Timing Everyday Tasks and Events

by

Janel Fergusson

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2017

Abstract

Many of the tasks we complete every day require us to attend to the passing of time or to use time information in some way. Everyday tasks frequently require us to use time information in a strategic, deliberate, and explicit way, such as when we wish to steep a cup of tea for 3 minutes or must leave to meet a colleague in 5 minutes. Much of the previous research on timing has used very short duration tasks, in the range of seconds. Several models have been developed to account for timing in short duration tasks, but it is not known which model(s) best fit timing of everyday intervals. This dissertation research was designed to determine whether the AGM or memory storage models can account for timing in the range of everyday intervals. Experiment 1 investigated the underlying pattern of estimates in the range of 1-5 minutes. Participants produced underestimates for all intervals. Experiments 2 and 3 used a framing manipulation to investigate the role of memory chunking in timing everyday durations. Participants who were instructed to focus on the present task produced intervals that were longer than participants who were instructed to focus on a future task, consistent with the idea that a present frame serves as a better organizing structure for the interval and results in fewer chunks. According to memory storage models, the number of items in memory is compared to a stored value in reference memory to determine how much time has passed. When stimuli are organized more cohesively, resulting in fewer chunks, it takes a longer time to accumulate the target number of chunks. Experiment 4 used feedback and an attentional manipulation to investigate the role of reference memory and attention to timing. Participants produced more accurate intervals following feedback, consistent with the idea that reference memory for everyday intervals is inaccurate. Attention had no effect on estimate duration, which is inconsistent with the AGM. Taken together, the results from these experiments suggest that memory storage models are a better fit for timing in the range of everyday intervals.

Lay Summary

When you steep a cup of tea for 3 minutes, do you set a timer, or do you try to keep track of the time without one? When you try to keep track of time without a clock, how accurate are you? What factors influence your accuracy? Most of the previous research on timing has focused on very short intervals, so we know very little about how we perform on timing tasks in the range of minutes. The research reported in this dissertation was conducted to determine how accurate people are in timing these types of tasks and whether memory or attention plays a role in our performance. Results suggest that memory, both for what happened during the interval and for how long the interval should be, plays a role in performance.

Preface

This dissertation is an original intellectual product of the author, Janel Fergusson. The experiments reported in Chapters 2- 5 were covered by UBC Ethics Certificate number H03-80566.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor, Dr. Peter Graf, for his many comments, suggestions, and words of encouragement throughout my PhD program and in preparing this manuscript. I would also like to thank my committee members, Dr. James Enns and Dr. Eric Eich, for their insightful and challenging questions and comments. My labmates, Michelle Crease Lark, Natasha Pestonji, and Anna Maslany provided much needed encouragement and commiseration throughout the process of conducting my dissertation research and preparing this manuscript.

I must also extend my deepest thanks to my family and friends, near and far. Ever patient and encouraging, I could not have done this without them.

Chapter 1 – Timing Everyday Tasks and Events

Timing, defined as the process of sequencing and coordinating the elementary components or stages of an event or activity, is required for many common tasks. Some tasks have very brief durations, such as when we swing a racquet to meet an oncoming ball, or when we reach out to shake a colleague's hand. Other tasks have very long durations, such as when we need to pay rent on time or to submit an abstract for a conference by the end of the month. My dissertation focuses on tasks that fall between these two extremes, that is, on everyday intermediate duration tasks such as steeping a cup of tea or leaving the room during a commercial break and then returning on time. The overall goal of my dissertation is to identify the cognitive processes required for timing these intermediate duration tasks, which fall in the range of 2 to 6 minutes.

Intermediate duration tasks are interesting because of the conscious, deliberate way we time their completion. For very short duration tasks, timing seems to be conducted automatically, without much conscious reflection. Swinging a racquet to meet a ball is done automatically, without consciously thinking about how much time we have for completing the action. For very long duration tasks, such as paying rent on time, we do not attempt to monitor the entire interval consciously and deliberately. Instead, we make a plan to complete the task at a future time and then typically rely on external cues to execute the task. In contrast, everyday tasks such as steeping a perfect cup of tea tend to require us to consciously and deliberately direct our attention towards timing and to make a response when a desired interval has elapsed. We resolve to time the interval, maintain conscious attention on the passage of time, and then respond when we think the interval has elapsed. This conscious attention to the passing of time sets timing intermediate duration tasks apart from timing shorter and longer duration tasks.

Intermediate duration tasks are also interesting because we seem to be confident in our ability to time them. One possible reason for our fairly high confidence is the presence of feedback, both during and at the conclusion of the task. For tasks such as steeping a cup of tea, we can make use of feedback such as colour or odour in addition to our internal sense of how much time has passed. Another possible reason for this high confidence may be that the exact duration of such tasks is not always critical. When steeping a cup of tea, it does not make much difference in flavour if the tea bag remains in the water for 3.75 minutes or 4 minutes. For other tasks, however, such as the last few minutes of an exam after the instructor gives a 5 minute warning, timing the exact duration is more important. It is possible that we feel confident in timing intermediate duration tasks because we have a lot of experience in completing them with acceptable accuracy. Yet another possibility is that a combination of feedback and experience drives our high confidence. When learning how long to steep a cup of tea, it is easy to attach time markers to the darkness of the tea, eventually using the shade of the tea as a proxy for duration.

My motivation to focus on the timing of intermediate duration tasks also stems from the fact that we know so little about them. Despite our confidence in timing intermediate duration tasks, we have no evidence to justify this confidence. Are we capable of accurately timing tasks in the range of 2-6 minutes, or is this confidence misplaced? To my knowledge, only three studies have addressed timing of intermediate duration tasks or events. The first of these studies, by Poynter (1983), used a retrospective interval comparison task in which naïve participants were exposed to an interval of 170 seconds while engaged in a secondary memory task that was either continuous or segmented, that is, broken up into distinct groups of words. Participants were not informed that they would be asked about the duration of the interval until after they had been exposed to it, and then they were instructed to estimate its duration. For this purpose, they were exposed to a 30-second sample interval. Following the target interval and sample interval, participants were asked to mark a line to indicate how long the target interval was in relation to the sample interval. In a second experiment, Poynter used the same procedure for timing a target event lasting 225 seconds, with a sample interval of 45 seconds. Poynter found that participants produced shorter estimates when the secondary task was continuous rather than segmented. Poynter reported that participants underestimated the target in both experiments, but did not report the amount by which the intervals were underestimated. Moreover, it is unclear whether this pattern of (under)estimates is specific to the time-estimation task Poynter's participants had to complete. Equally important, Poynter used a retrospective timing task (i.e., a task in which participants do not know at the outset that they would be asked to estimate duration), thus raising questions such as: Would participants also underestimate the duration of prospective tasks (i.e., when they anticipate having to estimate the duration of a task or event ahead of time)? Do participants underestimate when having to make verbal estimates about the duration of an event? Is the accuracy of participants' estimates correlated with the duration of events (e.g., 4-6 minutes)?

In the second previous study on the timing of intermediate duration tasks, Block and Reed (1978) had participants estimate the duration of intervals from 60, 120, and 300 seconds, while engaged in a secondary task that used either unmixed deep processing ("count the number of words that are a part of a building"), unmixed shallow processing (e.g., "count the number of words that use capital italic type"), or mixed processing (half of blocks with deep processing instructions, half with shallow processing instructions). Participants in the mixed condition produced longer estimates than participants in either the unmixed deep or unmixed shallow conditions. Data regarding the precise length of the estimates were not reported, so it is not possible to determine the size of the underestimates.

Block (1992) had participants reproduce durations of 165 seconds while engaged in a secondary, attention-demanding task, under either prospective or retrospective conditions. In prospective reproduction tasks, participants are aware from the onset of the sample interval that they will have to reproduce it. In retrospective reproduction tasks, naïve participants experience the sample interval and are then asked to reproduce it. Block reported that participants' reproductions were significantly shorter than the target of 165 seconds in both the easy and difficult ongoing task conditions. Participants in the prospective condition with the easy ongoing task responded, on average, 23.2 seconds before the target of 165 seconds, while participants in the prospective condition with the difficult ongoing task responded, on average, 34.8 seconds before the target of 165 seconds. Underestimates in the retrospective conditions were longer than those in the prospective condition, but task difficulty did not significantly affect reproductions. Block's study extends the results by Poynter (1983) to reproduction tasks, as well as to prospective timing tasks. Nevertheless, Block's study leaves several questions unanswered, such as: Do participants also underestimate the duration of production tasks? Does the underestimation habit extend to longer intermediate duration tasks? Does the difficulty of the secondary task influence the duration of participants' time estimates?

Block and Reed (1978), Block (1992), and Poynter (1983) interpreted their findings in light of existing models of timing. Specifically, Poynter interpreted his findings as supporting memory storage models (Ornstein, 1969), which are described in more detail below. Such models predict that estimates would be longer when there is more segmentation in the task, and this was confirmed by Poynter. Block and Reed interpreted their results in line with contextual change models (Block, 1978; Block & Reed, 1978), close relatives to memory change models, which predict that the number of changes in context (e.g., switches between deep and shallow processing) in an interval is positively related to the length of the estimate. Block's finding that prospective reproductions were longer when participants were engaged in the easy ongoing task was also interpreted in line with contextual change models (Block, 1978; Block & Reed, 1978). Given Block's findings on the prospective underestimation of 165-second intervals and the similarity of this outcome with research on the prospective timing of shorter intervals, it might be the case that an existing timing model, developed to account for timing in the range of seconds, can also account for timing in the range of minutes. Block's finding of the effect of attention on prospective time estimates can also be explained by the Attentional Gate Model (AGM) (Zakay & Block, 1996, 1998), a now dominant model in the literature on timing short durations, which I will discuss in greater detail later in this chapter.

The bulk of this chapter is devoted to a review of the available timing models with a focus on identifying a model or models that might give insight into the mechanism used for deliberate, conscious timing. Is there a model developed to explain timing of shorter duration tasks that can also explain the timing of intermediate duration tasks? Do intermediate duration, everyday tasks like steeping a cup of tea and short duration timing like swinging a racquet or crossing a street rely on the same cognitive mechanism? In contrast to short duration tasks, everyday intermediate duration tasks involve conscious, deliberate, and strategic timing. It seems unlikely that the same mechanism could be responsible for tasks that both involve conscious and deliberate timing and those that do not. My goal is to focus on identifying a model or models which can potentially account for performance on prospective production and estimation tasks in the intermediate range of minutes and which make predictions regarding accuracy. In order to identify such models, I will review the common features that many timing models share. I will then describe the existing models of timing and examine how suitable they are for explaining timing in the range of minutes.

Criteria for Evaluating Models

The main goal of my dissertation research was to increase our understanding of the mechanisms involved in timing everyday intermediate duration tasks. For this purpose, I needed to identify a model that is able to explain our ability to consciously, deliberately, and strategically time everyday events in the range of minutes. This conscious and deliberate manner in which we time everyday intermediate duration tasks is the most important feature that distinguishes them from short and long duration tasks. A suitable model must be able to account for prospective, deliberate timing, and it must include a role for attention and strategic control, while allowing participants to explicitly reference minutes/seconds when making time judgments.

Many of our everyday intermediate duration timing tasks are prospective in nature, and thus, in my review, I give special attention to models that can account for prospective timing. For prospective tasks, we are informed prior to beginning the task that we will be required to produce or reproduce its duration. Evidence from multiple studies suggests that prospective and retrospective judgments are significantly different in a number of ways (cf. Block & Zakay, 1997; Block, Hancock, & Zakay, 2010; Zakay & Block, 1997). As my own research focuses on the use of time in a prospective manner, rather than on retrospective timing, I will focus on models that are able to provide an account for prospective timing.

Many intermediate duration tasks require us to provide explicit information about minutes and seconds (e.g., I need to leave for a meeting in exactly 5 minutes; I instruct a colleague to meet me in 3 minutes). For this reason, my review will seek out timing models that include a mechanism for translating time durations into declarative statements.

Everyday timing tasks often require us to explicitly make reference to minutes and seconds. When steeping a cup of tea for 3 minutes, we must explicitly make reference to the duration. This is in contrast with other types of tasks which require us to determine which of two intervals is or was longer (or shorter), without explicit reference to time units (e.g., minutes and seconds). A suitable model must include some mechanism for making these explicit judgments (e.g., how many minutes has it been) that reference minutes and seconds, and not solely comparison judgments (e.g., which of these two stimuli was longer?). A model that can only account for discrimination between two intervals, with no reference to long term memory for intervals learned in the past and verbal labels, cannot explain participants` ability to make explicit judgements about how much time has passed.

A suitable model also must have a role for attention and strategic control, that is, it must recognize that attention to time is under voluntary control, and that the amount of attention dedicated to timing influences the length or accuracy of temporal estimates. “Time flies when we’re having fun” and “a watched pot never boils” both suggest that we are aware that our subjective experience of time is heavily influenced by concurrent tasks. When the concurrent task is interesting and engaging, less attention is paid to time, and it appears to fly by. When we are waiting for something to happen, time slows to a crawl. When a participant is required to engage in two concurrent tasks, the amount of attention that is required by the secondary, non-temporal task influences the length of temporal estimates (e.g., Brown, 2006; Zakay, Nitzan, & Glicksohn, 1983).

