predictably. Therefore, we assume that when short target times are presented in the reproduction task automatic timing resources are utilized whereas when longer target times are presented, cognitive timing is utilized. These different types of timing likely contribute to the behavioral differences found between ms target times and ss target times.

When combined, the overestimation of ms target times and the underestimation of ss target times would create an example of *true* Vierordt's law (as seen in chapter 1), but is it truly

Vierordt's? The data in the two previous experiments shows that ss target times can be underestimated and still migrate towards each other. The behavioral differences in ss and ms target times may be independent from the categorical learning found when multiple target times are presented together. We decided to re-run the first experiment using ms target times. We expected that all of the target times would be overestimated (being in the ms range) and that they would migrate towards each other. Specifically, the shortest would be more overestimated, while the longest would be less so. Additionally, we anticipated finding a similar pattern of results at testing. The shortest target time would not change, but the longer two target times would be even longer than at training.

### Experiment 3

#### Methods

*Participants:* Twenty-eight undergraduates participated in this study for course credit. Their average age (and standard deviation) was 18.82 (0.98). Participants were naïve to the purpose of the experiment and were compensated for their time with course credit.

*Procedure and Materials:* The presentation of experimental events was identical to Experiment 1 except that the target times presented and reproduced were different. The three target times were 270, 590, and 900 ms. The blue circle, 1.5 inches in diameter, corresponded with the 270ms target. The green square, 1.5 inches wide, corresponded with the 590ms target, and the red isosceles triangle, 1.5 inches high corresponded with the 900ms target. As with previous experiments, participants were run individually in isolated testing rooms and were instructed not to count. They were informed that the goal of the task was to learn the time intervals that corresponded to each shape. They were given no feedback on their performance.

*Data Analysis:* As in the previous studies, two measures of temporal perception were calculated. Proportional error (PE) was also calculated as a measure of temporal accuracy by subtracting the target duration from the estimate and dividing this difference by the target duration. CVs were calculated by dividing the standard deviation by the mean for each condition for each participant. Data points that fell beyond two standard deviations of the mean for each condition for each participant were removed with the assumption that estimates extending beyond this distribution did not represent temporal estimates, but instead represented non-task related distraction. This resulted in removing 3.92% of the data. PEs and CVs were submitted to separate 2 (timing phase) x 3 (target) repeated measures ANOVAs.

## Results

PEs decreased with increasing target time ( $F(2,54) = 115.51, p < .001$ ). As can be seen in Figure 6, the PE for the longest target time (900ms) was underestimated with respect to the target time (negative PE), while estimates for the other two target times were overestimated. There was also a main effect of test ( $F(1,27) = 39.61, p < .001$ ). PEs were larger for the testing phase than for the training phase. The interaction was not significant ( $F(2,54) < 1$ ).

There was a target time by phase interaction for CVs ( $F(2,54) = 3.78, p < .05$ ). In Figure 7, CVs at testing appear to be similar to each other, whereas the CVs for training decrease with increasing target time. There was a main effect of target time for the training phase ( $F(2,54) = 9.74, p < .001$ ). However, the CV for 270ms was significantly larger than the other two CVs (270ms & 590ms:  $F(1,27) = 9.29, p < .01$ ; 270ms & 900ms:  $14.39, p < .01$ ), but the CVs for 590ms and 900ms did not differ ( $F(1,27) = 1.23, p > .2$ ). In contrast, at testing, there was no significant difference across target times ( $F(2,54) < 1$ ).

On average, the distracter task lasted 3.00 minutes (standard deviation of 1.17 minutes) with a median time of 2.66 minutes. The length of the distracter task ranged from 1.75 minutes to 6.90 minutes. To verify that changes in estimates were not a direct result of the varying length of the distracter period, we correlated PEs and CVs at each level of timing phase with the total length of the distracter task. None of these variables correlated with the length of the distracter task (Table 7).

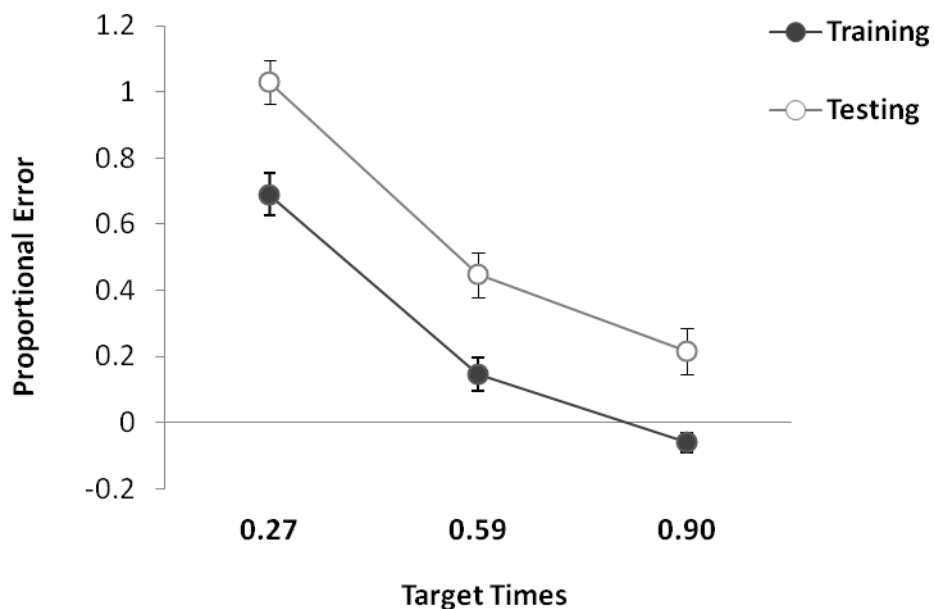


Figure 6: Proportional errors for each of the three target times for the two timing phases in Experiment 3. Error bars are standard error.

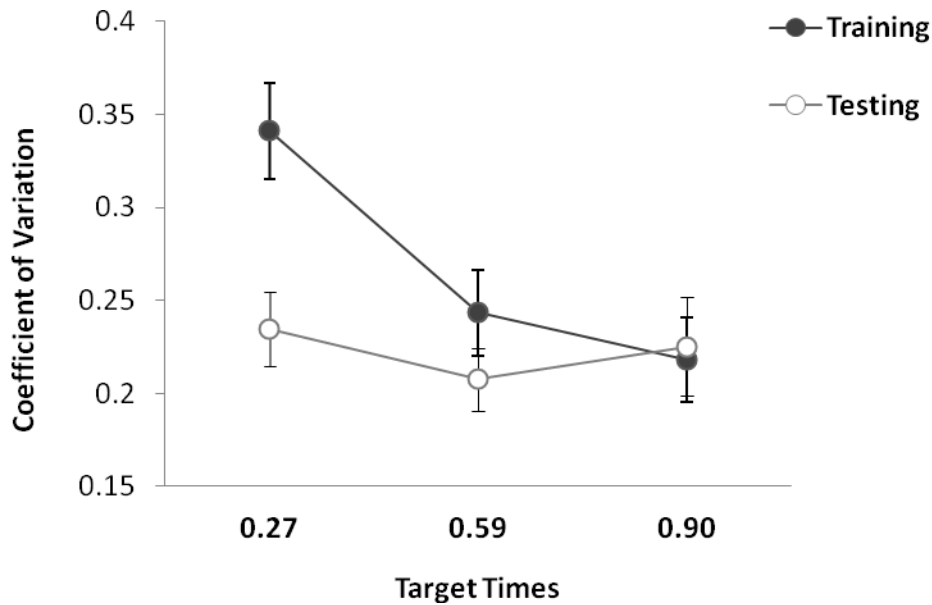


Figure 7: CVs for each of the three target times for the two timing phases in Experiment 3. Error bars are standard error.

	PE		CV	
	r	Significance	r	Significance
0.27s	-0.05	0.79	0.26	0.18
0.59s	-0.26	0.18	0.21	0.29
0.90s	-0.30	0.12	0.09	0.65
Ave.	-0.27	0.17	0.25	0.20

Table 7: Pearson correlation coefficients for each target time at testing for both PEs and CVs for correlations conducted with the length of the distracter task (Experiment 3). 'Ave.' is the average of the 3 target times correlated with the length of the distracter task. There were no significant correlations.

## Discussion

We expected that estimates of the three ms target times would be overestimated at training. However, we were surprised to find that reproductions of the 0.9s target time were underestimated instead. We had expected that all of the ms target times would be overestimated. However, the distinction between ms timing and ss timing may not occur at exactly 1s. As we discussed before, automatic timing is used to make estimates of target times in the ms range and cognitive timing is used to make estimates of target times in the ss range. Automatic timing is used when estimating ms target times and using tasks that require motor movements and in which stimulus presentation is predictable. The implication is that the motor timing system is used to make timing estimates (Lewis & Miall, 2003b). Evidence suggests that the cerebellum is necessary for timing of ms target times (Koch, et al, 2007; Mangels, et al, 1998; Ivry & Keele, 1989; Ivry, 1996). While 0.9s is technically less than a second, and therefore a “ms target time”, because of its proximity to 1s it may behave more like ss target times than ms target times. It may be too long to utilize the motor timing system, and may instead rely on the cognitive system. The implication in cognitive timing is that attention and memory processes are needed to make estimates (Lewis & Miall, 2003b). Further evidence for this is seen in the activation and necessity of these regions to complete estimates of intervals in the ss range (Harrington, et al., 1998; Koch, et al, 2003; Jones, et al. 2004; Mangels et al, 1998; Wiener, et al., 2010).

An alternative argument is that the overestimation of the .27s and .59s target times is an artifact of the reproduction task. Reproduction represents the estimate combined with the response time which would create a noticeably larger reproduction when the target time is sufficiently short. If the response time is constant across trials, it would represent a decreasing percentage of the total reproduction time as the target time increases (Wearden, 2003).



Therefore, shorter target times would be more overestimated because of response time, than longer target times. While this is logical, a drawback to this interpretation is that it does not account for why ss target times are then underestimated, whereas the reliance on separate systems (automatic and cognitive) would account for the behavioral differences.

In the previous studies, the results demonstrated a loss of the migratory effect after a few minutes delay, but in this study that did not happen. Instead, all estimates increased so that the negative slope was maintained, but shifted vertically upwards (Figure 6). One explanation for this difference could be that the delay period was not sufficiently long. For this experiment the delay period lasted, on average, 3.00 minutes, whereas in the previous two experiments it lasted 6.82 minutes and 7.96 minutes. This is likely because in the previous two experiments, many participants failed to respond after immediately hearing the word 'bat' causing the length of the distracter task to be longer. In this study, all participants responded quickly and accurately, shortening the length of the distracter task (i.e. the delay period). It is also possible that this occurred in this experiment and not the others because estimating ms target times served as a prime for making rapid responses to the stimulus in the distracter task. Although the distracter task we used had been shown in previous studies not to interfere with timing processes (Wearden, 1999; pilot data), these studies did not address how timing might influence the distracter task. Future work should compare reaction times in a simple reaction time task when the task is preceded by estimating ms target times or by estimating ss target times.

Finally, this experiment did not replicate the null difference between timing phases for CV. Instead, in training there was a decrease in the CV with increasing target time (Figure 7). Others have reported violations of the scalar property when using ms target times in reproduction tasks (Wearden, 2003) and this is likely due to added variability from the motor component

involved in reproduction. However, what is unique to our results, is the fact that at testing these differences in variability disappeared, even with an increase in the estimates. Variability became scalar at testing. This provides additional evidence that the scalar property arises during retrieval.

## General Discussion

These experiments provided evidence for the theory that Vierordt's law is not specifically the overestimation of the shortest target time and the underestimation of the longest target time. When multiple target times are learned together the representations of these different durations migrate towards each other. Other researchers have already described the migration effect as Vierordt's law (Lejeune & Wearden, 2007; Yarmey, 2000). The typical overestimation of the shortest target time and the underestimation of the longest target time seem to hold specifically when the shortest target time is a ms target time (up to about .9s) and longest is a ss target time. Target times in the ms range tend to be overestimated while target times larger than about .9s tend to be underestimated. It is important to keep in mind that the origination of the 'true form' of Vierordt's law came from experiments completed by Vierordt on himself and one of his graduate students and have not been widely explored or replicated since (Lejeune & Wearden, 2007).

The implication that when multiple target times are learned together, they migrate towards each other contradicts the veridical nature of timing. Estimates of target times closely represent the actual target times (Jones & Wearden, 2003; Wearden & Jones, 2007; Wearden & Lejeune, 2008). However, in these studies participants were presented with a single target time and were required to make shorter/longer/same judgments with comparison durations. Since multiple target times were not presented together, it does not contradict our theory that when

target times are presented together, the representations migrate towards each other. In fact, if participants learn two standards, comparison judgments will be biased in the direction of the relationship between the two standards (Ogden, et al., 2008). If A is larger than B, and comparison judgments are being made with standard A, then longer comparison times are judged as shorter than A than when A is presented in isolation. In contrast, if A is smaller than B, the reverse is found. Essentially, participants spontaneously categorize the durations presented together and this may not be dependent on task type. Furthermore, target times presented outside of the learned grouping do not affect the grouped category, as our data shows with the retraining in Experiment 2. When trained in isolation the retraining interval was kept separate from the categorized group of target times.