An Overview of Timing Models

There is great diversity among existing timing models, probably because timing tasks used in the laboratory, like tasks in everyday life, come in a huge variety. Timing tasks can be divided into five types: production, estimation, reproduction, discrimination, and bisection. Of these tasks, production and estimation are most frequently used in investigating longer intervals, and more closely correspond to everyday, real life timing tasks in this range of intervals. Temporal production tasks require making a response to indicate that a given interval (e.g., 30 seconds) has elapsed (e.g., Brown, 1997; Esposito et al., 2007). In temporal estimation tasks, participants are exposed to some event (e.g., a sound clip lasting 30 seconds) and then asked to estimate its duration (e.g., Brown, 1985; Hicks, Miller, & Kinsbourne, 1976; Kladopoulos et al., 2004; Zakay & Tsal, 1989).

It is likely that the distinction between prospective and retrospective tasks also contributed to the variety of models developed to account for human timing (Grondin & Macar, 1992). In prospective timing tasks, participants are instructed to complete some sort of timing task before the task begins (Block & Gruber, 2014; Block & Zakay, 1997; Brown, 1985); for example, a participant might be instructed to press the spacebar when they believe 2 minutes have elapsed. For such tasks, participants are aware that they have to estimate or reproduce an interval before that interval is presented. By contrast, retrospective timing tasks are typically done with participants who are unaware that they will be asked to estimate or reproduce an interval. Participants are engaged in a secondary task in some way for a desired period of time (they are placed into a dual task condition), and then informed that they are to estimate or reproduce its duration.

Prospective and retrospective timing differ in more than purely methodological ways. Prospective timing appears to involve attentional resources, and prospective estimates, productions, and reproductions appear to be sensitive to the amount of simultaneous nontemporal processing (Brown, 1985). In contrast, retrospective timing is less sensitive to the effects of concurrent task demands, possibly because accurate retrospective timing is more reliant on the memory for information that was passively or automatically encoded (Brown, 1985).

These distinctions between the nature of the response (production, estimation, etc.) and the prior knowledge of the participant (prospective vs. retrospective) have resulted in a large number of timing tasks used in laboratory research. The need to account for the evidence collected with these tasks has inspired the rich variety of timing models currently in existence. My dissertation research employs only a subset of the available tasks, sometimes called prospective temporal production tasks, and thus the models created for explaining performance on such tasks are featured in my review.

The rich diversity among existing timing models is also the result of efforts to explain distinct aspects of timing. Some models of timing are focused primarily on participants’ ability to produce a given interval, whereas others are more concerned with predicting performance on reproduction or bisection tasks. Models that are focused on explaining reproduction or bisection task performance alone typically do not include a mechanism for translating duration into explicit judgments of minutes and seconds. Some models of timing focus on the length or accuracy of estimates, while others focus on the variability between multiple estimates of the same duration (e.g., Allan, 1979). My dissertation research is primarily focused on explaining the length or accuracy of estimates. Intermediate duration tasks often require us to produce a single response, rather than a number of consistent responses, and so accuracy is most important. The models created for explaining participants' performance on production tasks and that focus on length or accuracy of timing will be featured in my review.

A further distinction among existing models is the nature of the mechanism proposed for timing, and more specifically, whether timing is the product of a special purpose internal clock or similar device, or is a by-product of another cognitive process that is used as a signal for timing. Many models of timing make reference to a "mechanism" that is used to time intervals. In its broadest sense, a mechanism is simply the process by which something occurs, and timing models use this term to refer to the cognitive and/or biological processes that give rise to our ability to estimate time. In the models discussed below, mechanism is used in a broad sense, without explicit reference to particular brain structures or neural networks. Instead, mechanism is used to refer to a theoretical construct that is proposed to account for behavioural data. The use of the term mechanism varies in specificity from model to model. In many models the proposed mechanism is a dedicated, specific structure or process that is solely or primarily involved in timing intervals. In other models the term mechanism is used more generally, to describe the particular way in which cognitive processing or resources are recruited for timing.

Types of Event Generators

There are two main components of the proposed timing mechanism that nearly all timing models share, namely, that they describe some sort of event generator, which creates a signal of some sort, and an event transducer, which translates this signal into information about time. Event generators can be grouped into two main categories: those for which the signal is a by-product of other cognitive processing (e.g., the amount of information stored in memory, the number of contextual changes, psychological “moments” of some type), and those for which the signal is generated by a mechanism dedicated to timing. Dedicated event generators typically take the form of some sort of mechanism which emits pulses or signals at regularly spaced intervals, though the spacing of these pulses and whether or not the rate varies as a result of cognitive or physiological factors differs from model to model. Event transducers, mechanisms for translating the signal emitted by the event generator into information about time, are more varied, and depend on the exact nature of the signal received from the event generator. Some event transducers include a mechanism for explicitly labelling intervals or referencing long term memory for past intervals, and some only allow comparison to an immediately preceding interval, with no ability to label or make comparisons with information stored in long term memory.

By-product of cognitive processing. In the models described below, timing is an indirect result of processes that are primarily involved in other processing (e.g., processing related to perception or to memory). This class of models can be further sub-divided into those that rely on memory processes and those that rely on a broader range of processes.

Memory models. Ornstein’s memory storage model (1969) and the contextual change model (Block, 1978; Block & Reed, 1978) propose that timing is the result of the summation of the amount of information stored in memory during an interval or the number of contextual changes that occur during an interval. When we are required to make an interval estimate, we use the amount of information stored in episodic memory to reconstruct the duration of the interval. If there is a large amount of information in episodic memory, either because we have been exposed to more stimuli or the information is less well organized into chunks, the interval is judged to be longer than if there is a small amount of information in memory for the duration we are estimating. In order to make prospective judgments, it may be the case that we compare the amount of information stored in memory to some reference value stored in long-term memory that is based on prior experience in order to judge whether the target amount of time has passed.

In both models, information stored in episodic memory is used in lieu of a dedicated event generator. While the exact components required for making these types of comparisons are not described, Figure 1.1 represents a possible organization of these types of models, based on other models that involve reference memory based on past experience.

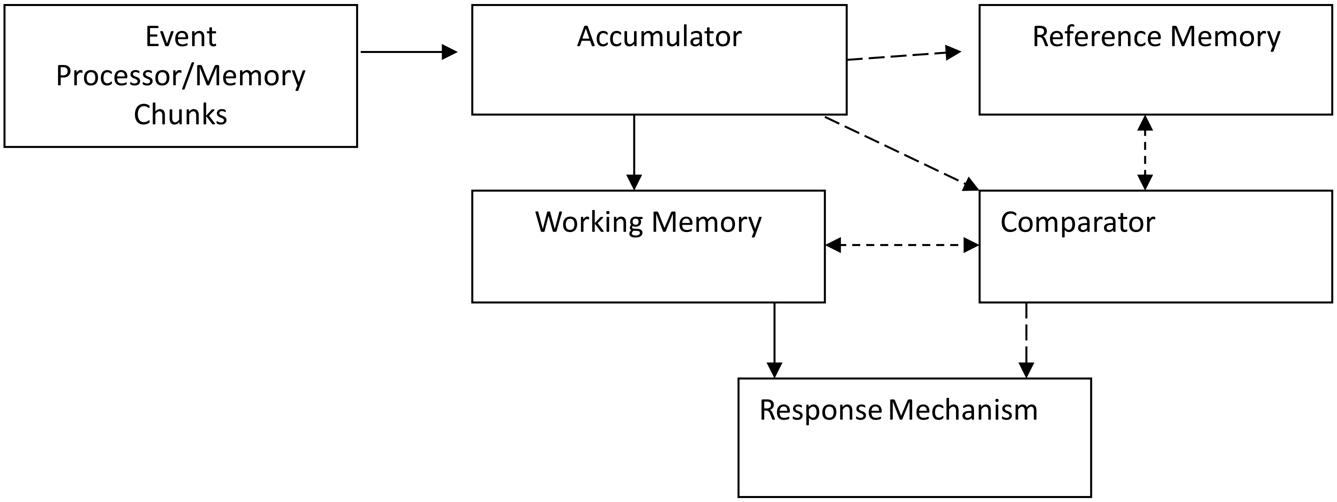

Figure 1.1. Possible organization of memory storage models

Ornstein's memory storage model. Ornstein's memory storage model (1969) is the simplest and oldest of the memory models and forms the basis for later memory models. The memory storage model proposes that our experience of duration is not biologically based (i.e., there is no dedicated internal clock). Ornstein argued that the experience of duration cannot be simply a matter of short-term memory, as the length of an interval often exceeds the limits of short-term memory; long-term episodic memory must be involved. According to the memory storage model, as information enters the sensory registers, it is segmented and then stored in short-term memory, and from there it is passed on to long-term episodic memory. The number of segments stored is then translated into a representation of duration.

More complex tasks result in a larger number of stored segments. The complexity of each task is determined, at least in part, by the participant's experience with the type of information being stored. If, for example, a participant is a chess expert, s/he can store a chess move in memory in a very simple way (e.g., Bd7), resulting in few segments. If the participant is a chess novice, the same move requires more complicated information to be stored (e.g., Bishop moves to the 7th row in the 4th column), resulting in more segments. The segments in short-term memory and long-term episodic memory serve as the raw data for timing in Ornstein's model. It is not clear what mechanism is responsible for segmenting memories.

In order to judge the amount of time that has passed, the participant sums the number of segments stored. The memory storage model predicts that intervals are judged as being longer when more segments have been stored and recalled than when fewer segments have been stored and recalled. According to the memory storage model, the chess expert described above would perceive a chess game as having a shorter duration than the chess novice (Ornstein, 1969). While the evidence seems to support the prediction that an interval that is filled with more information is judged as longer than an interval with less information, this appears to be mainly limited to retrospective judgments (Bueno Martinez, 1994).

Contextual-change model. Like the memory storage model, the contextual-change model asserts that the information stored in episodic long-term memory is the raw data used for timing. Rather than the number of stored memories serving as the raw data for timing (as in Ornstein's model), the contextual change model assumes that what is critical about memory segments is how often participants must switch between types of information or between different tasks (Block, 1978; Block & Reed, 1978; Poynter, 1983). By this view, each switch from one context to another serves as a temporal marker in memory, from which subjective duration is reconstructed. Both switching to a novel task and switching back and forth between two alternating tasks adds to the subjective experience of time. According to the model, an interval that contains more contextual changes will be judged as longer than an interval in which there are fewer contextual changes. Bueno Martinez (1994) describes an experiment in which participants were assigned to process words semantically, structurally, or a mix of semantic and structural processing, while engaged in a temporal reproduction task. The contextual change model predicts that participants in the mixed processing condition would experience more contextual changes, and would therefore have longer reproductions, and this prediction was confirmed by Bueno Martinez (1994).

Other cognitive by-product models. Memory is not the only cognitive process that has been proposed as an event generator. In other cognitive by-product models, the event generator is either described as the by-product of perceptual processing, as in the discrete moment hypothesis (Stroud, 1967) or in the travelling moment hypothesis (Allport, 1968), or as the by-product of subjective awareness, as in the emotional moment model (Craig, 2009).

Perceptual processing models. Two models include generators that rely on our perceptual processing of events: The discrete moment hypothesis and the travelling moment hypothesis. According to the discrete moment hypothesis, perceptual input is segmented into samples, each of which is temporally discrete (i.e., samples do not overlap) (Stroud, 1967). These perceptual samples are referred to as "moments". Each moment is assumed to be very brief, with each second containing approximately 10 psychological moments, though the number is not fixed and varies from 5-20 (Stroud, 1967). Within each moment all information about temporal order is assumed to be lost; as with a single frame of a movie, there is no information about the order of events portrayed. Psychological time appears continuous, rather than disjointed, because the perceptual moments pass by at a rate that is similar to the rate at which we perceive other perceptual input (Stroud, 1967). Much like the frames of a movie, time appears continuous because each moment flashes by faster than we can process it in a discrete manner.

Developed in response to Stroud's discrete moment hypothesis, Allport's (1968) travelling moment hypothesis retained the idea of a psychological or perceptual moment that is separate from outside events and references perceptual input, but made sweeping changes to what constitutes each moment. Rather than each moment being discrete, non-overlapping, and containing no information about the passing of time within the moment, like a single frame of a movie, the travelling moment hypothesis proposes moments which are comprised of a continuous sampling of perceptual input. Rather than regarding moments as something to be summed, Allport regarded each moment as a “period of temporal summation” (Allport, 1968, p. 397). The difference, then, between the discrete moment hypothesis and the travelling moment hypothesis concerns the temporal content of each moment. According to the discrete moment hypothesis, each moment contains no information about temporal order (Stroud, 1967). These moments, like the frames of a movie, give us a sense of temporal order because they are presented in sequence. According to the travelling moment hypothesis, moments do contain information about temporal order, as they are continuously sampled (Allport, 1968).

Emotional moment model. Craig’s emotional moment model (2009) extends his work on subjective awareness to perception of time. Similar to Stroud (1967) and Allport’s (1968) perceptual and moment models, the emotional moment model proposes that timing depends on the creation of units of perceptual experience which are then summed in some manner. Unlike Stroud and Allport’s models, the definition of a moment is broader and more complex in Craig’s account. Both Stroud and Allport describe moments that are created out of perceptual input. The moments in Craig’s model are the result of many types of input.

According to the model, each moment, referred to as a global emotional moment, is computed on the basis of multiple types of input, including motor conditions, environmental conditions, and motivational, social, and cognitive conditions. Integration of the neural activity associated with these types of input is assumed to occur in the anterior insular cortex (Craig, 2002; 2003; 2009). Each moment is a distinct entity which differs from the preceding and following moments. These global emotional moments combine to form a sense of the self existing across time. Craig proposes that the global emotional moments can be added together to construct subjective time, and that this construction is flexible and emotional. Simply stated, when emotions run high or there is a lot of external stimulation, subjective time passes much more quickly than objective time, as emotional moments accumulate quickly. During periods of decreased arousal or interest, subjective time passes much more slowly as global emotional moments are accumulated slowly. Craig's account predicts that activation should be seen in the anterior insular cortex when participants are completing time perception tasks, a pattern that has been found in multiple imaging studies on time perception in the range of sub-seconds to seconds (Craig, 2009).

Evaluation of memory and cognitive models. Neither the memory models nor the cognitive process models have a role for strategy and conscious attention to timing. In both cases, the raw data used for timing are by-products of long-term memory, or sensory segmentation (more broadly defined in the case of the emotional moment model). The generation of timing events may be influenced by outside conditions, emotional states, and arousal level, but no role for deliberate, conscious attention to timing is proposed. Models in this group do not have clearly described transducers; in fact, no model in this group describes a transducer mechanism beyond stating that the raw data described by the model (i.e., segments in memory, perceptual moments, emotional moments) are summed. With no transducer, no mechanism exists for translating this count into explicit time judgments. It is not clear whether any of the cognitive processing models are able to account for prospective timing. Memory storage and contextual-change models have been supported by data from retrospective timing tasks, but data from prospective timing tasks largely fail to support these models (Block, 1992; Grondin, 2001; Hicks et al., 1976). Neither the discrete moment nor the travelling moment model is well supported by evidence. The emotional moment model remains untested. Additionally, models in this group do not have any mechanism for making explicit judgments about time. Without such a mechanism, it is not clear how summed events are translated into a judgment that references minutes or seconds.

Dedicated event generators. In contrast with the models just described, models with dedicated event generators propose that the raw data that are used for timing are produced by a mechanism exclusively dedicated to timing, rather than being the byproduct of some other cognitive process. These event generators produce some sort of event (pulses, counts, oscillations, etc.) that can later be summed and somehow transformed into judgments about time. The rate of event generation, and the impact of cognitive or biological factors on the rate of event generation, varies from model to model. Models in this group tend to have clearly described transducers. The ability to produce explicit judgments of time, one of the criteria described above for a model of everyday timing, seems to rely in most models on the transducer mechanism. In addition, most models in this category suggest that any role for attention and strategy belong in the domain of the transducer. For these reasons, the majority of my evaluation of these models will be reserved for the following section on transducers. Models with dedicated event generators can be divided into 3 groups: pacemakers with a fixed rate, pacemakers with a variable rate, and oscillators.

Fixed rate pacemakers. A pacemaker is a special purpose device whose primary or sole function is to emit pulses at some rate. These pulses serve as the raw data for timing. Fixed rate pacemakers emit pulses at regularly spaced intervals. Fixed rate pacemaker models do not clearly define what a pulse is. Creelman (1962, p. 590) describes pulses as the result of “a large number of independent elements whose time of firing is randomly distributed” but does not describe what the pulses consist of. Abel (1972) later described the pulses in Creelman’s model as random neural firings that occur during the to-be-judged interval. Zakay and Block (1996; 1998) likewise do not describe exactly what makes up the pulses in their model. Any variability in pulse rate is attributed to random error. The event generator, or pacemaker, in the models described below is solely dedicated to generating pulses to be used by the transducer.

One of the earliest pacemaker models was Creelman’s 1962 model. This model was developed primarily to account for our ability to discriminate between intervals, specifically, for the discrimination of duration differences in short speech sounds. According to the model, the event generator consists of a large number of independent components firing at random (Creelman, 1962). The firings of the event generator are counted by a simple counting mechanism. Creelman is unclear about what these independent elements consist of, what their source might be, or how frequently they fire. The number of independent, randomly firing elements is large, resulting in a constant rate of pulse generation that is independent of the person's physiological or external state (Abel; 1972; Creelman, 1962). To understand how this random activity could result in a constant rate of firing, consider the rate of washroom use at a fair. Each attendee requires the washroom at a different time, which is independent of when any other attendee needs the washroom, but because the crowd is large, the rate of people entering the washroom is fairly constant.

Allan, Kristofferson and Wiens (1971), Kinchla (1972), and Killeen and Weiss (1987) have proposed modified versions of Creelman's model. However, none of the modified models include any changes to the event generator; all changes were made to the transducer. As with Creelman's model, Killeen and Weiss', Kinchla's and Allan et al.'s models call for a pacemaker, with pulses emitted at a constant rate by a large number of randomly firing elements. Like Creelman's model, the rate of pulse generation is unaffected by changes in physiological arousal, stimuli characteristics, or other external factors. According to the attentional gate model (AGM), timing makes use of a pacemaker, which emits pulses at regular intervals, and is uninfluenced by physiological arousal, stimuli characteristics, or other external factors (Zakay & Block, 1996, 1998) (see Figure 1.1). This event generator, or pacemaker, fires at a constant rate, whether or not the participant is engaged in any type of timing activity. The model also assumes that when a signal indicating a need for timing is received, a switch between the event generator and the transducer closes, allowing pulses to pass through. In addition to the switch, which acts in an all-or-none fashion, a gate which is controlled by attention to time modulates the number of pulses that are received by the transducer.

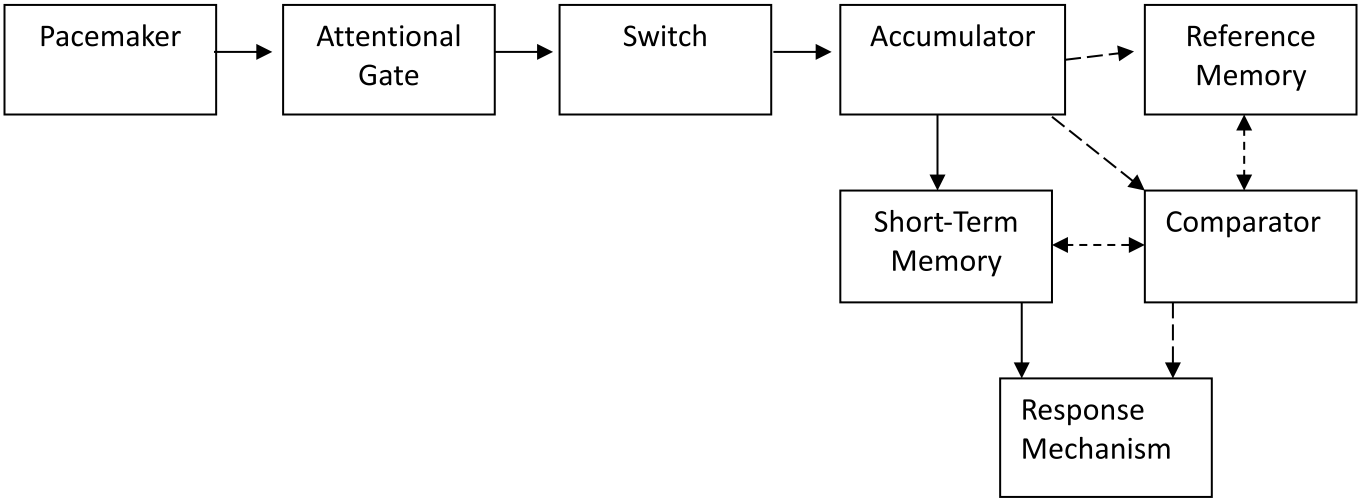

Figure 1.2. The attentional gate model (AGM)

Variable rate pacemakers. The second group of models with a dedicated event generator is comprised of models that call for a pacemaker with a variable rate, influenced by physiological arousal or by stimuli characteristics. Like the previous group of models, models in this group call for a pacemaker which emits pulses. Unlike the fixed rate pacemaker models, the rate of pulse generation in the following models is not a constant.

The majority of variable rate pacemaker models describe a pacemaker with a rate that varies with physiological arousal. As arousal increases, because of thoughts, moods, environmental events, or pharmacological influences, the firing rate of the pacemaker increases (Allman, Teki, Griffiths, & Meck, 2014). This results in subjective time flowing more quickly (i.e., an interval appears subjectively shorter when highly aroused than when moderately aroused).

Hoagland (1933) demonstrated that when participants were asked to tap at a rate of one tap per second, tapping rates increased as body temperature increased. This, along with other contemporary data showing it is possible to change participants' estimates with relatively simple manipulations, led Hoagland to propose a pacemaker with a rate that varied based on physiological arousal and the general state of the nervous system, which he called a master chemical clock (1933). Hoagland's model contains no explicitly described transducer; pulses are summed in some way, but it is not clear how this is done.

The second model in this group is scalar timing theory (Gibbon, 1991; Gibbon, Church, & Meck, 1984). Originally developed to account for animals' ability to complete time-based tasks, scalar timing theory has been extended to human timing (Grondin, 2001). Scalar timing theory, similar to Hoagland's master chemical clock, describes a pacemaker that emits pulses at a rate determined by physiological arousal (Gibbon et al., 1984).

Treisman's internal clock model (1963) has its basis in Hoagland's master chemical clock (1933), but, unlike Hoagland's model, makes no attempt to describe the underlying neural processes. According to Treisman (1963), the event generator is a pacemaker which emits pulses. When arousal is high, pulses are emitted at a faster rate than when arousal is low. This variation in pacemaker rate is produced by the specific arousal mechanism. Hoagland's master chemical clock model and scalar timing theory also posit a variation in pulse rate due to physiological arousal, but do not include a mechanism responsible for that variation.

Divenyi and Danner (1977) proposed another modified version of Creelman's original model. Unlike other models with a variable rate pacemaker, Divenyi and Danner's model asserts that stimuli characteristics, rather than physiological arousal, influence the rate of the pacemaker. Creelman's original model, described above, postulated a fixed rate pacemaker, with an invariable rate of pulse generation. Divenyi and Danner (1977) conducted a series of experiments where participants were required to judge which of two intervals was longer, and the auditory cues for each interval were systematically varied in pitch and intensity. When the cues were significantly different from one another in pitch or intensity, participants were less able to discriminate between the intervals. In order to account for their data, Divenyi and Danner proposed that the rate of pulse generation is influenced by stimulus characteristics, but do not specify how this occurs.

Oscillator models. Oscillator models, like fixed and variable rate pacemaker models, consist of a dedicated event generator which emits a signal to be counted and transformed by a transducer into a representation of time. The primary difference between oscillator and pacemaker models is the source of the timing signal. In pacemaker models, a pacemaker emits regular, discrete pulses, with a rate that is either constant or that varies based on some factor (arousal, stimulus duration). Oscillator models call for a series of mechanisms that produce cycles with regular periods. By observing the phase, or position, of the oscillator, the participant can calculate the duration that has passed. Similar to a metronome, an oscillator moves or cycles at a fixed rate. A metronome ticks back and forth at one beat per pass. A full cycle of the metronome arm (moving from left to right and back again) represents two full beats. When the participant looks at the metronome, a simple observation tells them where they are in the current beat. If the arm is at either the left or right edge, the current time is on the beat; if the arm is between the two sides, the current time is on the half- or quarter-beat.

Treisman's oscillator model. Treisman observed that it is possible for people to carry out two or more timing activities concurrently, a possibility which is difficult to account for with pacemaker models (Treisman, Faulkner, Naish, & Brogan, 1990). While it is possible that multiple pacemakers exist and are specialized for different intervals, Treisman concluded that the idea of multiple pacemakers distributed throughout the system was too complicated to be practical. According to Treisman, if multiple pacemakers all emit pulses at the same regular interval, it is difficult to time intervals outside of a small range; intervals shorter than the rate of the pacemaker cannot be accounted for, but a pacemaker with a quick pulse rate generates a massive number of pulses in a long interval. At the same time, if multiple pacemakers make use of different rates, it is difficult to coordinate behavior throughout the system and to make use of reference memory (Treisman et al., 1990).

In response to these two conflicting requirements (a stable rate and multiple rates), Treisman modified his original internal clock model to include an oscillator in place of the pacemaker (Treisman et al., 1990). The oscillator runs in cycles with a fixed, regular period Treisman called the characteristic frequency. The oscillator emits pulses as it crosses points in its cycle, much like the ticking of a metronome. In order to account for timing of a variety of intervals, Treisman described a calibration unit, adjacent to the oscillator, which takes information about the interval to be timed and the physiological arousal of the participant and calculates a modified pulse rate. The calibration unit receives pulses from the oscillator and transforms them into a function of the original rate (e.g., by multiplying by a factor of 2, or dividing by a factor of 3), depending on the requirements of the timing task and the level of arousal. When arousal is high, the calibration unit increases the speed of the pulses, and when arousal is low, the calibration unit decreases the speed of the pulses. This design allows the system to coordinate behavior – if one limb has to move twice as fast as another limb, the calibration unit simply has to modify the characteristic frequency for the faster moving limb (Treisman et al., 1990). Treisman proposed that the oscillator is composed of a network of neurons exerting excitatory and inhibitory effects on one another.