There are, however, contradictions to the assumption that multiple target times learned together migrate towards each other. For example, in Rakitin, et al.'s (2005) study, student participants do not demonstrate a migration of the two target times (6s & 17s). We intentionally chose target times for Experiments 1 and 2 that overlapped with the Rakitin et al. (2005) study to assess if larger ss target times were impervious to the migration effect and found that the overlapping target times (6.5s & 10s) did migrate. A major difference between the two studies is that in Rakitin, et al.'s (2005) study, the two target times were presented in blocks, whereas in our study, they were presented intermixed. It makes sense that the spontaneous categorization of target times would be less likely if the times are presented in separate blocks versus presented in conjunction. Additionally, estimates might be more prone to migration if more than 2 are presented together.

Our studies also demonstrated that the migration effect is fairly short lived when ss intervals are used. In Experiments 1 and 2, in which ss target times were used, estimates

increased for the 2 longer target times, but not for the shortest target time. The migration effect appeared to be transient and not to survive a delay of even a few minutes. This could imply that the migration effect arises from encoding processes instead of long-term memory storage. Yet, when ms target times were used, the migration effect remained at later recall. This could be another indication of different neurocognitive processes underlying the two time ranges, both in learning and in retrieval after storage.

Alternatively, it is possible that uneventful durations of time may expand in memory. In our day to day lives, time is very much intertwined to the events that occur in time- what we need to do, when we need to get it done, how long it will take. Learning a time interval that is relatively devoid of events may create a situation in which the cognitive architecture we typically use to estimate time is no longer useful. Without an event structure, the memory representations of time may dilate. By categorizing the representation of different durations, participants may be able to diminish this dilation. This is consistent with the results of Gooch, et al., (in revision). They showed that when participants learned a single target time, estimates were much longer at a later testing, than when the same target times was learned and tested with other target times.

Our results also demonstrate a unique finding regarding the scalar property of variance. Gibbon (1992) suggested that scalar variance arose from memory processes, but it has long been assumed it came specifically from the encoding phase (Meck, 1996). Our data show that the scalar property can also arise at retrieval. In the first two experiments, there is no difference in the CV for the two timing phases. If scalar variability only arose at encoding, then when estimates for the 6.5s and 10s target times increased at later testing, we would expect the CV to decrease. Since it did not, we assume that scalar variability can also arise at retrieval.

Furthermore, the violation of the scalar property of variance at training in Experiment 3 disappeared at testing. This, again, suggests that scalar variance arises during retrieval.

In conclusion, our experiments support the theory that ms and ss target times are processed via different systems and that these differences persist over a delay. We also make the first differentiation between overestimation and underestimation of time intervals and a migration effect across multiple target times and provide a testable theory for when Vierordt-like effects appear in timing data. Intervals learned together will exhibit migration effects in the short term. Finally, our results show that scalar variability arises during recall of durations.

## Chapter 3

### Implicit and Explicit Timing

The results of the previous experiments demonstrate that our memory representations of time increase when there are delays, both in the short-term and over longer periods. If this were an accurate reflection of our use of temporal information then we would rarely be late and we would schedule ourselves plenty of time to complete that paper or prepare that lecture because when we think back to how long it took us to do these tasks previously our memory of the duration would be longer. The dominant approach to assessing timing in the lab may not represent how we use timing outside of the lab because participants focus solely on the temporal domain during the experiments. Very rarely do we focus solely on timing in our daily lives, but we continuously learn and use temporal information.

The studies discussed so far have relied on explicit memory. While the discussion of implicit and explicit memory has proven to be an important discourse in memory research over the last decade, the distinction has not been discussed for memory for time. This is in part because the majority of research on interval timing has focused on locating and verifying the internal clock, not the underlying memory mechanisms involved in learning and remembering intervals of time. Explicit memory is studied when participants are aware of the relevant stimuli and can verbally describe the stimuli. Implicit memory is studied when participants are not aware of the relevant stimuli, but still exhibit learning, or cannot verbalize the relationship of stimuli, but can still respond to the patterns (Smith & Grossman, 2007; Smith, 2008). In their seminal paper demonstrating separate memory systems, Knowlton & Squire (1993) presented amnesiac patients and controls with displays of dot patterns. In later testing, the groups were able to

categorize new displays correctly, but the patient group was not able to recognize which items were presented during training. Following this paper, a substantial amount of neuroimaging work, patient, and behavioral studies have supported the distinction of multiple memory systems (e.g. for reviews: Ashby & Maddox, 2005; Poldrack & Foerde, 2008; Smith & Grossman, 2007).

We wanted to use temporal stimuli to change behavior without making time explicit. To help clarify what this would mean, it is important to understand another type of distinction commonly made in the domain of temporal information processing- prospective and retrospective timing. The previous tasks have all been prospective. Participants were aware that the goal of the task was to estimate presented intervals of time. In retrospective tasks participants are not aware that the goal of the task is to estimate a duration of time. Retrospective tasks present participants with another task and afterwards participants are asked to estimate how long they spent doing the task. The temporal information isn't learned implicitly in these types of tasks. Instead evidence suggests that participants use their memory of events to estimate how long the duration lasted (Block, 1992). If more events occurred during the interval to be timed, estimates increase (Block, 1992; Hicks, et al., 1976; Zakay & Block, 2004). While prospective tasks are clearly explicit, retrospective tasks are not always implicit. Retrospective tasks begin similarly to standard implicit memory tasks in that the goal of the task is not known, but retrospective timing tasks become explicit because once participants are asked to exhibit a behavior (estimate a duration) they are told explicitly what is expected of them and use their memory of the events to explicitly construct an estimate of the duration.

Timing studies using non-human subjects should also be considered explicit. While the instruction is not verbally explicit, subjects are taught that they must attend to time through hundreds of training trials and they do not continue on to testing phases until reaching a specified

accuracy criterion. Animal subjects learn that they must know *when* to respond, so while the instructions are not verbally explicit, they are behaviorally so. Many timing tasks for humans are derived from tasks used with animals (e.g. temporal generalization, bisection, peak interval), but use verbal instructions. Even with these instructional differences, results from human and non-human timing tasks are similar (Allan 1998; Allan & Gibbon, 1991; Rakitin, et al, 1998; Wearden, 2003b) supporting the idea that the training non-human subjects undergo prior to testing make the goals of the task explicit.

In order to study implicit memory for time we constructed a task in which participants were required to categorize Mr. Potato Head™ into one of two categories of hat preference. The usage of this stimulus was modeled after Shohamy, et al., (2004) in which participants were required to make ice cream preference choices for Mr. Potato Head™. The various interchangeable features of Mr. Potato Head™ (e.g. glasses, mustache, bowtie) were used as probabilistic predictors of category membership. Each feature predicted category membership (i.e. ice cream choice) with a specific probability, such that neither a single feature, nor a rule based on the combination of features, could be relied on to indicate category membership. We wanted to be able to compare performance on the implicit timing task with performance on an explicit version of the task, so instead of using a probabilistic structure, we created a rule-based structure based on Allen & Brooks' (1991) study which used three relevant features, but only 2 out of the 3 needed to be present to predict category membership. One of the three relevant features was presentation duration, while the other two were visual features. Short presentation duration predicted one hat, while a longer presentation duration predicted the other hat. The original Allen & Brooks (1991) study did not target implicit learning, but focused on different memory systems engaged in rule-based versus similarity learning. However, if the rule in a rule-



based task is too complex to be verbalizable, implicit learning can occur (Smith, 2008). We anticipated that the unexpected nature of the temporal feature in combination with the visual features would reduce participants' ability to construct a rule related to it, allowing implicit learning, which would be demonstrated by their effective categorization of the stimulus.

Following a training phase with feedback, participants were given a testing phase that contained no feedback. The testing session presented images of Mr. Potatohead™ without a hat at both the short and long presentation durations. We also incorporated a temporal bisection task (a type of a timing discrimination task) into the testing phase. In temporal bisection tasks, participants are trained on two target times, a short and a long one. Then participants are presented with the two target times as well as intermediary durations. For the intermediary durations participants must decide if they are more like the short target time or more like the long target time. Typically, as the presentation duration increases the probability of responding long also increases allowing us to assess temporal accuracy and precision. In this task participants did not know that they needed to decide if the intermediary durations were more like the short or the long presentation duration. We simply presented them with the different durations and asked them to predict Mr. Potatohead's™ hat choice. This way, we could temporal accuracy and precision during the testing phase.

New and repeat items were also included during testing. In order to verify that a category is learned, it should extend to novel exemplars (Smith, 2008; Smith & Grossman, 2007). We assumed that participants who knew the rule would categorize novel and repeat items similarly, relying on the use of the rule. Whereas participants who did not know the rule may rely on memory representations of the specific items to correctly categorize exemplars and not generalize learning to novel items. This would mean that they were not implicitly learning the

relationship. If participants responded similarly to repeat and novel items we would infer that they had implicitly learned the relationship between the 3 features and hat choice. However, if participants responded more quickly or more accurately to repeat items than novel items, then we would infer that participants relied on item specific memory representations.

## Experiment 1

### Methods

*Participants:* Twenty-nine participants participated for course credit. Their average age (and standard deviation) was 18.79 (1.10). Participants were naïve to the purpose of the experiment. Fourteen female participants were randomly assigned to the explicit group and 15 participants were randomly assigned to the implicit group, 1 of whom was male.

*Materials:* Stimuli were created using photographs of a Mr. Potato Head™ doll. Four visual features varied across the images. Two of the four features, bow tie (absent, present) and mustache (absent, present), were irrelevant features included to increase the number of exemplars and to create noise to decrease the likelihood of learning the rule. The other two visual features, glasses (absent, present), and shoes (blue, white) were the relevant features that participants needed to learn. The length of time that the image was on the screen (short, long) was the third relevant feature.

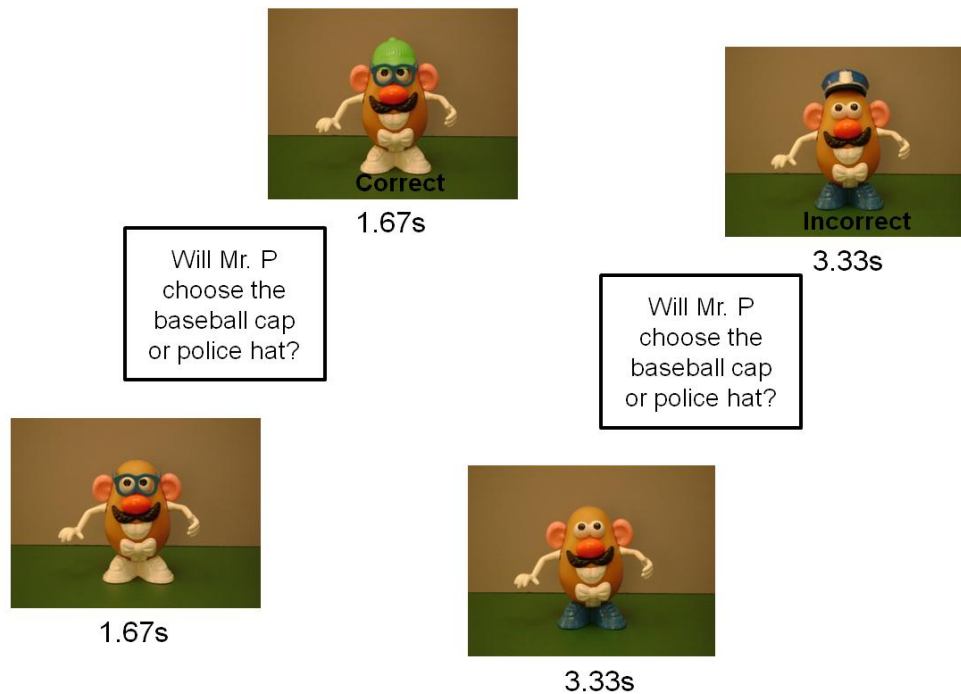
Category membership was defined by the presence of two of the three relevant features. Participants were required to use the features to predict if Mr. Potato Head™ would be wearing a police hat or a baseball cap in the following image. The three features that predicted the police hat were the presences of glasses, blue shoes, and a presentation time of 3.33s. The baseball cap was predicted by the absence of glasses, white shoes, and a presentation time of 1.67s. Since

category membership required two out of the three possible binary features, we could assess differences in learning when the two features were visual cues versus when the two features were temporal and visual cues. Each irrelevant feature occurred equally across both categories. A single relevant feature predicted the correct response 75% of the time.