Treisman attempted to determine the characteristic frequency of the oscillator by presenting participants with series of clicks or visual flickers at differing rates, and noting which click or visual flicker rates interfered with timing (Treisman & Brogan, 1992; Treisman et al., 1990; Treisman, 1993). In both modalities, patterns of interference suggested an oscillator with a characteristic frequency of 12-13 Hz.

Connectionist model. In response to criticism that the traditional scalar timing model could not account for animals' ability to time multiple intervals concurrently, Church and Broadbent (1990) developed a connectionist model (see Figure 1.3). In contrast to scalar timing and its single pacemaker, the connectionist model calls for a series of oscillators with different periods, operating in parallel. Information about the current phase of each oscillator is passed into the transducer. Unlike Treisman's oscillator model, the oscillators in the connectionist model do not emit pulses. Instead, the phase of each oscillator is noted by the status indicators (see Figure 1.3) and used to calculate the current time. To understand this more clearly, imagine a wall with three clocks, each of which has a single hand. The clock on the left has only an hour hand, the middle clock has only a minute hand, and the clock on the right has only a second hand. By observing the current position of the hands on all 3 clocks, one can easily calculate the current time.

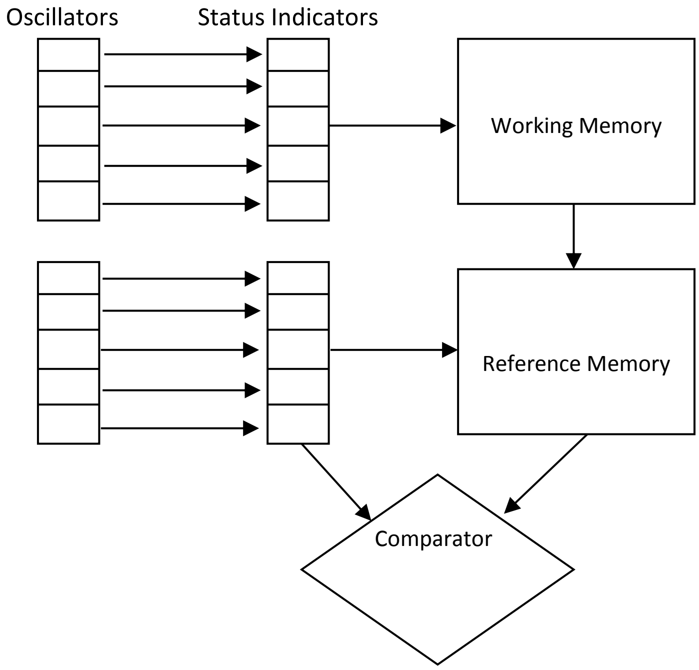

Figure 1.3. Connectionist model. Adapted from Church and Broadbent (1990).

Types of Transducers

Once a timing signal (i.e., pulses or oscillator position) has been generated by the event generator, the signal must be summed or transformed in some way to yield a judgment about the amount of time that has passed. Transducers, the mechanisms proposed for achieving this transformation, vary in complexity, from very simple summation and comparison units to complex mechanisms that make use of reference memory and include control elements which permit them to accommodate influences due to attention.

Simple summation. The simplest type of transducer adds up the number of pulses in an interval and compares that sum to the sum of pulses from a comparison interval (Allan et al., 1971; Allan & Kristofferson, 1974; Creelman, 1962; Divenyi & Danner, 1977; Killeen & Weiss, 1987; Kinchla, 1972). For purposes of illustration, imagine a very simple balance. If two weights are placed on either end of a simple balance, one can tell which one is heavier, but not what either one weighs. In the same way, these simple summation transducers allow for comparison of two intervals (e.g., which of these tones was longer?) but not explicit judgments about their duration (e.g., how many seconds was the tone?).

The models with this type of transducer (Allan et al., 1971; Creelman, 1962; Divenyi & Danner, 1977; Killeen & Weiss, 1987; Kinchla, 1972) were specifically designed to account for the ability to discriminate between the durations of very short intervals presented in quick succession. The mathematical underpinnings of these models vary slightly, but the structure of the transducer is identical in each model. The lack of any sort of long-term memory store with labels for different pulse counts results in models that can only account for discrimination between two intervals, not production, reproduction, or estimation of intervals. Transducers of this type are only capable of making discrimination judgments – judging whether one interval is longer or shorter than a recently presented interval.

The simplicity of these transducers is severely limiting; models with this type of transducer cannot account for explicit judgments. They can potentially account for prospective timing, but only in discrimination tasks. Models with simple summation transducers, then, do not provide a satisfactory explanation for everyday, intermediate duration tasks.

Transducers with memory components. Models in this group call for transducers with two memory components – some type of working memory or short term storage capacity, and some form of reference memory. The short term storage or working memory component allows participants to hold on to the counted pulses or position of the oscillators for the duration of the task. Reference memory involves long term storage of the number of pulses or position of oscillators that corresponds to particular durations. This allows participants to make explicit judgments (e.g., how many seconds have passed?) and productions (e.g., press the key when 30 seconds have passed). The sum of pulses or the position of the oscillators held in short term storage or working memory is compared to stored values in reference memory that represent known durations, and an explicit judgment can be made.

Treisman's internal clock and oscillator models. The transducer in Treisman's internal clock model (1963, see Figure 1.4) calls for a counter, which counts the pulses emitted from the pacemaker. The later oscillator model (Treisman et al., 1990) calls for the same transducer; modifications to the model were limited to the event generator stage. According to the internal clock model, each pulse serves to increment the counter. This count of pulses is then stored by the memory store, allowing the participant to compare multiple intervals or transform the pulse count into an explicit judgment. The verbal selective mechanism contains verbal labels for different pulse counts (e.g., 30 seconds) and allows us to make explicit judgments regarding the amount of time that has passed (Treisman, 1963).

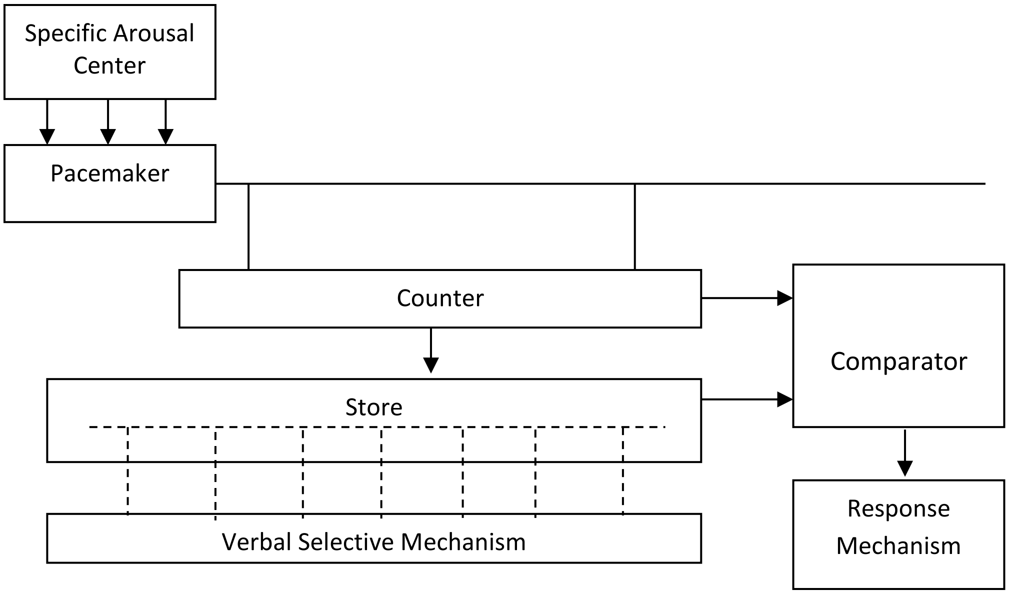

Figure 1.4. *Treisman's internal clock. Adapted from Treisman (1963).*

When an explicit judgment is needed (e.g., "the stimulus was presented for 2 minutes"), the comparator compares the pulse count stored in the memory store to the pulse counts stored in the verbal selective mechanism in order to make a decision about how much time has passed. When a relative judgment is required, the comparator compares the count in the counter to the store in order to determine which represents a longer interval. Once the comparator has compared the contents of the counter and the store, or the store and the verbal selective mechanism, the response mechanism selects and makes its response.

Scalar timing theory. Scalar timing theory (Gibbon et al., 1984, see Figure 1.5) describes a transducer with working memory and reference memory components. When the participant receives a signal to begin timing, the switch closes, and pulses begin to pass into the accumulator. The count in the accumulator is then passed into working memory. When an explicit judgment is necessary, the comparator compares the count stored in working memory with the label for each count in reference memory in order to make a decision. When a relative judgment is necessary, counts from two intervals can be held in working memory and compared. The comparator is responsible both for comparing the counts from two intervals or from an interval and reference memory and for making a decision or response.

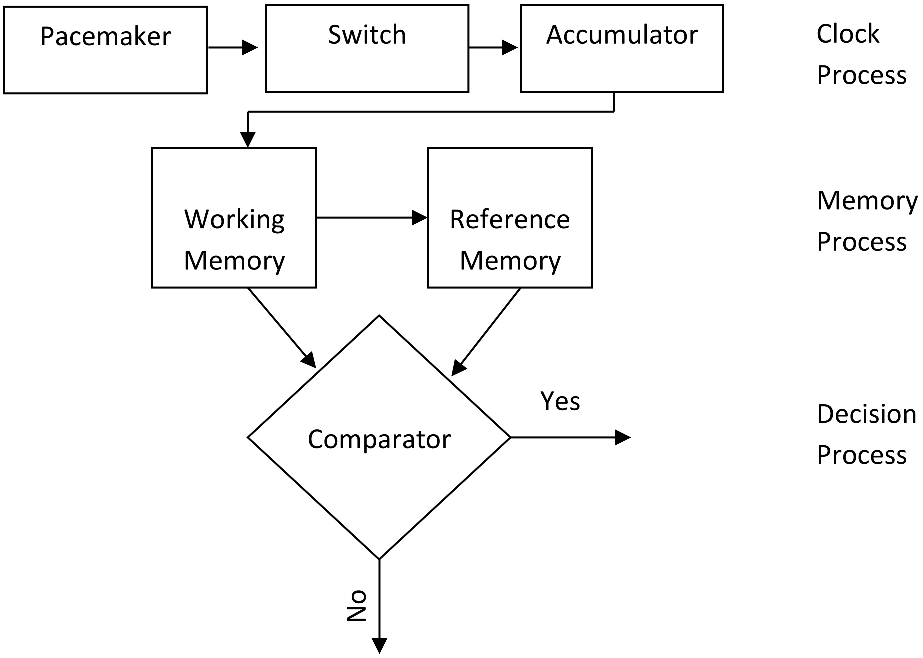

Figure 1.5. Scalar timing theory. Adapted from Gibbon, Church, and Meck (1984).

Connectionist model. The connectionist model (Church & Broadbent, 1990), unlike the models previously described, does not call for an event generator that results in discrete pulses. Instead, the position of multiple oscillators is tracked in the status indicators, and this information serves as the input for the transducer. Information about the current position of each oscillator is transmitted to the status indicators, and then stored in working memory. This information is then compared with reference memory by the comparator in order to determine how much time has passed.

Unlike the reference memory in other models in this group, which is composed of a series of values and their corresponding durations, the connectionist model describes a reference memory composed of a matrix of information about possible combinations of oscillator phases (i.e., if oscillator 1 is in position X and oscillator 2 is in position Y, Z time has passed). This matrix is assumed to simplify the processes of comparing the current position of the oscillators to reference memory, and greatly reduce the amount of information about past trials that must be stored (Church & Broadbent, 1990). Church and Broadbent argued that the expansion to multiple oscillators, and the corresponding reduction in the amount of information required in reference memory, could account for timing in the range of milliseconds to hours and beyond.

Treisman's internal clock and oscillator models (Treisman, 1963; Treisman et al., 1990), scalar timing theory (Gibbon et al., 1984), and the connectionist model (Church & Broadbent, 1990) all describe transducers with clear roles for a reference memory component. This reference memory component allows participants to access a verbal label for the product of the event generator, which allows participants to make explicit judgments regarding time. Participants can make estimates of a presented interval by referring to these labels (e.g., "I think 2 minutes have passed") and can produce intervals from verbal instructions (e.g., "press the spacebar when 4 minutes have passed").

The main shortfall of models with this type of transducer is the lack of a role for attention and strategy. The event generators and transducers in these models do not include components that would result in shorter or less accurate judgments when participants are engaged in a secondary, attention-demanding task. All of the transducers described are assumed to function in the same way whether or not the participant is attending to time.

A role for attention. The attentional gate model (AGM) includes an attentional gate to account for the effects of attention and strategy (Zakay & Block, 1996, 1998). As previously described, the AGM calls for a fixed rate pacemaker. Similar to scalar timing (Gibbon et al., 1984), the AGM calls for a switch located between the pacemaker and the accumulator. When the participant receives a signal to indicate that timing is relevant, the switch closes, much like the switch on an electrical circuit, allowing pulses to transfer into the accumulator.

The key element of the AGM that differentiates it from other internal clock models is the attentional gate, which is controlled by the amount of attention available for processing time information (Zakay & Block, 1996; 1998). When more attention is available for and allocated to temporal processing, the gate is open wider, allowing more pulses to pass through. When less attention is available for time, because it is allocated to processing nontemporal information, the gate is narrowed or closed, and fewer pulses pass through. When a secondary, nontemporal task requires more attentional resources, less attention is available for processing time information. Importantly, the allocation of attentional resources to temporal and nontemporal information can also be controlled voluntarily; when a participant dedicates more attention to timing, the gate is open wider, and pulses accumulate more quickly. A participant engaged in an attention-demanding nontemporal task (i.e., solving complex math problems) may allocate less attention to temporal information in order to succeed on the nontemporal task, resulting in a narrower gate aperture. Alternatively, if the participant feels the temporal task is more important, he may allocate more attention to the temporal task, resulting in a wider gate aperture.

The AGM does not explicitly describe what type of attention is involved in opening and closing the attentional gate, but it seems clear that attention is being conceptualized as a divisible but limited resource, consistent with a capacity model of attention (Kahneman, 1973). Many tests of the AGM involve a load manipulation, where participants are assigned to either an easy or difficult ongoing task. According to capacity models of attention, participants in the easy condition would have more attentional resources available to direct towards the timing task than would participants in the difficult condition. At the same time, the AGM also suggests that attention is conceptualized in line with spotlight models of visual attention (LaBerge, 1983). Attentional allocation manipulations, where participants complete the same ongoing and timing tasks but are instructed that one of the two tasks is more important, have been used to test the AGM.

Whether resource models or spotlight models best capture the meaning of “attention” in the AGM, it seems clear that the attention to timing that is controlling the attentional gate is top-down, endogenous attention. When participants are instructed to focus more on a timing task than an ongoing task, productions are shorter, indicating that the attentional gate is sensitive to goal-directed attentional focus. In addition, the AGM appears to fit data from prospective interval timing, where participants are aware that they will be asked to produce or estimate time, but not retrospective interval timing, where naïve participants are asked after the fact to estimate time. In the latter case, top-down attention to timing is not engaged, as participants are not aware that they will be asked about time, and the attentional gate aperture is not changed.

The accumulator counts the pulses that pass into it, and stores this pulse count in working memory. In order to make a decision about whether a given interval has elapsed (i.e., a production task) or to estimate how long an interval was (i.e., an estimation task), the comparator compares the accumulated pulse count stored in short term memory with a reference memory bank that contains pulse counts for different known intervals.

The role of reference memory and the comparator varies with the type of timing task being performed (Zakay & Block, 1998). Zakay and Block do not explicitly state how the reference memory and comparator work in different types of tasks, but extending the theory to these tasks is relatively simple. In a prospective time estimation task, participants are instructed that they will be exposed to a stimulus and then asked to estimate how long the stimulus lasted. In prospective time estimation tasks, the comparator compares the accumulated pulse count stored in short term memory with the reference memory values for a number of intervals, and decides which interval the accumulated pulse count most closely matches. If the accumulated pulse count is, for example, 1709, and the reference memory values state that 1600 pulses are equivalent to 2 minutes and 1800 pulses are equivalent to 2.25 minutes, the participant will make an estimate somewhere between 2 and 2.25 minutes. In a prospective time production task, participants are instructed to make a response when they believe a given interval has passed. In time production tasks, the comparator makes a series of comparisons between the accumulated pulse count and the reference memory for the target interval, and prompts a response when the accumulated pulse count is close to or equal to the reference memory pulse count. If, for example, the participant is to produce an interval of 2 minutes, the comparator makes a series of comparisons between the accumulated pulse count and the reference memory for the pulse count of 2 minutes (e.g., 1600 pulses), and when the accumulated pulse count nears 1600, prompts a response. In a prospective time reproduction task, participants are instructed that they will be required to reproduce the duration of an event. After being exposed to the event, participants are asked to make a response when they believe the same amount of time has elapsed. In time reproduction tasks, the comparator makes a series of comparisons between the reference memory for the experienced duration of the stimulus and the accumulated pulse count from the reproduction, prompting a response when the pulse counts match.

The AGM predicts that when more attention is available for timing, experienced durations should be shorter (i.e., subjective time is passing more slowly than objective time) than when attention is directed away from timing (Zakay & Block, 1998). When less attention is allocated to processing time information, the aperture of the attentional gate is narrowed, and pulses accumulate more slowly. In reproduction and production tasks, this results in longer response times when attention is directed away from timing, as it takes longer for the accumulated pulse count to match the reference memory for the target interval. In estimation tasks this results in shorter estimates when attention is directed away from timing, as the accumulated pulse count is lower and corresponds to the reference memory for a shorter interval. These predictions are consistent with a large body of experimental work on human timing that focuses on manipulating the attentional resources available for or directed towards timing (cf. Brown, 1985; Brown, 1997; Brown, Collier, & Night, 2013; Hemmes et al., 2004; Hicks et al, 1976; Kladopoulos et al., 2004; Macar et al., 1994; McClain, 1983; Zakay & Tsal, 1989).

The AGM has many strengths: it accounts for the role of attention and strategy in human timing, it can explain prospective timing, and it addresses how explicit judgments are made. The AGM is well supported by the literature on human timing.

Conclusions

The goal of this chapter was to identify a model or models best suited for explaining conscious, deliberate, and strategic timing in the range of minutes. In order to do so, a model must include a role for attention and strategy, be capable of accounting for prospective timing, and must include a mechanism for making explicit judgments. A summary of the models' ability to fit these three criteria is presented in Table 1.1.

Model type	Model	Attention and strategy	Prospective timing	Explicit judgments
By-product of Cognitive Processing	Discrete Moment (Stroud, 1967)	N	Y	N
	Travelling Moment (Allport, 1968)	N	Y	N
	Emotional Moment (Craig, 2009)	N	Y	N
Memory Models	Memory Storage (Ornstein, 1969)	N	N	N
	Contextual Change (Block, 1978; Block & Reed, 1978; Poynter, 1983)	N	N	N
Dedicated Event Generators	Master Chemical Clock (Hoagland, 1933)	N	Y	N
	Creelman and modifications (Allan et al., 1971; Creelman, 1962; Divenyi & Danner, 1977; Killeen & Weiss, 1987; Kinchla, 1972)	N	Y	N
	Treisman's Internal Clock (1963)	N	Y	Y
	Scalar Timing (Gibbon, Church, & Meck, 1984)	N	Y	Y
	Attentional Gate Model (Zakay & Block, 1996, 1998)	Y	Y	Y
Oscillator Models	Treisman's Oscillator Model (1990)	N	Y	Y
	Connectionist Model (Church & Broadbent, 1990)	N	Y	Y

Table 1.1. Ability of described models to meet criteria.

Models of timing typically consist of an event generator, which creates some sort of signal, and a transducer, which translates this signal into some sort of information about time. I have organized models first by event generator type. The first broad class of event generators, those which are a by-product of other processes, includes models that make use of long-term memory (Block, 1978; Block & Reed, 1978; Ornstein, 1969; Poynter, 1983) and models that make use of perceptual or other cognitive processes (Allport, 1968; Craig, 2009; Stroud, 1967). Models in this class do not describe transducers at all, and do not include a role for attention and strategy or the ability to make explicit judgments. Additionally, the memory storage and contextual change models appear to account for retrospective timing but fail to account for results in prospective timing tasks (Block, 1992; Hicks et al., 1976; Grondin, 2001).

The second broad class of event generators consists of dedicated event generators, both pacemakers and oscillators, all of which are special purpose mechanisms dedicated solely to timing. Models in this class, with the exception of Hoagland's master chemical clock (1933), have clearly described transducers. Several of the models have very simple transducers which are incapable of accounting for the role of attention or explicit timing (Allan et al., 1971; Creelman, 1962; Divenyi & Danner, 1977; Killeen & Weiss, 1987; Kinchla, 1972). Transducers with working memory and reference memory components can account for explicit judgments, but they lack a role for attention and strategy (Church & Broadbent, 1990; Gibbon et al., 1984; Treisman, 1963; Treisman et al., 1990). Only the AGM can account for both explicit judgments and the role of attention and strategy (Zakay & Block, 1996, 1998).

The AGM is the only model that includes a clear role for attention and strategy. Although it fits the criteria outlined above, the AGM was not developed with intermediate, everyday intervals in mind and to my knowledge has not been directly tested on intervals from 2-6 minutes. It seems implausible that participants would be able to sustain attention on timing for such long intervals. However, it may be possible to modify some of the components of the AGM in order to account for timing in this range.

The memory storage models, such as those developed by Ornstein (1969), and Block (1978; Block & Reed, 1978) also warrant further investigation. Poynter (1983) found support for memory storage models in retrospective timing in the range of intervals investigated in my dissertation. Block and Reed (1978) found support for closely related contextual change models in prospective intervals in this range. It is possible that in longer intervals, where attention cannot be sustained on timing for the duration of the task, modification of memory storage models may be able to account for the pattern of findings. Most research on memory storage models involves very short intervals, with support found for such models in retrospective, but not prospective, timing (Block, 1992; Hicks et al., 1976; Grondin, 2001). One possibility is that in short interval timing, the short term storage component of the AGM is capable of storing the pulse count gathered by the accumulator, but that in longer timing tasks that fall beyond the limits of short term memory, a reconstructive process similar to that used in retrospective tasks is used in order to estimate the pulse count.

We know very little about timing of intervals that are relevant for everyday tasks. Common sayings, like “time flies when we’re having fun” and “a watched pot never boils” seem to imply that our perception of time is influenced by concurrent activities. The AGM and memory models seem most consistent with intuition about the timing of everyday intervals. It seems likely that our time estimates would depend on whether or not we pay attention to the unfolding of time, with the duration of estimates varying based on whether the ongoing task requires more or less attention (Block, 1992). It also seems likely that our time estimates would depend on what or how much we remember about a period of time, consistent with memory models (Block & Reed, 1978; Poynter, 1983). When a period of time contains a large amount of information or a large number of contextual changes, our subjective experience of that interval seems longer.

In the following chapters, I will discuss a series of experiments that I conducted in order to determine whether the AGM and memory storage models can account for the duration of estimates produced in prospective timing of everyday intervals, whether they need to be modified to fit intervals in this range, and if modifications are necessary, which components require modification. The first three experiments focused on investigating the pattern of estimates produced in prospective timing of everyday intervals. In Experiment 1, I attempted to establish the underlying pattern of estimates in this range. My previous research (Fergusson, 2010) has shown that participants consistently underestimate intervals in this range, but the relationship between target interval duration and estimate duration has not yet been established. In order to determine which model(s) might best fit estimation of intervals in the range of everyday tasks, it is necessary to know whether the relationship between target interval duration and estimate duration, as different models make different predictions about the relationship. Participants in Experiment 1 produced intervals of 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 minutes in order to determine what pattern best fit the estimates produced. In Experiment 2, I introduced a framing manipulation in order to investigate the role of memory chunking in timing everyday durations and thereby investigate memory storage models in this range. If chunking influences estimation in this range, memory storage models may be the best fit. Participants were instructed to focus on the present or the future while producing intervals of 2, 4, and 6 minutes. Experiment 3 combined the methods of Experiments 1 and 2 in order to determine whether manipulating task framing would influence the underlying function or simply the magnitude of the underestimates. In Experiment 4, an attentional manipulation and accuracy feedback were used in order to evaluate the fit of the AGM to intervals in this range and the role of reference memory. If performance on the timing task is influenced by the difficulty of the ongoing task, as seen in many studies of shorter intervals, then the AGM may be a good fit for timing in this range. If feedback improves the accuracy of estimates in this range, it may be the case that we lack experience or have inaccurate stored pulse counts that lead to our timing errors.

.

Chapter 2 - Establishing the Pattern of Estimates for Everyday Intervals

Little is known about our ability to estimate the duration of everyday intervals. In order to investigate which timing models best fit the pattern of estimates in this range, it is first necessary to address questions about the relationship between target interval duration and duration of estimates. The Attentional Gate Model (AGM, Zakay & Block, 1996, 1998) and memory storage models (Block, 1978; Block & Reed, 1978; Ornstein, 1969) provide possible accounts for timing in this range, as discussed in the previous chapter, but without data on the relationship between target interval duration and estimate duration it is difficult to evaluate their explanatory power. Experiment 1 was designed to address two main objectives: the magnitude of estimate duration in everyday intervals and the relationship between target interval duration and estimate duration (i.e., is the estimate error constant across interval durations; is it linearly related to interval duration).

To my knowledge, only three previous studies, conducted by Block (1992), Block and Reed (1978), and Poynter (1983), addressed intervals in this range, and their results provide no insight into the nature of the relationship between estimates and interval durations. These studies investigated only a small number of intervals: Block (1992) investigated only intervals of 165 seconds, Poynter (1983) only intervals of 170 seconds, and Block and Reed (1978) only intervals of 60, 120, and 300 seconds. This small number of intervals makes it impossible to draw any informed conclusions about the relationship between target interval duration and estimate duration. More importantly, neither Block and Reed (1978) nor Poynter (1983) provided any data about the size of the temporal underestimates; they simply reported that participants underestimated and which condition resulted in larger underestimates. Only in the case of Block (1992) is data about the size of underestimates provided, but because that study included only one interval, it is not possible to draw any conclusion about the relationship between target interval duration and estimate duration.