In total there were 32 distinct exemplars based on the relevant visual and temporal features. Of these, 16 were presented during training and all 32 were presented during testing so that we could assess item specific memory. The images presented during training and testing were randomized across subjects.

*Procedure:* Images were presented using an executable program developed in Matlab using the psychophysical toolbox on PCs with the XP operating system (Brainard, 1997; Pelli, 1997). The training phase consisted of 80 trials, i.e. five presentations of the 16 exemplars. The training phase was designed in 5 blocks such that each block contained one presentation of each exemplar. The blocks ran continuously for the subject, but allowed us to assess if learning improved over time. A trial began with an image of Mr. Potato Head™ presented for one of the target times (1.67s or 3.33s) and was followed by a black screen with white text which stated: “Will Mr. Potatohead put on his police hat or his baseball cap? Press the blue key for the police hat and the yellow key for the baseball cap.” Responses were recorded on a standard keyboard to the nearest millisecond. The ‘f’ key was covered with a yellow sticker and the ‘j’ key was covered with a blue sticker. After participants made their choice they were shown an image of Mr. Potato Head™, identical to the previous one, except that now he was wearing the appropriate hat. Participants were told if they were correct or incorrect by the placement of the words ‘correct’ or ‘incorrect’ at the bottom of the photo (Fig. 8). The feedback remained on the screen for the appropriate target time (e.g. the time that the image was presented for previously)

so that participants could observe the relevant features during feedback. There was a fixed ITI of 4s between trials.



*Figure 8: Two examples of trials presented during training. First participants saw Mr. Potato Head™ without a hat and then were prompted to predict which hat he would choose to put on. The trial to the left, sequenced from bottom to top, demonstrates the series of events for a participant who correctly predicted that Mr. Potato Head™ would put on his baseball cap. ‘Correct’ is written at the bottom of the image. The trial to the right demonstrates the series of events for a participant who incorrectly predicted that Mr. Potato Head™ would put on his baseball cap. ‘Incorrect’ is written on the bottom of the image. Feedback lasted for the time used during presentation. Testing trials differed only in that there was no presentation of Mr. Potato Head™ in the correct hat or feedback.*

After the training session, subjects were shown white text on a black screen that told them that the testing session differed from the training session because there would no longer be feedback. This information remained on the screen indefinitely and participants were told to press any key when ready to continue. A temporal bisection task was incorporated into the discrimination task during testing. In a temporal bisection task intermediary comparison durations are included requiring participants to decide if the duration is closer to the longest target time or closer to the shortest target time. During testing, images of Mr. Potato Head™ were presented for the target times from training (1.67s, 3.33s) as well as for 5 comparison times (1.95s, 2.33s, 2.50s, 2.78s, 3.05s). These 7 target times were each presented for 16 trials randomly assigned to the 32 exemplars. Participants were not told that the presentation duration would vary during testing. The ITI was the same as in training (4s) and their response prompted the next trial.

At the beginning of the experiment participants were randomly assigned to one of two groups. The explicit group was told the categorization rules and given the following story that explained the rules and made clear how time was relevant. The implicit group was not given the categorization rules but the text in italics was presented to both groups:

*You will be shown a series of images of different mornings in Mr. Potato Head's life. You are asked to decide if Mr. Potato Head™ is going to work that day or if he is going out to run personal errands. If he is going to work, he will put on his police hat, but if he is going out on personal errands, he will put on his baseball hat. You will see an image of Mr. Potato Head™ without a hat and then you will decide which hat he will put on.*

To help you make your decision, there are some things you should know. When Mr. Potato Head™ gets up in the morning you can often tell where he will be headed that day by how long he stands in front of his closet. Typically, when he has to go in to work, he is slower to get ready and will stand in front of his closet

for more time. However, there are some mornings that he feels rushed, so the time he spends in front of his closet is not the only cue that will tell you he will be going that day. Mr. Potato Head™ is farsighted and on days when he has to go to work he will wear his glasses because he has a lot of paperwork to read. However, there are some days in which his personal errands also require him to wear his glasses and some days when he does not wear his glasses when he leaves for work. Finally, it is required by the Potato Head™ Police Department that all employees wear blue shoes. If Mr. Potato Head™ is wearing blue shoes, he is likely going to work that day. However, there are days when he forgets and wears his white shoes to work and his blue shoes to run errands. In order to know if Mr. Potato Head™ is going to work, there will need to be at least 2 of the 3 features that will signal his going to work.

After participants completed the training and testing phases of the experiment they were given a brief questionnaire. The implicit group was asked if they noticed the different presentation durations and if so, if they used them to make their choice. Both groups were asked to provide verbal estimates of the presentation time of Mr. Potato Head™ in his police hat and in his baseball cap.

*Data Analysis:* Percent correct and reaction times (RTs) were calculated for each condition for each participant for each block in training. Percent correct and RTs were submitted to separate 2 (cue: temporal, visual) by 5 (blocks) by 2 (group: explicit, implicit) repeated measures ANOVAs with group as a between subjects factor. Half of the participants (8/15) in the implicit group reported in the post-test questionnaire that they noticed the different presentation durations of Mr. Potato Head™. None of them reported using this information to make their choice. To assess if this awareness affected responses during training, we submitted accuracy measures to a 2 (cue) by 5 (block) by 2 (group: aware, unaware) repeated measures ANOVA with group as a between subjects factor.

Multiple analyses were completed to assess accuracy and response rate during testing. Percent correct and RTs were calculated for novel and repeat stimuli in trials in which category

membership was predicted by temporal and visual cues and the target time was one of the two standards (1.67s, 3.33s) presented during training. We did not calculate percent correct when the target time was one of the comparator durations, because none of them were actually correct. We also calculated percent correct and RTs for trials in which category membership was predicted by visual features across all levels of target time (since target time was an irrelevant feature in these cases). Percent correct & RTs were submitted to separate 2 (cue: visual, temporal) by 2 (new, repeat) by 2 group (explicit, implicit) repeated measures ANOVAs with group as the between subjects factor. This allowed us to see how accurate participants were at a later testing and if there were differences in novel and repeat stimuli.

We also calculated participants' probability of responding long (in this case, police hat) for each target time in which category membership was predicted by the temporal and visual feature. The probability of responding long (P(L)) was submitted to a 2 (new, repeat) by 7 (target time) by 2 (group) repeated measures ANOVA. Means of the verbal estimates from the post-test questionnaire were calculated and submitted to a 2 (hat) x 2 (group) repeated measures ANOVA. Accuracy from testing, P(L), and verbal estimates were submitted to separate ANOVAs testing for group differences in the aware and unaware participants in the implicit group.

## Results

The term 'visual cues' will refer to conditions in which both relevant features were visual cues and the term 'temporal cues' will refer to conditions in which one relevant cue was temporal and the other was visual.

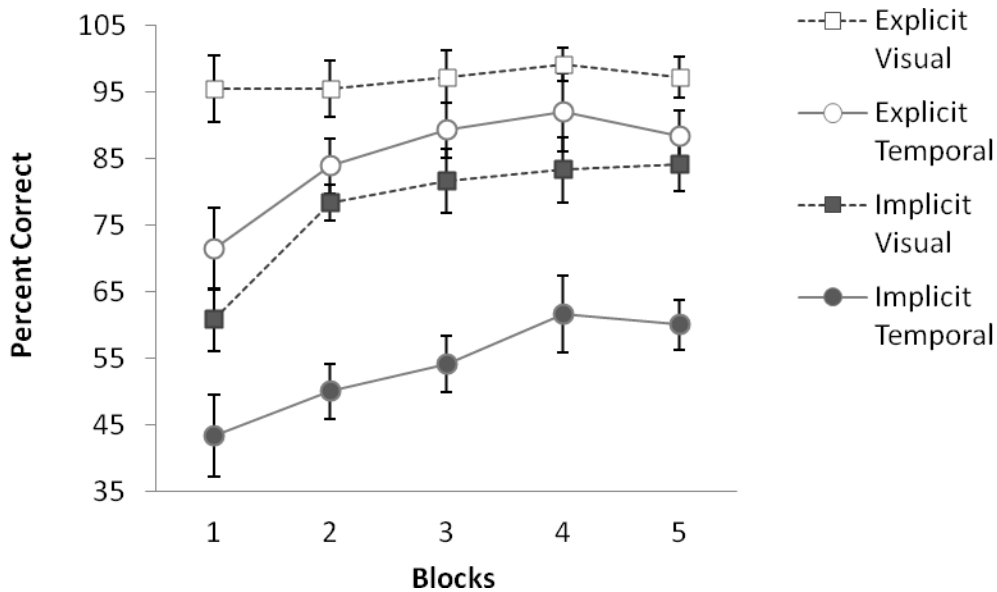


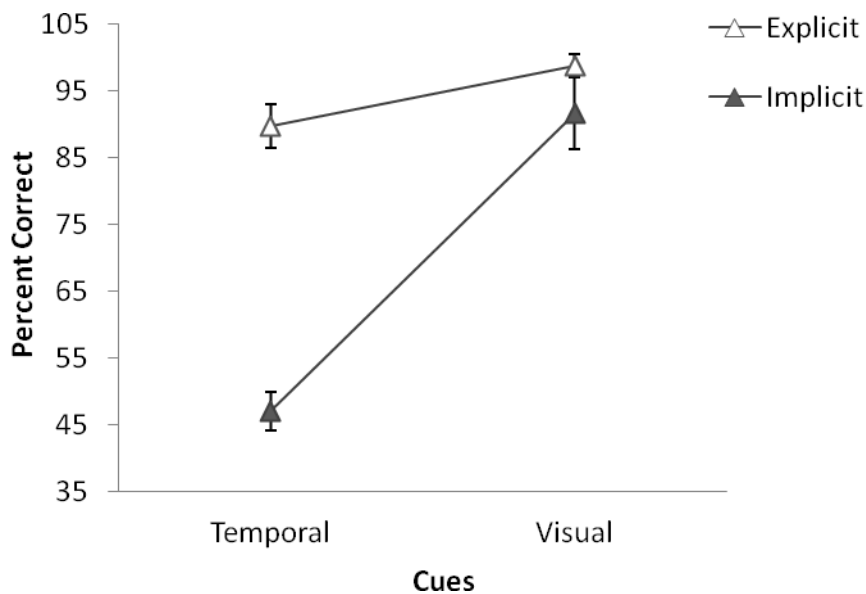
Figure 9: Percent correct for each cue/group relationship across blocks presented during training in Experiment 1. Error bars are standard error.

*Training:* There was a cue by block by group interaction for accuracy ( $F(4,108) = 2.49, p < .05$ ; see Figure 9). For the implicit group, accuracy was higher when presented with visual cues than when presented with a temporal cues ( $F(1,14) = 25.80, p < .001$ ). Also, for the implicit group overall accuracy increased as the task progressed ( $F(4,56) = 12.10, p < .001$ ). In contrast, for the explicit group there was a cue by block interaction ( $F(4,52) = 5.07, p < .01$ ). Accuracy for the visual cues for the explicit group didn't change across blocks, ( $F(4,52) < 1$ ), but accuracy for the temporal cues increased across blocks ( $F(4,52) = 9.26, p < .001$ ). Performance was better when presented with visual cues than with temporal cues for the explicit group ( $F(1,13) = 18.75, p < .01$ ). Additionally, the overall ANOVA showed a cue by group interaction ( $F(1,27) = 4.58, p < .05$ ), a main effect of cue ( $F(1,27) = 41.67, p < .001$ ), a main effect of group ( $F(1,27) = 82.99, p < .001$ ), and a main effect of block ( $F(4,108) = 20.09, p < .001$ ). Accuracy was higher for the



visual cues, and the explicit group was more accurate than the implicit group. Overall accuracy improved over blocks of trials.

RTs decreased over the blocks of trials ( $F(4,108) = 53.22, p < .001$ ). RTs were longer for the temporal cues than for the visual cues ( $F(1,27) = 21.21, p < .001$ ). Groups did not differ in their RTs ( $F(1,27) < 1$ ) and there were no other significant effects. The only significant effect of self-reported awareness of time in the implicit group was a cue by block by group interaction ( $F(4,52) = 4.34, p < .01$ ). Looking at the temporal cues only, there was not a significant effect of group ( $F(1,13) < 1$ ), which would be expected if participants were effectively using their awareness to categorize the images.



*Figure 10: Percent correct for explicit and implicit groups for stimuli that contained visually predictive cues (Visual) or combined visual and temporal cues (Temporal) for the testing phase of Experiment 1. There was a significant difference between groups for the temporal cues, but not the visual cues. Error bars are standard error.*

*Testing:* As seen in Figure 10, there was a significant cue by group interaction for accuracy ( $F(1,27) = 41.77, p < .001$ ). The two groups did not differ in accuracy for the visual cues ( $F(1,27) = 3.49, p = 0.07$ ), but they did differ when for the temporal cues ( $F(1,27) = 134.15, p < .001$ ). The explicit group was more accurate with the temporal cues than the implicit group. The overall ANOVA showed a main effect of cue, with accuracy higher for the visual cues than the temporal cues ( $F(1,27) = 95.04, p < .001$ ). Accuracy was also higher for the explicit group than for the implicit group ( $F(1,27) = 95.13, p < .001$ ). There were no significant effects of novel items (main effect:  $F(1,27) < 1$ ; interaction with group ( $F(1,27) = 1.53, p > .22$ ); interaction with cue:  $F(1,27) < 1$ ), suggesting that neither group used item specific memory to categorize the stimuli.

The only significant effect of RTs was a cue by novel by group interaction ( $F(1,27) = 5.45, p < .05$ ). For the implicit group RTs for the temporal cues were significantly longer than for the visual cues ( $F(1,14) = 7.10, p < .05$ ), however, this was not significant for the explicit group ( $F(1,13) < 1$ ).

There was a significant interaction between target time and group for the P(L) ( $F(6,162) = 46.69, p < .001$ ). As can be seen in Figure 11, the P(L) significantly increased as the presentation duration increased for the explicit group ( $F(6,78) = 65.25, p < .001$ ), but not for the implicit group ( $F(6,84) < 1$ ). There were no significant effects of novel items (main effect:  $F(1,27) = 1.15, p > .2$ ) nor any other significant effects.

A significant hat by group interaction ( $F(1,23) = 5.41, p < .05$ ) was found with verbal estimates collected in the post-test questionnaire (Table 8). Four participants did not fill out this question on the questionnaire, leaving a total of 25 participants (12 explicit, 13 implicit) in the

analysis. The verbal estimates of the explicit group were significantly larger for the police hat than for the baseball hat, as we would expect ( $F(1,11) = 342.29, p < .001$ ). However, verbal estimates for the implicit group did not differ across hats ( $F(1,12) < 1$ ). There were no significant differences between aware and unaware participants in the implicit group for accuracy during testing ( $F(1,13) < 1$ ), for P(L) ( $F(1,13) = 1.36, p > 0.26$ ), or for verbal estimates ( $F(1,11) < 1$ ).

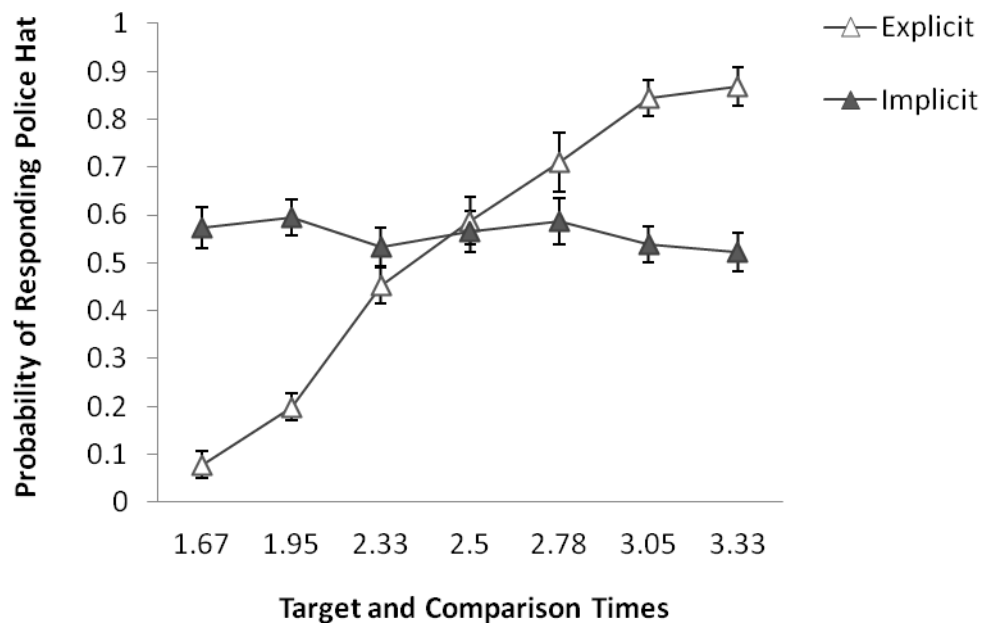


Figure 11: Probability of responding police hat (long, 3.33s) for the explicit and implicit groups for the target times (1.67s & 3.33s) and the comparison times in Experiment 1. Error bars are standard error.

Group	Police Hat	Baseball Cap
Explicit	4.42 (0.63)	2.62 (0.65)
Implicit	4.77 (4.73)	4.31 (3.65)

Table 8: Means (standard deviations) of the verbal estimates of the presentation duration of Mr. Potato Head™ in his respective hat (police hat or baseball cap) provided in a post-test questionnaire from Experiment 1.

## Discussion

The explicit group performed as expected. During training their performance on the visual cues was near ceiling, and there was an increase in performance on the temporal cues as training increased. This is most likely due to the instructions they were given. The instructions clearly specified the visually predictive cues, but the temporally predictive cues were relative (short time vs. long time). Participants would need to develop familiarity with the relation before performing at ceiling. This is especially noticeable in the steep increase from blocks 1 to 2 for the explicit group's accuracy of the temporally cued stimuli seen in Figure 9. Nonetheless, accuracy for temporal cues did not reach the same level as visual cues (although they did not significantly differ) for the explicit group suggesting that categorization tasks using these cues are more difficult. One reason that people may have been better at categorizing the visual cues than the temporal cues is that time perception can be variable across trials because it is affected by attention levels. When more attention is directed towards time, estimates will increase, and with less attention, estimates decrease (e.g. Macar, et al., 1994). Time is unlike other stimuli because the entire duration must elapse before it can be identified and stored in memory. In contrast, once visual cues are perceived, participants know whether or not Mr. Potatohead™ is wearing glasses, for example, and do not need to continue to focus their attention towards that. Attention needs to be directed towards time during the accumulation process. However, it is probably difficult to sustain that level of attention trial after trial, and therefore, it is likely that attention towards time fluctuated from trial to trial such that on some trials, the shorter target time could have seemed longer and the longer target time could have seemed shorter.

The implicit group was able to learn to categorize using the visual cues, but not the temporal cues. During training, accuracy for the visual cues for the implicit group was near the

accuracy levels for both cues for the explicit group (Fig. 9) suggesting that the implicit group was able to learn the to categorize the visual cues. This was corroborated by the fact that accuracy for the visual cues at testing was the same for the two groups suggesting that the parameters of the task were sufficient for learning to occur in the visual domain. The poor performance in the temporal domain is likely due to the fact that the visual domain is a dominant perceptual system in humans and the importance of the visual domain was further emphasized by the use of a toy that is a conglomeration of visual cues. Likely, subjects had some prior understanding or experience with Mr. Potato Head™ that emphasized the importance of the visual configurations and, perhaps, increased the irrelevance of time. However, these facts were intentionally utilized as a means of distracting participants from the relevance of time and preventing them from recognizing, and therefore verbalizing, the rules used to correctly choose Mr. Potato Head's hat.

While the implicit group did not learn the temporal cues, they also did not rely on a single visual cue to predict hat choice. If they had relied consistently on a single visual cue, accuracy would have been at 75%, since each individual cue predicted hat membership 75% of the time. Instead, participants in the implicit group had an accuracy level of 47% for the temporally cued stimuli. Additionally slower RTs during testing for the implicit group for the temporal cues suggested that participants were using more cognitive resources to surmise the correct answer, rather than responding based on a single specific visual cue. Some participants in the implicit group self-reported an awareness of different presentation durations, but analysis comparing the group who noticed with the group who did not, showed that self-reported knowledge did not alter accuracy.

Interestingly, the verbal estimates given by the implicit group for both hats that were larger than the actual presentation time for either hat (Table 8). The verbal estimates were most similar to the fixed ITI of 4s. We kept the ITI at a fixed duration since our previous experiments (e.g. Working Memory for Time) demonstrated that altering non-target intervals can affect estimates of the target time and we wanted to minimize changes in estimates due to changes in trial based intervals. However, since the fixed ITI was presented after every trial, it was the most frequently presented duration, and the implicit group may have been influenced by this information when making their verbal estimates. Additionally, the verbal estimates of the explicit group were higher than the actual presentation times (Table 8), so it is possible that verbal estimation yields overestimation, whereas reproduction yields underestimation of time intervals in a similar range.

Finally, both groups relied on rule-based mechanisms to complete the task. Accuracy for novel items was not different from accuracy for repeat items presented during testing suggesting that participants in both groups relied heavily on rule-based categorization strategies rather than item specific memory of the stimuli. The improved performance for visual cues versus temporal cues for the implicit group combined with the null result between novel and repeat items demonstrates that the implicit group strived to construct a rule for completing the task and did not rely on other memory systems.

The results of Experiment 1 suggested that the visual cues used to prevent the verbalizability of the time-based rule were too distracting for participants in the implicit group to learn the time-based cue, implicitly or explicitly. For Experiment 2 we decided to increase the length of the longer duration from 3.33s to 5.01s to make the intervals more discriminable. As discussed previously, attention towards time can affect the perceived length of duration.

Participants in the implicit group do not attend to time, because we have not told them to, and the similarity of the two target times may make it difficult to differentiate two times. We hoped that the increased difference between the short and long target times would enable the implicit group to utilize the temporal information implicitly.

## Experiment 2

### Methods

*Participants:* Thirty-seven participants (mean age 19.03(2.6)) participated in the study for course credit. Their average age (standard deviation) was 19.03 (2.6). Nineteen participants were randomly assigned to the explicit group (18 female, 1 male) and 18 to the implicit group (all female). Participants were naïve to the purpose of the experiment.

*Procedure and Materials:* The only difference between Experiments 1 and 2 was that the target time used as a predictor for the police hat was increased from 3.33s to 5.00s. The target time used to predict the baseball cap remained at 1.67s. During the testing phase, there were 5 different comparison times: 2.23s, 2.78s, 3.34s, 3.89s, and 4.45s. All other events and stimuli in Experiment 2 were identical to Experiment 1.

*Data Analysis:* All data was calculated and analyzed identically to Experiment 1. Only 3 out of the 18 participants in the implicit group reported awareness of the differences in presentation duration on the post-test questionnaire and therefore, analysis assessing the differences in estimates in the implicit sub-groups was excluded.

## Results

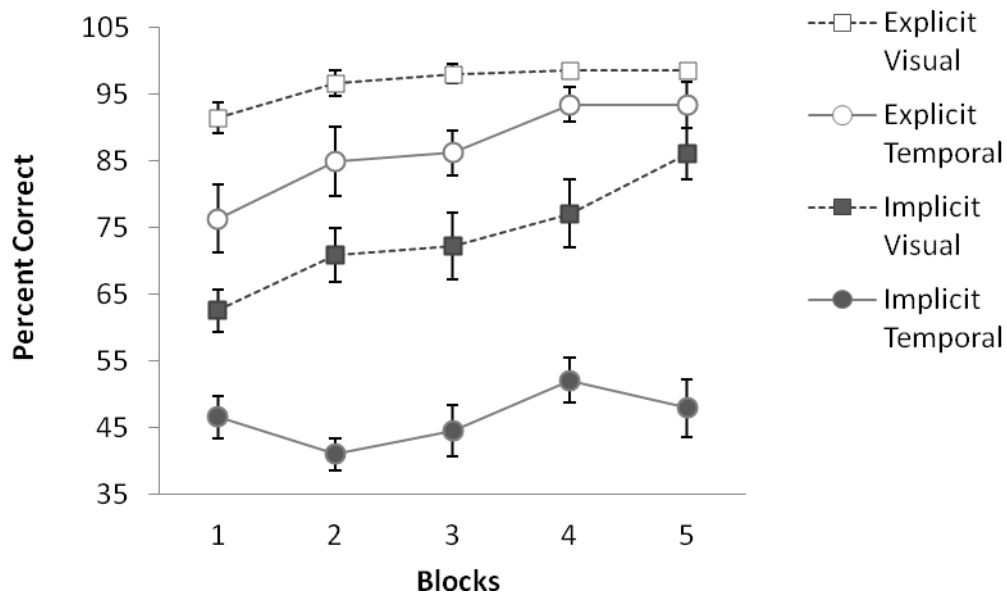


Figure 12: Percent correct for each cue/group relationship across blocks presented during training in Experiment 2. Error bars are standard error.