What might be the relationship between interval duration and participants' estimates? As reported in the preceding paragraph, Block and Reed (1978), Poynter (1983), and Block (1992) concluded that the duration of everyday intervals is consistently underestimated. By contrast, research on short duration intervals has consistently shown that such intervals are overestimated (Grondin, 2001). In light of this combination of findings, two possibilities are conceivable. First, it might be that participants overestimate short intervals, while under estimating long intervals, resulting in a linear relationship between the estimated and actual duration of intervals. By this possibility, we can expect that somewhere between very short intervals and long, everyday intervals, participants provide estimates that are accurate (i.e., neither over nor under estimates of their actual duration). A second and completely different possibility, based only on the limited research on everyday task durations, is that participants' always underestimate such intervals. It is possible that participants' time estimates are a constant proportion of such interval durations, that the relation between them is a Weber fraction (Shouval, Shuler, Agarwal, & Gavornik, 2014).

The possibility that longer duration intervals are underestimated is also consistent with the research by Lewis and Miall (2009) who focused on the *reproduction* of durations. As explained in Chapter 1, reproduction tasks differ from production tasks primarily in that verbal labels are not used. In a reproduction task, participants are exposed to an interval and asked to reproduce it. In a production task, participants are given instructions to respond after a given amount of time, requiring them to map verbal labels onto an interval. It is not known to what extent this verbal mapping influences estimates, and it is possible that the mechanism involved in reproducing an interval differs from that involved in producing an interval. Nevertheless, there are two studies by Lewis and Miall (2009) where reproduction tasks were used to estimate longer intervals (3.59 minutes and 16.7 minutes). The results showed clear evidence of underestimates for longer intervals, though the report by Lewis and Miall does not provide numeric data on the size of the under estimate and does not indicate whether the underestimate was different for different intervals.

Research on time estimation is not concerned with participants' ability to estimate time when they make use of a deliberate strategy (i.e. counting one Mississippi, two Mississippi). Instead, the focus is and has been to examine time estimation under conditions where participants are engaged in an ongoing activity that is attention demanding and prevents strategic time monitoring. Despite this declared focus, many previous experiments employed concurrent tasks which did not permit performance monitoring, or the published reports did not include any data on the ongoing task (Brown, 2006). In the absence of such data, it is not possible to make firm conclusions about the attentional demands of the ongoing task.

To address this method issue, Experiment 1 was also designed to pilot a novel ongoing task which would allow for monitoring of accuracy and response time data, as well as allow for manipulation of difficulty. Specifically, I used a dot counting task, in which a series of 6x6 grids, each filled with between 8-15 blue dots, was displayed in the center of a computer screen. For each grid, participants were required to respond via the computer keyboard with the number of dots they counted. Participants' performance on this task could be examined for accuracy and response speed, as well as for the relationship between responding on this task and the duration of their time estimates. The dot counting task was designed to prevent participants from strategic time monitoring (e.g., counting seconds), by keeping them occupied with counting dots. In my previous work, many participants reported that they tried to count seconds in order to make their estimates more accurate, but were largely unable to do so because of the attentional demands of the ongoing task.

Method

Participants and design. Participants were 66 undergraduate student volunteers from the University of British Columbia who participated in this experiment in return for course credit. The experiment had one within-subjects factor (target interval duration) with nine levels: 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 minutes.

Procedure. All participants were tested individually using e-Prime on a desktop computer. At the beginning of each session, participants were asked to remove watches and to place any time-tracking devices out of sight. The experiment took approximately two hours to complete. The experiment was conducted with the approval of the University of British Columbia Behavioural Research Ethics Board.

Each participant was required to carry out two concurrent tasks: an ongoing task which involved dot counting and a prospective timing task. The ongoing task consisted of a large number of trials. On each trial, participants were presented with a 6x6 grid (see Figure 2.1 for a representative example), which occupied a 6-nch square in the center of the screen. The grid was filled with between 8-15 blue dots; the dots were placed at random in the boxes of the grid. Participants were instructed to count the number of dots, enter it (i.e. the count) on the number pad on a computer keyboard, and pressing the Enter key to submit each response. Immediately after pressing the Enter key, the next grid appeared.

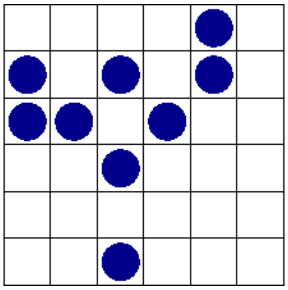

Figure 2.1. Representative example of a grid used in the dot counting task in Experiment 1.

At the beginning of the experiment, participants were told that they would be producing different intervals of time, and that they should estimate each interval as accurately as possible. Participants were required to make 3 estimates for each of 9 different intervals: 1, 1.5, 2, 2.5 etc., for a total of 27 estimates. The ordering of these 27 interval-trials was completely random for each participant. For each new interval, instructions appeared in the middle of the screen indicating what the target interval was and directing participants to press the spacebar when they believed the target interval had elapsed (i.e., “Press the Spacebar when 2 minutes have passed”). While engaged in the interval production task, participants completed a series of trials of the self-paced dot counting task. When participants believed the target interval had elapsed, they pressed the spacebar and the dot counting task terminated. Participants then repeated this procedure for the next target interval. See Figure 2.2 for a visual depiction of the trial sequence. If participants failed to make a response to the target interval production task after 8 minutes, the current target interval production task terminated and the next began immediately.

After completing the interval production task, participants completed a questionnaire about their confidence in timing intervals lasting 2, 5, 10, 15, 20, 30, 45, and 60 minutes. Participants rated their confidence in their ability to accurately time each interval on a scale of 1 (not at all confident) to 8 (very confident). Experimental sessions lasted approximately 2 hours.

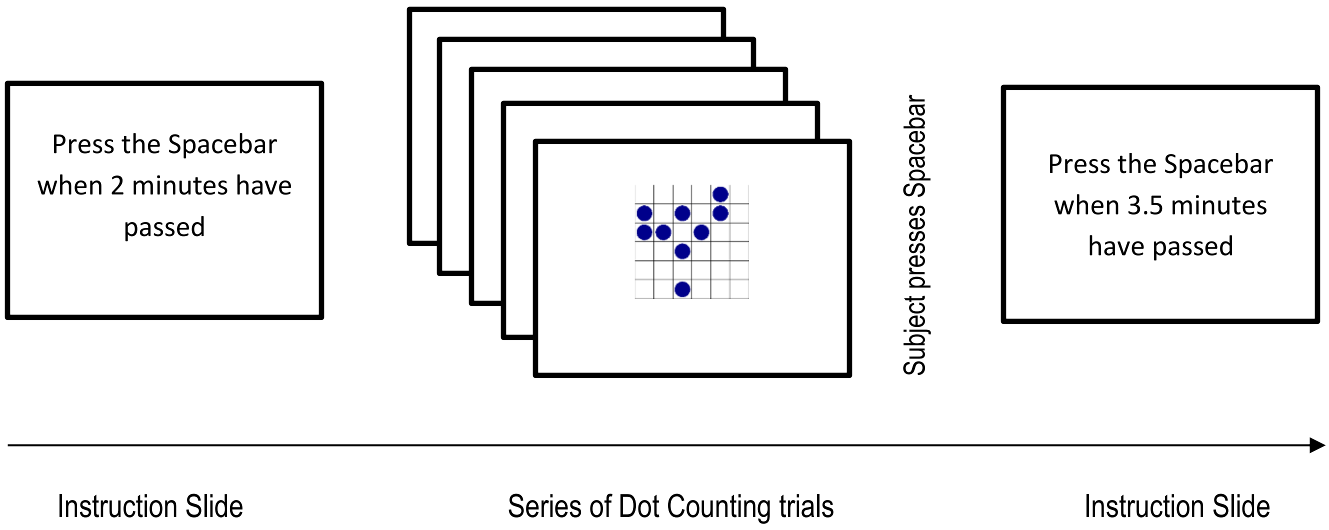

Figure 2.2. Visual depiction of the trial sequence in Experiment 2. Participants saw a slide with instructions to begin timing. Participants then completed a series of dot counting trials, which continued until the participant pressed the spacebar to indicate they believed the target interval had lapsed. The block terminated, and the instruction slide for the next target interval was presented.

Results

Dot counting task. For each participant, the mean accuracy and response time on the dot counting task was calculated. The mean accuracy across all participants was 64.51% (*SD* = 13.91). With 8 possible response options, random responding would result in an average accuracy of 12.5%. Participants' accuracy suggests that they were engaged with the task, not merely responding randomly. The mean response time on all trials was 1706.07 ms (*SD* = 678.16). The mean response time on accurate trials only was 1708.08 ms (*SD* = 790.108). Response time was not significantly different on accurate trials compared to all trials.

Confidence ratings. Participants rated their confidence in their ability to time intervals from 2 to 60 minutes on a scale of 1(not at all confident) to 8 (very confident). Participants indicated confidence above the middle of the scale for only the 3 shortest intervals (2, 5, and 10 minutes) (see Figure 2.3). These ratings, however, were not correlated with the length of estimates for 2 and 5 minute intervals, *r* (56) = -.058, *p* = .67 and *r* (56) = -.173, *p* = .20, respectively.

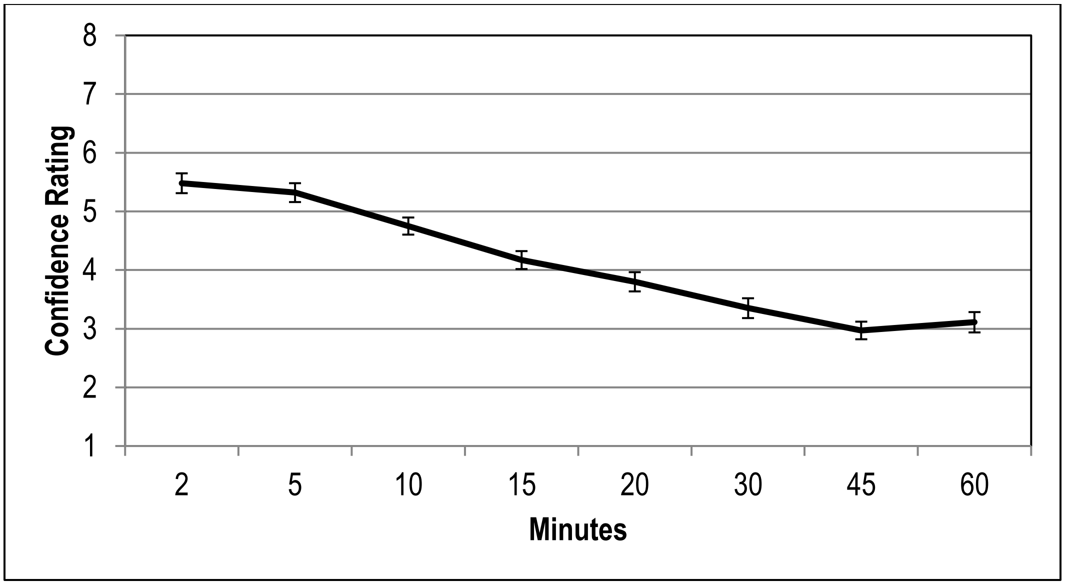

Figure 2.3. *Mean confidence ratings. Error bars represent standard error of the mean.*

Interval production performance. The key dependent variable of interest was the closeness of the time estimate produced by each participant for each target interval. The means of the median estimates produced by each participant for each target interval are shown in Figure 2.4.

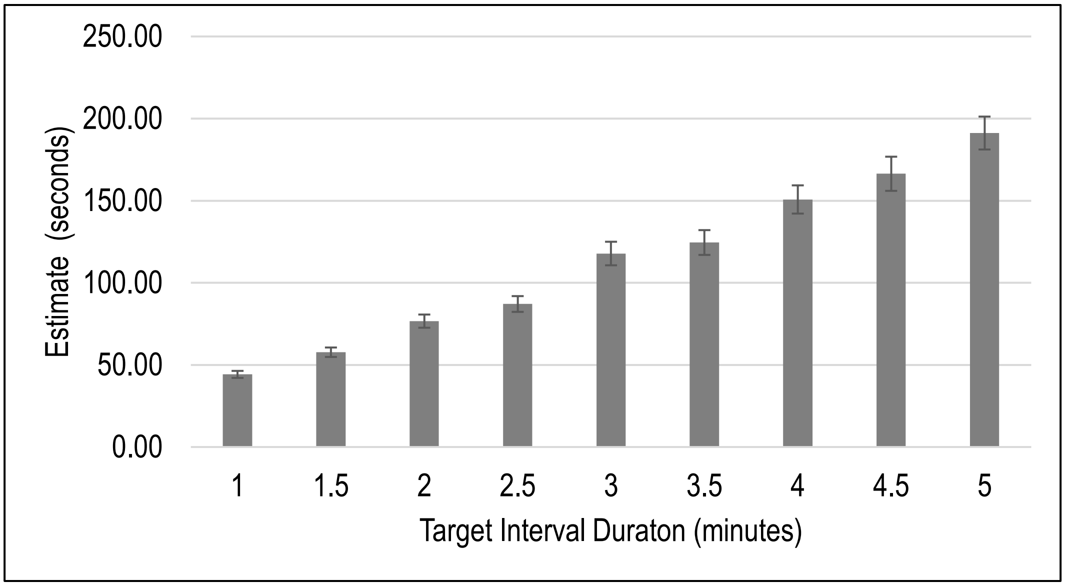

Figure 2.4. Means of median estimates for target intervals of 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 minutes in seconds. Error bars represent the standard error of the mean.