*Training:* There was a significant cue by block by group interaction ( $F(4,140) = 4.62, p < .01$ ), demonstrated in Figure 12. For the implicit group there was a significant cue by block interaction ( $F(4,68) = 3.08, p < .05$ ). Accuracy increased over blocks for the visual cues, but not for the temporal cues. For the explicit group this interaction was marginally significant ( $F(4,72) = 2.45, p = 0.054$ ). As in Experiment 1, accuracy for the temporal cues for the explicit group increased during training ( $F(4,72) = 10.21, p < .001$ ). Overall, there was also a significant cue by group interaction ( $F(1,35) = 9.98, p < .01$ ). Accuracy increased as across blocks ( $F(4,140) = 13.84, p < .001$ ). Accuracy scores were higher for the explicit group than the implicit group ( $F(1,35) = 184.93, p < .001$ ) and accuracy scores were also higher for the visual cues stimuli than for the temporal cues ( $F(1,35) = 45.18, p < .001$ ).

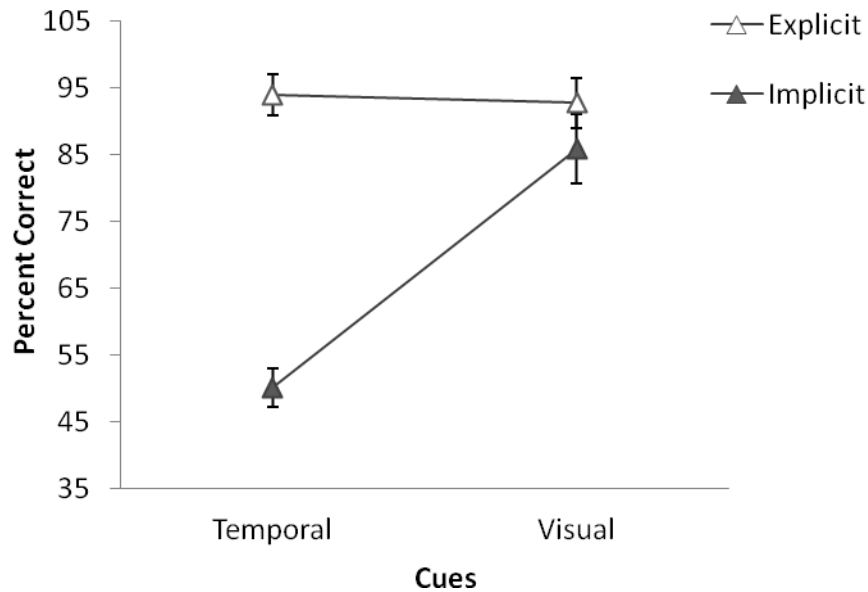


RTs significantly decreased over the blocks of trials ( $F(4,140) = 48.69, p < .001$ ). There was also a significant block by group interaction ( $F(4,140) = 3.84, p < .01$ ). As can be seen in Table 9, RTs decreased substantially more for the explicit group than for the implicit group as the training session progressed. There were no other significant effects of RTs. This is a deviation from Experiment 1 in which there were no differences between the two groups.

Group	Training				
	Block 1	Block 2	Block 3	Block 4	Block 5
Explicit	1.64 (0.12)	0.84 (0.09)	0.71 (0.06)	0.64 (0.07)	0.62 (0.07)
Implicit	1.39 (0.12)	0.86 (0.09)	0.79 (0.07)	0.84 (0.07)	0.84 (0.07)

*Table 9: Means (standard deviations) of reaction times at each block for each group in Experiment 2.*

*Testing:* There was a significant cue by group interaction for accuracy ( $F(1,35)=58.31, p < .001$ ). As demonstrated in Figure 13, participants in the implicit group performed similarly to the explicit group when the hat choice was predicted by visual cues only ( $F(1,35)=2.42, p > .12$ ), but performed much worse when hat choice was predicted by temporal cues ( $F(1,35)=132.80, p < .001$ ). The overall ANOVA showed that participants were more accurate when presented with the visually predictive cues than with the temporally predictive cues ( $F(1,35)=50.95, p < .001$ ). The explicit group was significantly more accurate than the implicit group ( $F(1,35)= 56.86, p < .001$ ). There were no significant effects of novelty (main:  $F(1,35)<1$ ), suggesting that episodic memories were not used by either group.



*Figure 13: Percent correct for explicit and implicit groups for stimuli that contained visually predictive cues (Visual) or combined visual and temporal cues (Temporal) for the testing phase of Experiment 2. There was a significant difference between groups for the temporal cues, but not the visual cues. Error bars are standard error.*

There were no significant effects of RTs. Participants responded at a similar rate regardless of stimulus predictor or group membership, which demonstrated a deviation from the three-way interaction (cue by novelty by group) from Experiment 1.

There was a significant target time by group interaction for the probability of responding long ( $F(6,210)=61.586$ ,  $p < .001$ ; Fig. 14). The P(L) significantly increased as the presentation duration increased for the explicit group ( $F(6,108)=93.36$ ,  $p < .001$ ), but not for the implicit group ( $F(6,102)=1.24$ ,  $p > .29$ ). There was also a main effect of P(L) ( $F(6,210)=61.59$ ,  $p < .001$ ), but not of group ( $F(1,35) = 2.99$ ,  $p > .09$ ).

Two participants from the implicit group did not provide verbal estimates on the post-test questionnaire so that the analysis contained 16 participants in the implicit group. There was a main effect of hat ( $F(1,33) = 25.73, p < .001$ ) and significant hat by group interaction ( $F(1,33) = 8.73, p < .01$ ). Verbal estimates for the police hat were significantly higher than for the baseball cap for the explicit group ( $F(1,18) = 82.13, p < .001$ ), but there was no effect of hat for the implicit group ( $F(1,15) = 1.23, p > .2$ ; Table 10).

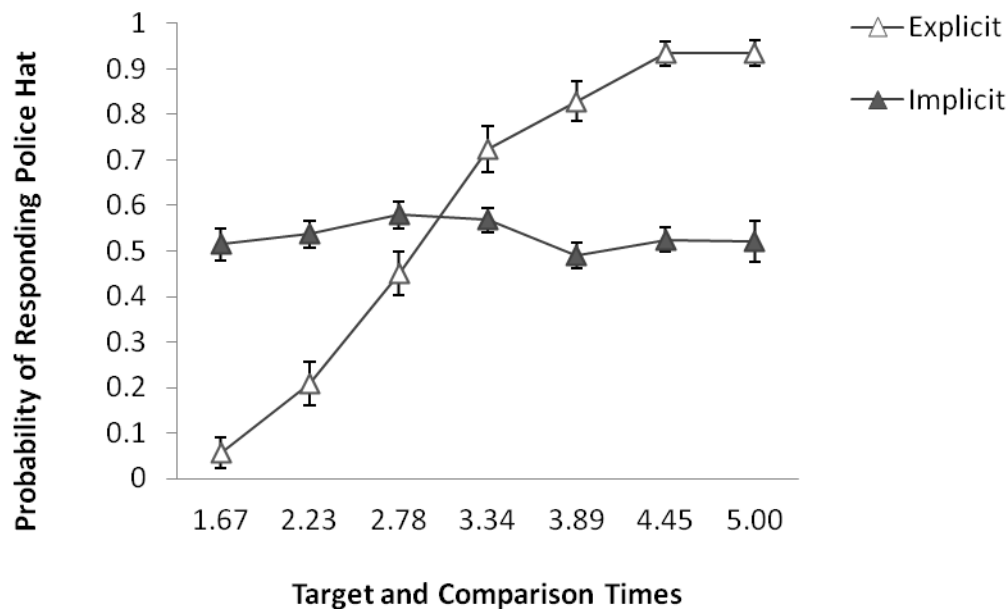


Figure 14: Probability of responding police hat (long, 5.00s) for the explicit and implicit groups for the target times (1.67s & 5.00s) and the comparison times in Experiment 2. Error bars are standard error.

Group	Police Hat	Baseball Cap
Explicit	4.87 (1.03)	2.74 (0.92)
Implicit	3.81 (2.71)	3.25 (1.48)

*Table 10: Means (standard deviations) of the verbal estimates of the presentation duration of Mr. Potato Head™ in his respective hat (police hat or baseball cap) provided in a post-test questionnaire from Experiment 2.*

## Discussion

The results of this experiment replicated the previous experiment with few deviations. The increased discriminability of the intervals caused no major differences in the overall results for either group. Specifically, the larger range for the two target times did not improve performance on the temporal cues for the implicit group. Since we incorporated visual cues into the task to serve as distracters from the temporal cue with the intention of creating a situation in which temporal information would be used to predict hat choice implicitly, we decided to use the visual cues as distracters only.

In this final experiment, we follow the same procedure with the same set of stimuli, except that the predictor for hat choice is the presentation duration only, i.e. the temporal cue. Visual cues are no longer relevant and are presented equally across the two presentation durations. Therefore, the visual cues serve as distracters, but do not also serve as a solution for part of the categorization task.

## Experiment 3

### Methods

*Participants:* Twenty participants participated for course credit. Their average age (and standard deviation) was 19.5 (1.85). Ten were randomly assigned to the explicit group (7 female, 3 male) and ten were assigned to the implicit group (9 female, 1 male). Participants were naïve to the purpose of the experiment.

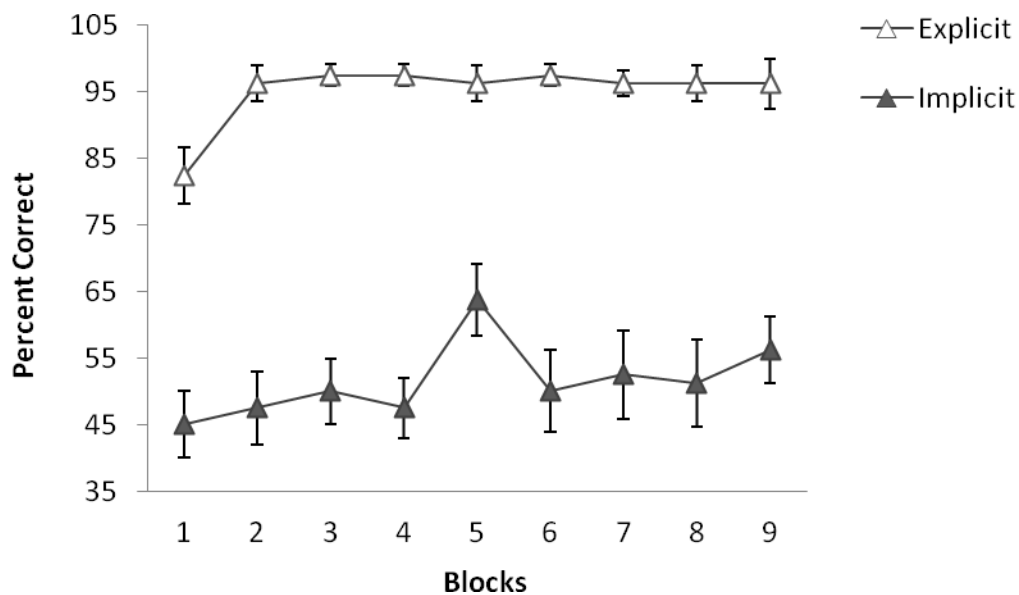
*Materials:* Half of the images used in the previous experiments were used in Experiment 3 because not all images of Mr. Potato Head™ without a hat had an accompanying image of Mr. Potato Head™ in both hats. To pilot this study, we used the images of Mr. Potato Head™ without a hat if we also had the same images of Mr. Potato Head™ in both hats. Visual features were irrelevant and predicted hat choice 50% of the time.

*Procedure:* There were fewer exemplars, so we included more blocks in training. Instead of 5 blocks of 16 exemplars (as in the previous experiments), there were 9 blocks of 8 exemplars in training. The other events of the trials in the training phase were the same as discussed in Experiments 1 and 2. Target times were 1.67s and 3.33s for the baseball cap and police hat respectively, the same targets used in Experiment 1. Images were presented for the same 7 target times during testing as in Experiment 1 (1.67s, 3.33s, and 1.95s, 2.33s, 2.50s, 2.78s, and 3.05s). Each target time was presented 14 times and randomly assigned to the 16 exemplars.

*Data Analysis:* The same measures were calculated as in Experiment 1. For training, percent correct and RTs were submitted to separate 9 (block) by 2 (group) ANOVAs with group as the between subjects factor. For testing, percent correct and RTs were calculated for trials in which one of the two standard target times (1.67s & 3.33s) was presented. These measures were submitted to separate 2 (new, repeat) by 2 (group) ANOVAs. The probability of responding long

(P(L)) was calculated for all trials at testing and submitted to a 2 (new, repeat) by 7 (target time) by 2 (group) ANOVA. One participant from the implicit group failed to provide verbal estimates of the presentation duration of each hat type in a post-test questionnaire. The provided verbal estimates were submitted to a 2 (hat) by 2 (group) ANOVA. Six out of the 10 participants in the implicit group self-reported noticing differences in presentation duration. However, because the implicit group was so small, we did not analyze the sub-groups separately in this experiment.

## Results

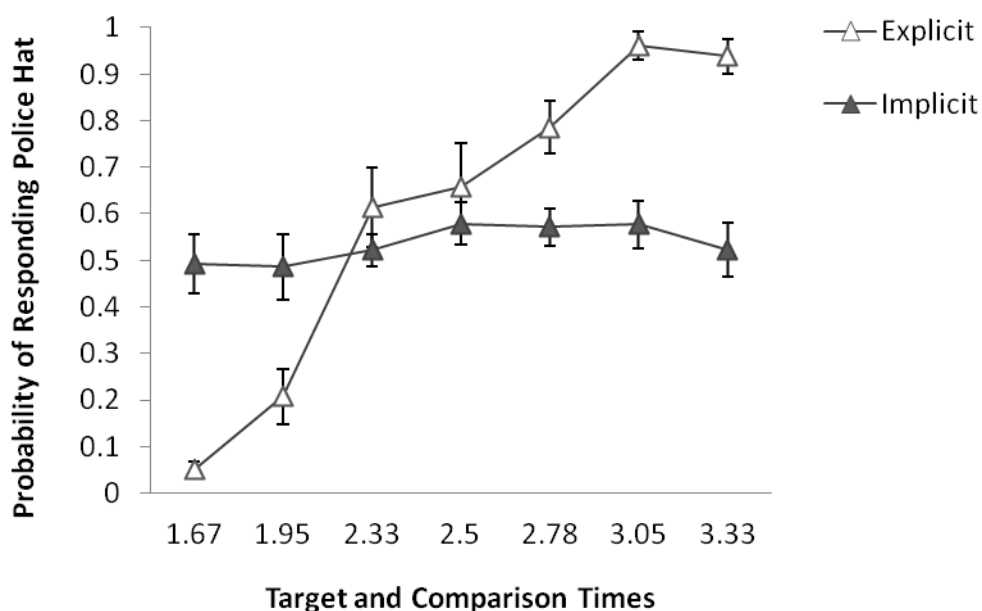


*Figure 15: Percent correct for each group across blocks presented during training. Time was the only predictive cue in Experiment 3. Error bars are standard error.*

*Training:* Accuracy increased as training progressed ( $F(8,144) = 2.31, p < .05$ ). The explicit group was significantly more accurate than the implicit group ( $F(1,18) = 198.00, p < .001$ ).

Although Figure 15 demonstrates a peak in accuracy at block 5 for the implicit group, there was

not a significant block by group interaction ( $F(8,144) = 1.05, p > .39$ ). RTs tended to decrease as training progressed ( $F(8,144) = 27.64, p < .001$ ). RTs were also faster for the explicit group than for the implicit group ( $F(1,18) = 7.51, p < .05$ ). The interaction was not significant ( $F(8,144) < 1$ ).



*Figure 16: Probability of responding police hat (long, 3.33s) for the explicit and implicit groups for the target times (1.67s & 3.33s) and the comparison times in Experiment 3. Error bars are standard error.*

*Testing:* As in the previous experiments, the explicit group was significantly more accurate than the implicit group ( $F(1,18) = 64.82, p < .001$ ). There was no difference in accuracy for new and repeat items ( $F(1,18) = 1.34, p > .26$ ). The explicit group responded significantly faster than the implicit group ( $F(1,18) = 9.26, p < .01$ ), and there were no other significant effects of RTs. The probability of responding long (P(L)) increased with increasing target time ( $F(6,108) = 25.36, p$

< .001). There was also a significant group by target time interaction ( $F(6,108) = 19.01, p < .001$ ). The P(L) increased with increasing target time for the explicit group ( $F(6,54) = 49.63, p < .001$ ), but this was not the case with the implicit group ( $F(6,54) < 1$ ; Fig. 16). There were no other significant effects for P(L).

All participants provided larger verbal estimates for the police hat than for the baseball cap ( $F(1,17) = 29.41, p < .001$ ). There was no difference between the groups ( $F(1,17) = 1.65, p > .2$ ) and no significant group by hat interaction ( $F(1,17) = 1.34, p > .2$ ; Table 11).

Group	Police Hat	Baseball Cap
Explicit	4.05 (0.86)	2.25 (0.54)
Implicit	3.28 (1.56)	2.11 (0.65)

*Table 11: Means (standard deviations) of the verbal estimates of the presentation duration of Mr. Potato Head™ in his respective hat (police hat or baseball cap) provided in a post-test questionnaire from Experiment 3.*

## Discussion

The final experiment differed from the previous two in that presentation duration was the only predictor of hat choice. The visual cues were used as distracters to prevent participants in the implicit group from forming rules about the predictor of hat choice. As in the previous two experiments, participants in the implicit group were not able to accurately perform the categorization task. However, when asked about the duration presentation of each hat, they were able to distinguish the presentation durations of each hat and this performance did not differ from controls.



In this case, the verbal estimates are an example of retrospective timing. After the task participants were asked to provide temporal estimates of the presentation duration of each hat. Since the manner of collecting this measure makes time explicit, it is unlikely that the temporal information was learned implicitly. Retrospective timing is thought to rely on memory of the events that took place during the interval to be timed (Block, 1992). In fact, the reliance on memory for events for retrospective timing is so strong that when more events are included in an interval it is estimated as longer than when fewer events are included in the same interval (Hicks, et al, 1976; Zakay & Block, 2004). What is interesting about the current experiment is that the experimental events for the presentation of the two different hats were identical except for the difference in target time. It is possible that internal mental events occurring during the experiment could be utilized by participants to reconstruct the duration retrospectively. We assume that task related mental processes are relatively similar for the two hat presentations, however, the longer presentation duration would allow for more mental processes. Therefore, it is difficult to assume that the mechanisms typically involved in retrospective timing tasks were used in this task.

Likewise, it is equally valid to speculate that their performance on verbal estimates was based on implicit learning of the target times. It is possible that participants in the implicit group were able to implicitly learn the target times and utilize that information when explicitly required, but they did not extend that information to the categorization task. This is consistent with the idea, supported by the previous experiments, that performing a categorization task of this type using temporal information is more difficult for participants.

## General Discussion

Overall, participants in the implicit group were unable to categorize Mr. Potato Head's™ hat choice using temporal information. However, when temporal information was the only predictor for hat choice, participants in the implicit group were able to differentiate the durations when giving retrospective verbal estimates of the time each hat was presented for. This likely didn't happen when time was not the only predictor for hat choice because participants were heavily focused on the relationship between the visual stimuli and predicting hat choice and there was positive feedback that visual cues were important to the task. The intense focus on visual cues may have prevented any level of encoding (implicit or otherwise) of other elements of the trials, such as duration presentation. The visual stimuli were included as a means of distracting participants from time to prevent them from creating a verbal rule using duration. However, it seems that because we commonly think of categories as being constructed out of visual cues and Mr. Potato Head™ is a toy that is constructed out of visual parts, participants may have been too distracted by the visual cues to utilize other elements of the trials. We had anticipated that the visual features would dominate perceptual awareness, but we'd hoped that these cues would have served as distracters from the temporal domain, but not be so distracting as to prevent all forms of learning of temporal information for participants in the implicit group.

Still, the fact that groups did not differ in their verbal estimates in Experiment 3 is very intriguing. Verbal estimates given by the implicit group were the same for the two hat types in the first two experiments when both visual and temporal cues were used, but the implicit group was able to retrospectively differentiate presentation duration of the two hats when the visual cues were not predictors of hat choice. Typically, in retrospective timing tasks, participants use their memory of the events of the trials to reconstruct how long the to-be-estimated interval

lasted (Block, 1992). The more events included during the interval to be retrospectively timed, the longer estimates are (Hicks, et al, 1976; Zakay & Block, 2004). However, the events in the two trials were identical, except for presentation duration. This makes it difficult to assume that participants made their estimates using the same memory mechanisms used in other retrospective timing tasks. It is also possible that participants in the implicit group learned the two presentation durations implicitly, but did not apply that information to the categorization task, until explicitly asked to do so. This didn't happen in the first two experiments because visual information allowed them to solve the task some of the time. When that cue was removed from the experiment, participants were able to learn more about the events of the trials, hence they may have learned the times implicitly, but they did not use that information to categorize hat choice. Alterations in the event structure during the trials could help clarify if the temporal information was learned implicitly and not used in categorization, or if the information was reconstructed from event based information in memory. By embedding trials with different numbers of events into the experiment, such as beeps of sound that could be used as a mental architecture for recalling the duration, we could replicate the retrospective phenomenon showing that more events during the trial led to increased estimates, thereby supporting the use of retrospective memory for recalling the durations in the current experiment. In contrast, if this effect is not found, we may assume that without other predictive cues, time was learned implicitly but not used.

It is possible that the stimulus we chose was not the most conducive to categorization by combining the modalities of vision and time. There were no differences between new and repeat stimuli for any measure at testing which implies that participants were heavily focused on creating verbal rules for completing the task and did not attempt to rely on item specific memory

for effective categorization. In contrast, the Allen & Brooks (1991) study that these experiments were partially based on, showed that participants relied on episodic memory for categorizing repeat stimuli. The fact that our participants did not utilize episodic memory for repeat stimuli implies that the features of our stimuli were not sufficiently integrated as a unified whole.

Participants in the explicit group relied on the rules rather than their memory of the stimuli.

Participants in the implicit group relied on rules for the visual cues and tried to create a rule for the temporal cues rather than rely on their memory of the stimuli. The conglomeration of features of Mr. Potato Head™ create well integrated stimuli (Shohamy, et al, 2004), however, the inclusion of the temporal domain may have diminished the perceptibility of wholeness.

Integrated stimuli are important for categorization, although they are not necessarily important for implicit learning, but this may explain why both groups relied on rule-based strategies. This may also explain why temporal information, if learned implicitly in Experiment 3, was not used to categorize items. A more integrated stimulus might have allowed for effective categorization in the implicit group. The existing task could be modified to use visual, auditory, and temporal information. For example, participants could be told that they need to categorize the significance of calls produced by artificially created animals and be presented with different types of artificially constructed animals making different types of sounds for different lengths of time. This would create a categorization task that crosses 3 modalities (visual, auditory, and temporal), increasing the likelihood that there would be multi-modal integration in categorization. This potential experiment also uses a real-world analogy which may be more straight forward than how long one takes to get ready, as was used in the current studies.

Additionally, the implicit learning of time could be explored using probabilistic tasks, as used in the Weather Prediction task or the task that originally used of Mr. Potato Head™

(Shohamy, et al, 2004; Shohamy, et al, 2008). For example, target time  $a$  may predict category  $A$  with a .8 probability, whereas target time  $b$  may predict category  $A$  with a .2 probability, with target times in between predicting the outcome with probabilities in between. Some additional sensory modality must be used to present the temporal information. If using the auditory modality, the auditory stimuli would also have a specified probability structure. In such a task, it would be possible to inform participants of the relevance of duration, just as participants are told to pay attention to the features of Mr. Potato Head™ or the Weather Prediction cards. However, having explicit knowledge of the importance of time would not enable participants to complete the probabilistic structure of the task explicitly. Participants would likely respond based on the probabilities but not have explicit knowledge of what motivates their responses.

In conclusion, these studies have provided insight into the importance of understanding how different types of memory may affect our learning, recognition, and recall of durations of time. These studies show that creating categories does not occur implicitly across perceptual domains, especially when one of the domains (vision) dominates the other (time). Furthermore, when feedback indicates that visual cues will predict the outcome some of the time, nothing else is learned implicitly about the trials. In contrast, when visual cues do not predict the outcome, temporal information about the trials is learned, but not used to complete the categorization task, even though it is available for explicit recall later.

## Final Conclusion

In the three chapters we demonstrated the importance of understanding temporal perception across memory systems. In the first chapter, Working Memory for Time, we presented two experiments with increasing delay periods inserted into a timing reproduction task between presentation and reproduction. A roving referent paradigm was used to target working memory processes involved in timing. A millisecond (ms) and supra-second (ss) target time were compared. The ms target time was overestimated and ss target time was underestimated, however, estimates of both target times increased with increasing delay period. In Experiment 1 we proposed that the delay period interfered with the memory representation of the target time and in Experiment 2 we demonstrated that the amount of interference depended on the proportional relationship between the delay period and the target time. SET (scalar expectancy theory) was used to explain this interference effect. Some amount of temporal information accumulated during the delay period even though it was an irrelevant duration and this additional accumulation was added to the working memory representation of the target time leading to longer estimates. The primary contribution of this study demonstrated that the time period between duration presentation and reproduction systematically interferes with the working memory representation of the target time by causing it to increase.