Estimates were converted to relative difference scores, expressed as a percentage of the target interval, using the following formula:

$$\left[\frac{Estimate - Target\ Interval}{Target\ Interval} \times 100\right]$$

Use of this formula yields a negative value for estimates that were shorter than the target interval (i.e., an underestimate), zero for estimates that were perfect, and a positive value for estimates that were longer than the target interval (i.e., an overestimate). Each participant's median estimate for each target interval (i.e., the median estimate over all of the one minute blocks, the median estimate over all of the one and a half minute blocks, etc.) was used in the following analyses. As described above, participants responded by key press to indicate when they thought the target interval had elapsed. In the event that participants did not respond and the program terminated the timing task, the program recorded a 0 for the participant's response time. These estimates were excluded from analysis. Terminations occurred for a total of 13 productions (0.72% of total) across all participants and intervals. No participants were eliminated as a result of having more than one program-terminated trial

Figure 2.5 shows the means of the median estimates for each interval. Participants underestimated all intervals, with underestimates growing between 1 and 2.5 minutes, shrinking for 3 minute intervals, and then leveling off between 3.5 and 5 minutes. The results of a trend analysis confirmed this summary of the findings by revealing a significant linear trend, $F(1, 65) = 4.30$, $p < .05$, $\eta^2 = .06$, with a slope of -.15 and intercept of -2.51, and significant quadratic trend, $F(1, 65) = 11.58$, $p < .01$, $\eta^2 = .15$. No other effects were significant.

One possible explanation for this pattern is simply that the 1 minute estimates were different from the estimates for all longer intervals, leading to the significant linear and quadratic trend. To examine this possibility, I re-ran the analyses without the data from the 1 minute condition. These follow—up analyses revealed no significant trends among the remaining conditions.

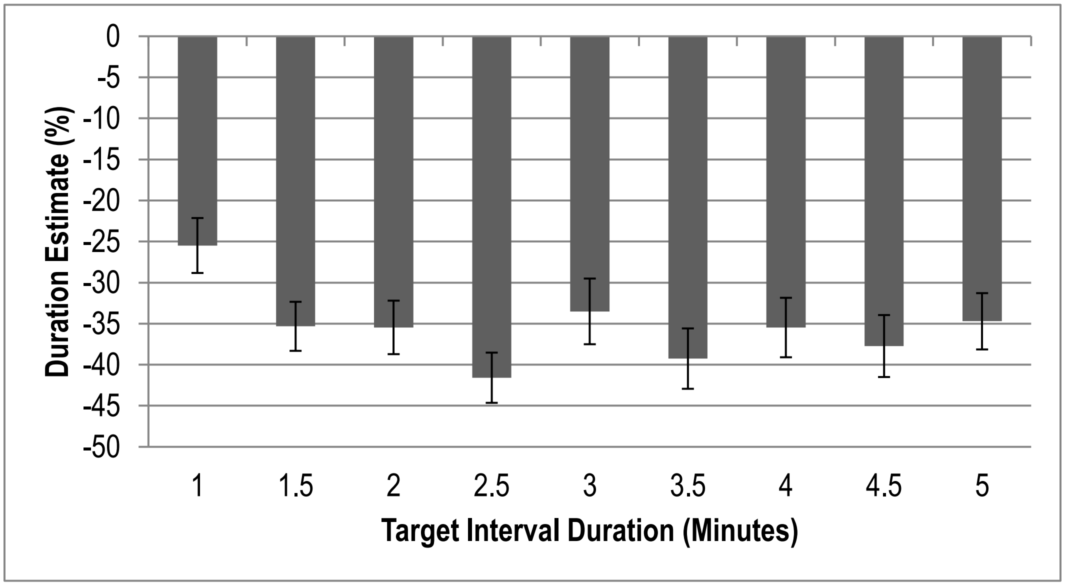

Figure 2.5. Duration estimates for target intervals of 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 minutes. All values are underestimates. Error bars represent the standard error of the mean.

Interval production performance was also analyzed as a function of dot counting task performance, in order to rule out the possibility that participants who found the dot counting task easy and those who found the dot counting task difficult produced significantly different intervals. Such a difference could indicate that the dot-counting task was not adequately attention-demanding to prevent strategic time-monitoring or counting. Participants were separated into high dot counting accuracy and low dot counting accuracy groups by a median split. Participants in the high dot counting accuracy group had a mean accuracy of 76.35% (*SD* = 5.63) and a mean response time of 1735.10 ms (*SD* = 654.73). Participants in the low dot counting accuracy group had a mean accuracy of 52.81% (*SD* = 9.65) and a mean response time of 1577.41 ms (*SD* = 501.73). Duration estimates produced by these two groups are presented in Figure 2.6. The results of a repeated measures ANOVA with interval as a within-subjects factor and dot counting performance (high vs. low) as a between-subjects factor revealed a significant effect of interval, $F(8, 240) = 3.77$, $p < .001$, $\eta^2 = .02$. No other effects were significant. Participants in the high dot counting accuracy group did not produce significantly different estimates of time than participants in the low dot counting accuracy group.

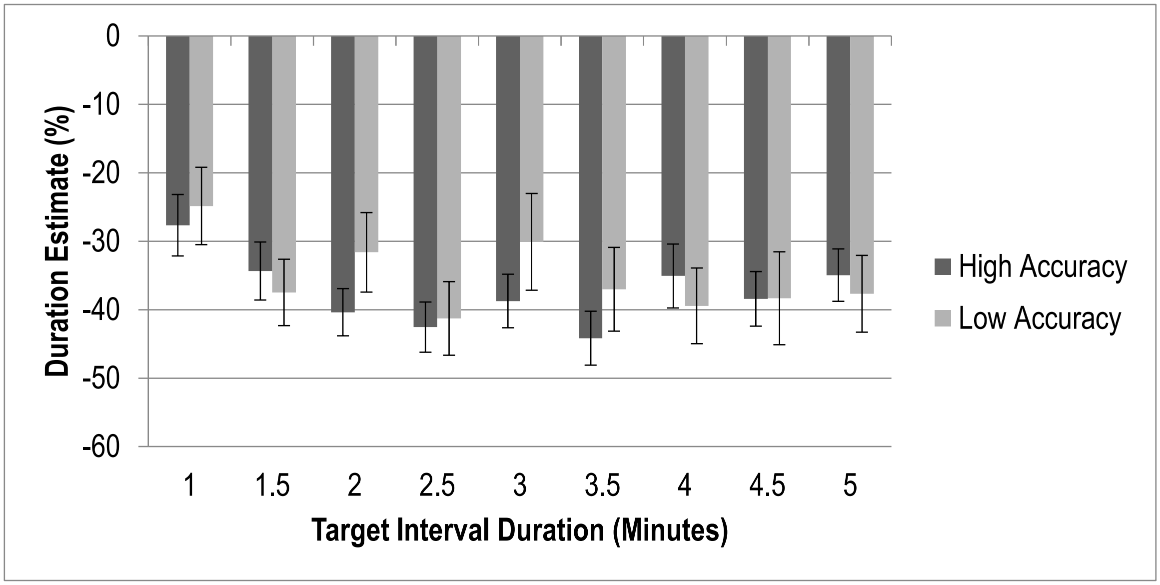

Figure 2.5. Duration estimates for 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, and 5 minute target as a function of performance on the dot counting task. Negative values are underestimates. Error bars represent the standard error of the mean.

Discussion

Experiment 1 was designed to investigate the magnitude of underestimates in the range of everyday intervals, to investigate the relationship between target interval duration and estimate duration, to test a new ongoing task, and to investigate participants' confidence in their ability to estimate time in this range. In Experiment 1, participants produced underestimates for all intervals, and these underestimates were large. The size of the underestimates did not differ between target interval durations, with the exception of the shortest 1 minute interval. Preliminary analyses appeared to indicate that the relationship is not strictly linear, but rather contains a linear and a quadratic component. Specifically, differences in the duration of estimates appeared between 1 and 1.5 minutes, and between 2.5 and 3 minutes. However, when the shortest interval was eliminated from analysis, the pattern disappeared completely. In addition, the size of the underestimate was not related to accuracy on the ongoing task, and was not correlated with participants' confidence in their ability to estimate these intervals.

Two findings require further discussion: The large underestimates produced for all intervals and the nature of the difference between the shortest 1 minute interval and the longer intervals. Previous research on timing in the range of everyday intervals has suggested that participants produce underestimates (Block & Reed, 1978; Block, 1992; Poynter, 1983). However, only Block (1992) provided data on the size of the underestimates. He reported that participants responded, on average, 23.2 seconds before the target of 165 seconds in the easy ongoing task condition and 34.8 seconds before the target of 165 seconds in the difficult ongoing task condition. When converted to percentages using the same formula I used in Experiment 1, this results in underestimates of 14.06% and 21.09%, respectively. The underestimates found in Experiment 1 were larger (*M* = 35.39%) than those reported by Block.

A number of factors might explain the difference between in the estimates from Block (1992) and the estimates produced in Experiment 1. One possible explanation lies in the nature of the ongoing task. In Experiment 1, participants were engaged in a dot counting task that was designed, in part, to disrupt their ability to count seconds. For the ongoing task in Block's experiment, participants were presented with lists of words. In the easy condition, they were instructed to passively view these words; in the difficult condition, they were instructed to name an action associated with each displayed word [e.g., if the word was beer, they could say "drink" or "pour" (Block, 1992)]. No data were reported regarding participants' performance on this ongoing task. At the conclusion of the experiment, Block gave a recognition test for the words from the ongoing task. Participants in the difficult condition performed better on the recognition test for the presented words (66% in the easy condition vs. 92% in the difficult condition). In view of the low level of recognition performance, participants in the easy condition may not have attended to the ongoing task as instructed. Instead they, or at least some of them, may have been monitoring time by counting seconds, and this strategy might explain the smaller underestimates reported by Block.

A second possible reason for the difference in underestimates in Experiment 1 and Block's (1992) experiment arises from the number of intervals participants were asked to produce. In Block's experiment, it is unclear how many times participants produced estimates of the 165 second interval; it is possible that only one production of a single interval was required. By contrast, participants in my Experiment 1 had to produce 27 intervals, 3 for each of the 9 interval durations. The need to produce this many intervals might have resulted in some type of proactive interference. When making the 15th estimate, for example, participants might have lost track of the duration that needed to be estimated on the current trial, and this might account for the larger underestimates in Experiment 1. It is also possible that 'estimating fatigue' played an important role in Experiment 1 and that more fatigued participants make shorter estimates than less fatigued participants.

The pattern of underestimates in Experiment 1 points to a flat relationship between target interval duration and estimate duration (i.e., if the 1 minute interval is eliminated from analysis). This pattern raises the possibility that estimates and actual durations may form a type of Weber fraction. By contrast, in previous research examining the relationship between estimates and actual durations, a negative linear relationship has typically been found in connection with very short intervals (c.f., Church & Gibbon, 1982; Treisman, 1990). That is, previous research showed that estimates (note: short duration intervals are typically overestimated), expressed as a proportions, become longer with a decrease in the actual duration of intervals. The existence of a different relationship between estimates and actual durations in the present study which focused on everyday intervals, versus in previous research, where the focus was on very short intervals, strengthens the possibility that different mechanisms are used for timing long versus short intervals. However, such a conclusion must be made with caution, especially in view of the fact that in

Experiment 1, time estimates were made in the context of a specific concurrent task which was designed to prevent strategic time monitoring. By contrast, in many previous studies time estimates were made in the absence of a concurrent task, or in the presence of concurrent tasks which may or may not have prevented strategic time monitoring.

Experiment 1 revealed a difference between estimates of 1 minute intervals versus longer intervals, suggesting a boundary between short and long intervals located around the 1 minute point. Boundaries between intervals of differing durations have been noted in other areas of research on timing. One such case is the Kappa Effect, the finding that when stimuli are presented at longer temporal delays, their spatial distance is judged as larger. This temporal dilation appears to have a lower boundary of approximately .6 seconds and an upper boundary of approximately 7-11 seconds (Jones & Huang, 1982).

Experiment 2 was designed to investigate further the possibility of a boundary in estimating relatively short everyday intervals versus longer every day intervals. Experiment 2 had a number of additional objectives: to replicate the findings of Experiment 1 with regards to the relationship between target interval and estimate duration, to limit the possibility that proactive interference might have caused the large underestimates found in Experiment 1, and to explore the involvement of memory processes in estimating durations. As in Experiment 1, participants produced intervals of differing lengths while engaged in a secondary task. In order to limit the role of proactive interference, the number of intervals was reduced to three (2, 4, and 6 minutes) and participants were given a reminder of the current interval to produce in the corner of the computer screen.