This is different from previous work demonstrating subjective shortening after delay periods in discrimination tasks and suggests that task type plays an important role in understanding memory processes. Discrimination timing tasks and reproduction timing tasks are analogous to recognition and recall memory tasks discussed in the non-temporal domain. Recall memory tasks are assumed to require more processing resources because participants must retrieve the full representation rather than recognize it for comparison, as in the discrimination

tasks. Research has suggested that different timing tasks do not always yield the same results (Allan, 1979; Bangert, et al, 2011; Jones, et al. 2008). Distinguishing reproduction and discrimination timing tasks as recall and recognition tasks retrospectively, helps to explain why there are differences and points to the importance of approaching an understanding of time perception by utilizing what is understood about non-temporal memory.

In the second chapter, Memory of Multiple Time Intervals, we presented three experiments with a delay-causing distracter task between a block of training trials and a block of testing trials of a reproduction task. In the first two experiments participants learned three ss target times, all of which were underestimated with respect to the target time. However, estimates were less underestimated for the shortest target time than the longest target time providing evidence for Vierordt-like migration effects. A similar effect was found for the third experiment using ms target times. The two shorter target times were overestimated, as would be expected for ms target times, and the longest was underestimated.

One of the primary contributions of the three experiments from chapter 2 comes from the training block showing that the migration effect commonly referred to as Vierordt's law is independent from the respective overestimation and underestimation of ms and ss target times. It is likely that the context of learning multiple target times together caused participants to organize the presentations into a category and that this mental organization caused the memory representations of the target times to drift towards each other. This was corroborated in Experiment 2 when another target time was presented during a retraining phase using the same cue. Participants kept the memory representation of the group of intervals presented at training independent from the retraining target time. This unique understanding of the mechanism involved in Vierordt-like migration effects would be difficult to isolate using a discrimination

task. Interestingly, after a delay of a few minutes, the Vierordt-like migration effect diminished and was replaced by an increase in estimates. This could mean that the migration effect is related to encoding processes instead of long-term memory storage. However, the ms target times migrated at a later recall. This may be another distinction underlying the two time ranges.

The overestimation of ms target times and the underestimation of ss target times, as well as Vierordt-like effects, are not accounted for in SET. Typically, it is assumed that estimates of time are veridical with actual time (e.g. Wearden, 2003) and both of these phenomena contradict that assumption. It is difficult to reconcile the overestimation of ms target times and underestimation of ss target times with the veridical nature of time estimates proposed by SET. However, Vierordt's law need not be contradictory to SET. If it is assumed that Vierordt-like migration effects arise due to categorization of multiple target times when learned together, this would qualify a difference between learning time intervals in isolation and in a group. This means that estimates are veridical to actual time when learned in isolation, but not when learned in a group.

Another main contribution of these experiments was that the scalar property of interval timing became evident during the later testing block, suggesting that the scalar property derives from retrieval processes. In SET the internal clock emits pulses following a Poisson distribution, so as estimates increase they would not remain scalar (Gibbon, 1992; Staddon & Higa, 1999). Therefore, it has been argued that the scalar property arises from memory or decision processes (Gibbon, 1992; Gibbon & Church, 1984). However, this has been difficult to verify. The results from chapter two effectively isolate later retrieval from encoding processes and the scalar property is not violated, even though estimates increase between training and later testing. This



provides evidence that the scalar property arises from memory processes and specifically from the retrieval phase of memory.

In the third and final chapter, *Implicit and Explicit Timing*, three experiments attempted to target the implicit learning of temporal information by presenting images of a Mr. Potato Head™ figure and requiring participants to predict one of two outcomes. Half of the participants were told the relationship between the predictors and the outcomes (explicit group), while the other half were not given this information, and were expected to learn it through observation with feedback (implicit group). The distinction between implicit and explicit memory is not made in SET. For this reason, we felt it was important to understand if time perception was similar to or different from other types of perceptual domains that can be learned using implicit paradigms. In the first two experiments both visual and temporal cues served as outcome predictors and we assumed that this combination of cues would make it difficult for the implicit group to explicitly learn the relationship between the temporal predictor and the outcomes. Instead, we found that the implicit group was able to learn the relationship only when visual cues predicted the outcome, but not when temporal cues predicted the outcome. In Experiment 3, temporal information was the only predictor. The implicit group was not able to use this information to successfully predict hat choice, but when asked afterwards how long they thought the presentation durations were, their estimates matched those given by the explicit group, even though this didn't happen with the first two experiments. Isolating time as the only feature gave participants in the implicit group sufficient information to learn the different presentation durations, even though they didn't use that information for categorization. Future work should explore how the explicitly elicited estimate given by the implicit group was encoded. They may have learned the times implicitly, but not used them in the task. Alternatively, they may have

used retrospective timing memory to recall the different presentation durations, but that doesn't completely account for why they were unable to do this for the first two experiments.

Combined, these separate studies demonstrate that there are a variety of methods for productive experimental exploration using what we understand from memory research and applying that information to the temporal domain. Although these separate chapters target very distinct memory systems, there are a few ways in which the results relate. At the end of chapter 1 (Working Memory for Time) we described that the overestimation and underestimation of ms and ss target times respectively could be an example of Vierordt's law, instead of evidence that the two target ranges are processed differently. In chapter 1, because the delay affected the two target times similarly (estimates for both increased with increasing delay length), we thought this could be evidence that the two target ranges (ms & ss) were not processed differently, and as an alternative explanation, we suggested that the overestimation of ms target times and the underestimation of ss target times were evidence of Vierordt's law. However, in chapter 2 (Memory of Multiple Time Intervals), we demonstrated that ms target times are overestimated and adhere to the Vierordt-like migration effect and that ss target times are underestimated and also adhere to the migration effect definitively showing that these effects are distinct.

Throughout chapters 1 and 2 we also speculated that the overestimation of ms target times could be a result of the reproduction task. Wearden (2003) demonstrates how both the overestimation of ms target times and the underestimation of ss target times could be achieved with one mechanism. When participants are going to make a temporal estimate, the amount of accumulating information must reach a specific criterion prior to initiating a response, according to SET. When that criterion is met, a response is initiated. Following Wearden's (2003) example, let's assume that criterion is 60% of the target time and that the reaction time is 300ms. If the

target time is 500ms, participants will initiate a response when 300ms has elapsed, and after adding the response time, the final estimate will be 600ms, which is larger than the target time. In contrast, if the target time is 1200ms, the response will be initiated at 720ms, and with the reaction time included, the final estimate will be 1020ms; smaller than the target time. While this accounts for the data at short ss intervals, it does not account for it when ss target times are larger. For example, if the target time is 10s, then the response would be initiated at 6s and the final estimate would be 6.3s. We can easily see from chapter 2 that the underestimation of the larger ss target times is not so extreme. For this to hold, we must assume that the response criterion changes for larger ss target times.

A more parsimonious explanation would be that clock mechanisms for the two target ranges differ, resulting in the contrasting overestimation and underestimation of the target ranges, while working memory mechanisms for the two ranges do not differ, resulting in increases in estimates across the target ranges found in chapter 1. This explanation is supported by neuroimaging findings that show activation in parietal and frontal regions, areas related to working memory, for both target ranges (Macar & Vidal, 2009).

The migration effect found in chapter 2 when multiple target times are learned together is not a unique finding (Lejeune & Wearden, 2009), but some argue that it is a result of the reproduction task following the same logic above (Wearden, 2003). However, there is evidence of Vierordt's law from verbal estimation tasks (Lejeune & Wearden, 2009), although these tasks tended to include ms and ss target times, confounding if the effect was due to migration or differences in how these ranges are processed. We argue that the migration is a result of multiple target times being learned together, independent of the target ranges being presented, and that participants spontaneously organize the target times into a category. In general, the nature of the

reproduction task easily allows for the presentation of multiple intervals, because estimates are typically made immediately after presentation, as seen in chapter 1. In contrast, with discrimination tasks, participants learn one or two standards, where the two standards are presented relative to each other (e.g. one is short, and one is long) and are then presented with a number of comparison durations. Migration effects wouldn't be predicted when there is only one standard target time and have not been assessed when two standards are presented in a discrimination task. Theoretically, the comparator times used in the discrimination task could influence the memory representation of the standard, if participants are categorizing them together. The common use of multiple target times in the reproduction task would explain why Vierordt's law tends to be reported more often in reproduction tasks.

However, discrimination tasks could be used to test for the Vierordt-like migration effects by utilizing the shape-target time association we developed in chapter 2. For example, three different standard times would be presented, each associated with a specific shape. Comparison times would be presented with the shape that corresponded to the target time that participants should make the comparison to. Comparison gradients would be compared across conditions in which a single standard time was learned with conditions in which multiple standard times were learned. Additionally, changing the range and quantity of comparators could also influence the memory representation of the standard, providing evidence that intervals presented together influence the memory representation of each other, which we argue is the mechanism underlying Vierordt's law.

In chapter 3 (Implicit and Explicit Timing) we designed a paradigm to target the implicit learning of time, and found that the design allowed for temporal estimates of target times when specifically prompted for them at the end of the experiment. If the estimates were made using

retrospective timing mechanisms, then this task could be used to study retrospective timing further. Typically, retrospective judgments are obtained after participants complete an unrelated task. One of the reasons that retrospective timing has been relatively less studied than prospective timing is because of the difficulty of obtaining estimates for different target times from the same participants, because once a single judgment has been requested participants are aware of the relevance of the temporal domain. Grondin & Plourde (2007) were able to get multiple estimates from a single participant by requiring participants to complete different tasks for different target times and give retrospective timing judgments when all the tasks were completed. A unique aspect to our design is that we were able to obtain multiple estimates from a single participant, while the participant completed the same task, categorizing Mr. Potato Head™, for both target times. Additionally, our task allowed us to use shorter target times than is typical in retrospective tasks. In order to give participants time to complete the non-timing task, studied durations tend to be in the minutes range (e.g. 2-4 min for Grondin & Plourde, 2007). Our task allowed us to assess estimates for 1.67s and 3.33s. This task could potentially be used to assess even shorter target times (e.g. ms) to see if there are range differences in retrospective timing. Finally, we could incorporate a long delay between the categorization phase and the retrospective judgments, perhaps filled with a distracter task shown not to interfere with timing, to see how retrospective estimates may change at a later recall.

An alternate explanation is that the presentation durations were learned implicitly, but not used to complete the categorization task. Therefore, participants could use the implicitly learned memories to give estimates after the task. Future experiments should focus on isolating the implicit learning of time. A potential experiment could use a more realistic example that would combine multiple perceptual domains. As mentioned previously, participants could be told that

they need to categorize the significance of calls produced by artificially created animals and be presented with different types of artificially constructed animals making different types of sounds for different lengths of time. This would allow for multi-modal categorization (across the visual, auditory, and temporal domains), and has a straightforward real world analogy that would increase the likelihood that participants would use the implicitly learned information to complete the task.

In conclusion, these three chapters have demonstrated a number of important findings. Task type is very important for understanding different memory systems in time perception. Discrimination tasks do not facilitate an understanding of the mechanisms underlying Vierordt's law the way a reproduction task does because discrimination tasks do not require participants to learn multiple time intervals simultaneously. Discrimination tasks demonstrate subjective shortening but reproduction tasks demonstrate lengthening when working memory processes are targeted, likely because these tasks are analogous to recognition and recall tasks found in non-temporal memory studies. These studies have also shown that differences in processing ms and ss target times lead to differences in estimates, but these differences are independent from effects of working memory. Estimates for both target ranges will increase with an increasing delay, even though ms target times are consistently overestimated and ss target times are consistently underestimated. Additionally the overestimation of ms target times and the underestimation of ss target times is independent of migration effects found when multiple target times are learned together. Finally, the last study, showed us that participants can learn a duration implicitly, but do not apply that information to the categorization task. In general, these three chapters have demonstrated that many of the approaches used to understand different memory systems can be applied to understanding the memory of time.

## References

- Allan, L.G. (1977). The time-order error in judgements of duration. *Canadian Journal of Psychology, 31*, 24-31.
- Allan, L.G. (1979). The perception of time. *Perception & Psychophysics, 26*, 340-354.
- Allan, L.G. (1998). The influence of the scalar timing model on human timing research. *Behavioral Processes, 44*, 101-117.
- Allan, L.G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning & Motivation, 22*, 39-58.
- Allen, S.W., & Brooks, L.R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120*, 3-19.
- Ashby, F.G., & Maddox, W.T. (2005). Human category learning. *Annual Review Psychology, 56*, 149-178.
- Bangert, A., Reuter-Lorenz, P.A., & Seidler, R.D. (2011). Dissecting the clock: Understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychologica, 136*, 20-34.
- Block, R.A. (1992). Prospective & retrospective duration judgment: The role of information processing and memory. In F. Macar, Pouthas, V., Friedman, W.J. (Eds.) *Time, Action, and Cognition: Towards Bridging the Gap* (pp. 141-152). Kluwer Academic Publishers, Dordrecht, the Netherlands.
- Bobko, D.J., Schiffman, H.R., Castino, R.J., & Chiapetta, W. (1977). Contextual effects in duration experience. *American Journal of Psychology, 90*, 577-586.
- Brady, T.F., & Alvarez, G.A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science, 22*, 384-392.

- Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.
- Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perceptual Psychophysics*, *59*, 1118–1140.
- Brown, S.W. (2008). Time and attention: Review of literature. In S. Grondin (Ed.), *Psychology of time*. Bingley, U.K.: Emerald Group, pp. 111-138.
- Brown, S. W., Newcomb, D. C., & Kahrl, K. G. (1995). Temporal signal detection and individual differences in timing. *Perception*, *24*, 525-538.
- Church, R. M. (1984). Properties of the internal clock. *Annals of the New York Academy of Sciences*, *423*, 566-582.
- Church R.M. (2006). Time Perception, Introductory article. *Encyclopedia of Cognitive Science* (pp. 1-3). John Wiley & Sons, Ltd.
- Coull, J.T., Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, *303*, 1506–1508.
- Dorrance, B.R., Kaiser, D.H., & Zentall, T.R. (2000). Event-duration discrimination by pigeons: The choose-short effect may result from retention-test novelty. *Animal Learning & Behavior*, *28*, 344-353.
- Fetterman, J. G., & Killeen, P. R. (1990). A componential analysis of pacemaker-counter timing systems. *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 766-780.
- Fetterman, J. G., & Killeen, P. R. (1995). Categorical scaling of time: Implications for clock-counter models. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 43-63.
- Fortin, C., & Breton, R. (1995). Temporal interval production and processing in working memory. *Perception & Psychophysics*, *56*, 203-215.



- Fortin C, & Couture E. (2002). Short-term memory and time estimation: Beyond the 2-second “critical” value. *Canadian Journal of Experimental Psychology*, 56,120-127.
- Fortin, C., Rousseau, R., Bourque, P. E., & Kirouac, E. (1993). Time estimation and concurrent nontemporal processing: Specific interference from short-term-memory demands. *Perception & Psychophysics*, 53, 536-548.
- Franssen, V., Vandierendonck, A., & Van Hiel, A. (2006). Duration estimation and the phonological loop: Articulatory suppression and irrelevant sounds. *Psychological Research*, 70, 304-316.
- Gibbon, J. (1977). Scalar expectancy theory and Weber’s law in animal timing. *Psychological Review*, 84, 279-325.
- Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology*, 35, 283-293.
- 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). Hillsdale, NJ: Erlbaum.
- Gibbon J, Church RM, & Meck WH. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences* 423, 52–77.
- Goldstone, S., Lhamon, W.T., & Boardman, W.K. (1957). The time sense: Anchor effects and apparent duration. *Journal of Psychology*, 44, 145-153.
- Gooch, C.M., Balsam, P.D., Taylor, K., Stern, Y., & Rakitin, B.C. (in revision). Effects of number of standards and retention interval on temporal memory.
- Grondin S. (2005). Overloading Temporal Memory. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 869-879.

- Grondin, S., & Plourde, M. (2007). Judging multi-minute intervals retrospectively. *Quarterly Journal of Experimental Psychology*, *60*, 1303-1312.
- Harrington, D.L., Haaland, K.Y., & Knight, R.T. (1998). Cortical networks underlying mechanisms of time perception. *Journal of Neuroscience*, *18*, 1085-1095.
- Hicks, R.E., Miller, G.W., & Kinsbourne, M. (1976). Prospective & retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology*, *89*, 719-730.
- Ivry R. (1996). The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology*, *6*, 851-857.
- Ivry, R.B., & Hazeltine, R.E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 3-18.
- Ivry R.,B., & Keele, S.,W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*, 136-52.
- Jones, C.R., Malone, T.J., Dirnberger, G., Edwards, M., & Jahanshahi, M. (2008). Basal ganglia, dopamine, and temporal processing: Performance on three timing tasks on and off medication in Parkinson's Disease. *Brain & Cognition*, *68*, 30-41.
- Jones, C.R., Rosenkrantz, K., Rothwell, J.C., & Jahanshahi, M. (2004). The right dorsolateral prefrontal cortex is essential in time reproduction: An investigation with repetitive transcranial magnetic stimulation. *Experimental Brain Research*, *158*, 366-372.
- Jones, L.A., & Wearden, J.H. (2003). More is not necessarily better: Examining the nature of the temporal reference memory component in timing. *Quarterly Journal of Experimental Psychology*, *56*, 321-343.

- Jonides, J., Lewis, R.L., Nee, D.E., Lustig, C.A., Berman, M.G., & Moore, K.S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 1-32.
- Koch, G., Oliveri, M., Torriero, S., & Caltagirone, C. (2003). Underestimation of time perception after repetitive transcranial magnetic stimulation. *Neurology*, *60*, 1844-1846.
- Koch, G., Oliveri, M., Torriero, S., Salerno, S., Gerfo, E.L., & Caltagirone, C. (2007). Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research*, *197*, 291-299.
- Knowlton, B.J., & Squire, L.R. (1993). The learning of categories: Parallel brain systems for item memory and category knowledge. *Science*, *252*, 1747-1749.
- Lejeune H. (1998). Switching or gating? The attentional challenge in cognitive models of psychological time. *Behavioral Processes*, *44*, 127–145.
- Lejeune, H., & Wearden, J.H. (2009). Vierordt's *The Experimental Study of the Time Sense* (1868) and its legacy. *European Journal of Cognitive Psychology*, *21*, 941-960.
- Lewis, P.A., & Miall, R.C. (2003a). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current Opinion in Neurobiology*, *13*, 250-255.
- Lewis, P. A., & Miall, R. C. (2003b). Overview: An image of human neural timing. In Meck, W. H. (Ed.). *Functional and neural mechanisms of interval timing* (pp. 515-532). CRC Press, Boca Raton, FL.
- Lewis, P.A. & Miall, R.C. (2006). Remembering the time: A continuous clock. *Trends in Cognitive Sciences*, *10*, 401-406.
- Macar, F., Grondin, S., & Casini, L. (1994). Controlled attention sharing influences time estimation. *Memory & Cognition*, *22*, 673–686.

- Macar, F., & Vidal, F. (2009). Timing Processes: An outline of behavioral and neural indices not systematically considered in timing models. *Canadian Journal of Experimental Psychology*, *63*, 227-239.
- Malapani, C., Deweer, B., & Gibbon, J. (2002). Separating storage from retrieval dysfunction of temporal memory in parkinson's disease. *Journal of Cognitive Neuroscience*, *14*, 311-322.
- Mangels, J.A., Ivry, R.B., & Shimizu, N. (1998). Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Cognitive Brain Research*, *7*, 15-39.
- Matell, M. S., & Meck, W.H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research*, *21*, 139-170.
- Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive Brain Research*, *3*, 227-242.
- Nobre A.C., & O'Reilly J. (2004). Time is of the essence. *Trends in Cognitive Science*, *8*, 387-389.
- Ogden, R.S., Wearden, J.H., & Jones, L.A. (2008). The remembrance of times past: Interference in temporal reference. *Journal of Exp. Psychology: Human Perception & Performance*, *34*, 1524-1544.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437-442.
- Penton-Voak, I.S., Edwards, H., Percival, A., & Wearden, J.H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology*, *22*, 307-320.
- Poldrack, R.A., & Foerde, K. (2008). Category learning and the memory systems debate. *Neuroscience & Biobehavioral Reviews*, *22*, 197-205.

- Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., & Meck, W. H. (1998). Peak-interval timing in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 19-33.
- Rakitin, B.C., Stern, Y., & Malapani, C. (2005). The effects of aging on time reproduction in delayed free-recall. *Brain & Cognition*, *58*, 17-34.
- Schiffman, H.R., & Bobko, D.J. (1974). Effects of stimulus complexity of the perception of brief temporal intervals. *Journal of Experimental Psychology*, *103*, 156-159.
- Shea-Brown, E., Rinzel, J., Rakitin, B.C., & Malapani, C. (2006). A firing rate model of Parkinsonian deficits in interval timing. *Brain Research*, *1070*, 189-201.
- Shohamy, D., Myers, C.E., Kalanithi, J., & Gluck, M.A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience and Biobehavioral Review*, *32*, 219-236.
- Smith, E.E. (2008). The case for implicit category learning. *Cognitive, Affective, & Behavioral Neuroscience*, *8*, 3-16.
- Smith, E.E., & Grossman, M. (2007). Multiple systems of category learning. *Neuroscience & Biobehavioral Reviews*, *32*, 249-264.
- Smith, E.E., & Jonides, J., (1991). Storage and executive processes in the frontal lobes. *Science*, *283*, 1657-1661.
- Spetch, M.L., & Rusak, B., (1992). Temporal context effects in pigeons' memory for event duration. *Learning and Motivation*, *23*, 117-144.
- Spetch, M.L., & Wilkie, D.M., (1983). Subjective shortening: A model of pigeons' memory for event duration. *Journal Experimental Psychology: Animal Behavioural Processes*, *9*, 14-30.
- Staddon, J. E. R., & Higa, J. J. (1999). Time and memory: Towards a pace-maker free model of

interval timing. *Journal of the Experimental Analysis of Behavior*, 71, 215-251.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs*, 77, 1-31.

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception*, 19, 705-748.

Van Volkinburg, H., Balsam, P., & Malapani, C. (unpublished). Schizophrenia and time perception.

Wearden, J.H. (1999). Beyond the fields we know...: Exploring and developing scalar timing theory. *Behavioural Processes*, 45, 3-21.

Wearden, J. H. (2003a). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and mind II: Information Processing Perspectives* (pp. 21–39). Hogrefe & Huber Publishers, Gottingen, Germany.

Wearden, J.H. (2003b). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.) *Time and Mind II: Information Processing Perspectives* (pp. 21-39). Hogrefe & Huber Publishers, Gottingen, Germany.

Wearden, J.H., & Culpin, V. (1998). Exploring scalar timing theory with human subjects. In DeKeyser, V., d'Ydewalle, G., Vandierendonck, A. (Eds.) *Time and the Dynamic Control of Behavior* (pp. 33-49). Hogrefe & Huber Publishers, Gottingen, Germany.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology B*, 51, 97-120.

Wearden, J. H. & Ferrara, A. (1993). Subjective shortening in humans' memory for stimulus

- duration. *The Quarterly Journal of Experimental Psychology Section B*, *46*, 163-186.
- Wearden, J.H. & Jones, L.A. (2007). Is the growth of subjective time in humans a linear or nonlinear function of real time? *Quarterly Journal of Experimental Psychology*, *60*, 1289-1302.
- Wearden, J.H. & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal Of Experimental Psychology*, *61*, 569-587.
- Wearden, J.H., Parry, A., & Stamp, L. (2002). Is subjective shortening in human memory unique to time representations? *Quarterly Journal of Experimental Psychology*, *55*, 1-25.
- Wearden, J. H., & Penton-Voak, I. S. (1995). Feeling the heat: Body temperature and the rate of subjective time, revisited. *Quarterly Journal of Experimental Psychology B*, *48*, 129-141.
- Wearden, J.H., Piklinton, R., & Carter, E. (1999). 'Subjective lengthening' during repeated testing of a simple temporal discrimination. *Behavioral Processes*, *46*, 25-38.
- Wiener, M., Turkeltaub, P., & Coslett, H.B. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage*, *49*, 1728-1740.
- Yarmey, A.D. (2000). Retrospective duration estimations for variant and invariant events in field situations. *Applied Cognitive Psychology*, *14*, 45-57.
- Zakay, D. (1993). Relative and absolute duration judgments under prospective and retrospective paradigms. *Perception and Psychophysics*, *54*, 656-664.
- Zakay, D. (1998). Attention allocation policy influence prospective timing. *Psychonomic Bulletin & Review*, *5*, 114-118.
- Zakay D, & Block RA, (1996). The role of attention in time estimation processes. In Pastor, MA, Artieda, J.(eds). *Time, Internal Clocks and Movement* (pp 143–164). Elsevier Sciences: New York, New York.

Zakay, D., & Block, R.A. (1997). Temporal Cognition. *Current Directions in Psychological Science*, 6, 12-16.

Zakay, D., & Block, R.A. (2004). Prospective and retrospective duration judgments: An executive-control perspective. *Acta Neurobiologica Experimental*, 64, 319-328.

Zentall, T.R., (2006). Timing, memory for intervals, and memory for untimed stimuli: The role of instructional ambiguity. *Behavioural Processes*, 71, 88-97.