Consistent with the possibility that different mechanisms might be involved in timing very short intervals (of the sort investigated by previous research) versus in timing every day intervals, Experiment 2 was also intended to investigate assumptions about timing that are suggested by memory storage models

(Block, 1978; Block & Reed, 1978; Ornstein, 1969). As reported in Chapter 1, memory storage models argue that when estimating intervals, we reconstruct the time that has passed based on the number of chunks stored in memory. If more chunks are stored, we estimate that more time has passed. Both Block and Reed (1978) and Poynter (1983) used ongoing tasks that were designed to manipulate the number of chunks stored in memory for the interval, but in both cases the manipulation could also have influenced the amount of attention directed towards timing, making it difficult to determine whether memory storage models or the AGM best explain timing in the range of everyday intervals. The manipulation used in Experiment 2 was designed to avoid this potential confound by having participants complete exactly the same ongoing task and using framing manipulation instead of task difficulty or task switching manipulation to influence the number of stored chunks. Participants were instructed to focus on the future or on the present, providing a different organizing framework for the ongoing task and potentially a different number of chunks stored in memory.

Chapter 3 – Future or Fun?

“Time flies when we’re having fun”. “A watched pot never boils”. Both are common sayings which, on their face, seem to describe our perception or experience of the duration of everyday tasks. Anecdotally, our experience of time seems heavily influenced by our environment, our expectations, our perspective, and our mood. When we are engaged in an enjoyable task, time seems to fly by. We look up from a conversation with a good friend or an interesting book only to find that what felt like minutes was actually hours passing by. By contrast, when we are anticipating something, like the water in the kettle to boil for the first cup of tea of the day, it seems to take forever.

These differing subjective experiences raise a number of important questions: Is our subjective experience of time reflected in behavior, that is, is it a valid index of how we act in situations that require timing? If our timing behavior is influenced by whether or not we are having fun, how large is this influence? If a “fun” effect exists, is it really a result of engaging in a fun or boring task, or is there some other property common to these experiences that is driving the effect? If there is a “fun effect” in our timing behavior, is the magnitude of this effect the same for short (i.e., two minutes) and long duration tasks (five or six minutes)? Are the AGM or memory storage models able to provide insight into how time estimates change with having or not having fun? In addition to building on Experiment 1 and further exploring the relationship between timing estimates and event durations, Experiment 2 was conducted to investigate whether prospective timing is influenced by task properties that may be at the heart of sayings like “time flies when we are having fun”.

If a timing fun effect exists, it might occur for at least three different reasons, and the AGM seems to be marginally consistent with only one of them. To recapitulate, the AGM consists of a pacemaker that emits pulses at regular intervals, and an accumulator that counts these pulses and compares them to stored values in reference memory (Zakay & Block, 1996, 1998). Between the pacemaker and the accumulator is a gate, which is controlled by the amount of attention allocated to timing. The AGM states that when more attention is placed on timing, the attentional gate is open more widely, and pulses accumulate more quickly. Consistent with the AGM, a first account of a fun effect in timing might go as follows. When we are engaged in an enjoyable or interesting task or when we are having fun, less attention (assumed to be top-down, divided attention) is placed on time passing, compared to when we are engaged in a boring or uninteresting task and allocate more attention to the passage of time. In the first case, the attentional gate is narrowed, with fewer pulses accumulating per unit of time (e.g., 1 minute); in the second case, the attentional gate is wide open and more pulses accumulate per unit of time. Therefore, when comparing an accumulated pulse count to stored values in reference memory, the "fun count" would correspond to a shorter interval than the "boring count".

Although consistent with the AGM, this type of interpretation is post hoc and depends on a questionable operationalization of what it means to allocate "attention to timing". It is possible that some fun tasks direct attention away from timing and some boring tasks direct attention to timing, but is not reasonable to assume that an enjoyable task always results in less attention directed towards timing, or that a boring task always results in more attention directed towards timing. It is easy enough to think of enjoyable tasks that focus attention on timing, such as when watching a fast-paced basketball game in which there is a shot clock. Similarly, it is easy to imagine boring tasks that draw attention away from timing, such as when entering data into a database. In view of such counterexamples, an AGM interpretation of the timing fun effect which treats fun and boring tasks as if they were mere proxies for the amount of attention directed towards to timing does not seem warranted.

A second explanation for a timing fun effect postulates that this effect is due to the difference in the physiological arousal that accompanies enjoyable and boring tasks. Internal clock models, such as Treisman's (1963) describe a pacemaker that emits pulses, much like the AGM, but the rate of pulse generation is determined by physiological arousal. High physiological arousal is assumed to result in a higher rate of pulse generation, and low physiological arousal results in lower pulse generation. It is possible that enjoyable tasks are associated with higher physiological arousal (Treisman 1963), resulting in a higher rate of pulse generation, and a corresponding higher pulse count. By contrast, boring tasks may be associated with lower physiological arousal, resulting in a lower rate of pulse generation, and a corresponding lower pulse count. However, there is no reason to assume that all enjoyable tasks involve higher physiological arousal. Two tasks, for example, sitting at home reading a favourite book, and hiking with friends, may be equally enjoyable, but they are associated with very different levels of physiological arousal. For such reasons, it does not appear that physiological arousal can account for the "fun effect".

A third interpretation of the timing fun effect builds on research on framing. A frame is a kind of cognitive scaffold, a particular way of describing or thinking about a problem that changes how we perceive and respond to it (Druckman, 2001). Our perceptions of the outcomes of different decisions, the probabilities of alternatives, and even the fundamental problem being posed all make up a frame that influences the decisions we make. For example, a 10 year old who had lunch at noon and asks for a snack before their 6:00 pm dinner could be told that she ate only three hours ago, or to wait because supper is only three hours away. Both statements are factually equivalent, but they frame the problem (i.e., the arrival of the next meal) in two different ways. In the first case, the child's problem is framed with respect to the past; in the second case, it is framed with respect to the future. In turn, the frame is likely to influence the child's perception of time (i.e., how long it will be till the next meal arrives). Consistent with this frame perspective, it is conceivable that the "fun effect" is not due to a difference in attention towards timing, but that it occurs because fun and boring tasks typically involve different frames, which result in different experiences of time.

Framing is known to influence choices or decisions in many different problem solving tasks (Druckman, 2001; Dunegan, 1993; Tversky & Kahneman, 1981). In a typical framing experiment, a problem is presented, together with a frame which focuses either on losses versus gains. In a study by Tversky and Kahneman (1981), half of the participants were asked to choose between a medical intervention that would save 200 out of 600 people versus a medical intervention that had a 1/3 probability that all 600 would be saved and 2/3 probability that nobody would be saved. The other half of their participants were asked to choose between an intervention that would result in 400 out of 600 dying versus an intervention that had a 2/3 chance that everybody would die and a 1/3 probability that nobody would die. Participants in the survival focused condition were more likely to choose the first option, and participants in the mortality focused condition were more likely to choose the second option, despite the numerical probabilities being identical for both groups.

In a time estimation task, framing may serve as a scaffold for organizing and retaining information, and thereby influence the perception of time. When we are having fun, we are focused on the present, and this ‘present’ frame might function as a top-down structure for organizing ongoing events into a relatively small number of chunks. By contrast, when we are not having fun, our focus is elsewhere (e.g., on the past or the future), and there is no frame for grouping ongoing events. Instead, ongoing events are encoded and remembered as a relatively large collection of disjoint units. According to memory storage models (Block, 1978; Block & Reed, 1978; Ornstein, 1969), when required to estimate the duration of an event, we respond based on the number of memory chunks associated with it. As a consequence, on prospective timing tasks we would expect participants who are focused on the present to signal the end of a target interval later than participants focused on the future, as they would have fewer stored chunks and believe less time had passed.

Indirect evidence that is consistent with this type of event chunking hypothesis comes from research on retrospective timing, which has demonstrated that intervals filled with more segmented information or more changes in context are judged to be longer than intervals filled with less segmented information or fewer changes in context (Block, 1992; Brown, 1985; Bueno Martinez, 1994; Poynter, 1983). Additional evidence favoring a framing hypothesis comes from a study by Liu, Li and Sun (2014). They presented participants with different descriptions of time intervals, and asked them to indicate the length of each interval, using a scale with endpoints marked, respectively, "very short" and "very long". The time intervals were either packed or unpacked. Packed intervals were defined in terms of nothing more than a start and end point (e.g., "From June 1st of this year to the end of the following fourth month") (Liu et al., 2014, p. 2). Unpacked intervals were defined in terms of several points between the start and end point (e.g., "From June 1st of this year, passed by the next month, the month after next, the following third month to the end of the following fourth month") (Liu et al., 2014, p. 2). Participants' estimates of duration were significantly longer in the unpacked condition compared to the packed condition. The finding that durations composed of a large number of chunks are perceived as longer than durations composed by a smaller number of chunks seems to support the hypothesis that the fun effect in timing could be due to the different frames that are instantiated by focusing on the present moment versus focusing on a future event.

Experiment 2 had three main objectives – to replicate the findings of Experiment 1 with regards to the relationship between target interval duration and estimate duration, to investigate the potential influence of proactive interference in Experiment 1, and most importantly, to investigate the effect of memory processes in timing everyday intervals. In Experiment 2, participants produced intervals of 2, 4, and 6 minutes while engaged in an ongoing task. To reduce the effects of proactive interference, participants had fewer target intervals to produce and were given a reminder of the current target interval in the corner of the computer monitor. In order to investigate the role of memory processes, I used a framing manipulation, which was intended to affect participants' overall framing of the timing task as either present- or future-oriented, without changing the amount of attention directed towards the timing task. The AGM makes no direct prediction about how this type of manipulation might affect the length of productions, since the main factor in the model that determines the length of productions is the amount of attention directed towards timing. Consistent with the framing hypothesis described earlier, I predicted that participants in the present frame condition would produce longer intervals.

Method

Participants and design. Thirty-six undergraduate student volunteers from the University of British Columbia participated in this experiment in return for course credit. The experiment had one between-subjects factor, task framing (present frame or future frame), and one within-subjects factor, target interval duration (two, four, or six minutes). Eighteen participants were randomly assigned to each task framing condition.

Procedure. All participants were tested individually using a combination of tasks presented with e-Prime on a desktop computer and paper-and-pencil measures. At the beginning of each session, participants were asked to remove watches and to place any time-tracking electronics out of sight. The experiment took approximately one hour to complete. The experiment was conducted with the approval of the University of British Columbia Behavioural Research Ethics Board.

Each participant performed two tasks – a word generation task and a trivia task, and each task consisted of a series of self-paced trials. Each series of trials of the same task formed a block. In the experiment, participants alternated between blocks of the word generation task and blocks of the trivia task. The experiment began with the word generation task. For each trial of the word generation task, a common English word with 6-8 letters was displayed in the center of the computer monitor. Participants were instructed to use letters from the displayed word in order to form new words that used two or more of the same letters, and to record these words on a sheet of lined paper. When they could not think of any new words, participants pressed the Enter key to display a new word.

The trivia task blocks consisted of a series of six trivia questions, selected from a website with high school level Canadian trivia questions (e.g., what is the official winter sport of Canada?). Question topics included provincial flowers, capital cities, sports, and other general knowledge. Each question was displayed in the center of the computer screen with four response options. Participants recorded their answer for each question on paper, and then pressed Enter to display the next question.

Participants in both the present frame and future frame conditions completed exactly the same tasks. An instructional manipulation was used in order to assess the effect of framing of the timing task (present- or future-oriented) on prospective interval production. Participants in the present frame condition were told that their main task was to continue the word generation block for the duration of the target interval (e.g., for two, four, or six minutes). Participants in the future frame condition were told that their main task was to start the trivia block on time, after the target interval had elapsed (e.g., after two, four, or six minutes). When the participant believed the target interval had elapsed, they pressed the spacebar to terminate the block. The participant then completed a trivia block. The trivia block, in which participants answered simple Canadian trivia questions, was included to ensure that participants in the future frame condition had a distinct task to look forward to that was unrelated to the word generation task.

Each participant completed three word generation blocks per target interval of two, four, and six minutes, for a total of nine word generation blocks. Word generation blocks were presented in random order. If participants failed to make a response to the target interval production task after 8 minutes, the block terminated and the trivia block began immediately.

Results

Word generation and trivia task performance. For each participant, the number of word generation trials and number of words generated was calculated. The mean number of word generation trials was 78.33 (*SD* = 27.18) in the present frame condition and 80.21 (*SD* = 23.25) in the future frame condition. The mean number of words generated was 181.33 (*SD* = 60.83) in the present frame condition and 216.95 (*SD* = 59.14) in the future frame condition. The mean number of words generated per trial was 2.87 (*SD* = 2.56) in the present frame condition and 3.41 (*SD* = 3.40) in the future frame condition. None of these differences were significant in *t*-tests with condition as the between-subjects factor. The largest obtained *t* value was for the number of words generated, *t* (35) = 1.81, *p* = .08. Accuracy on the trivia task also did not differ between the present and future frame conditions, with means of 8.67 (*SD* = 2.00) and 8.95 (*SD* = 3.01) respectively, *t* (35) = .332, *p* = .74.

Interval production performance. The key dependent variable of interest was the closeness of the time estimate produced by each participant. Estimates were converted to relative difference scores, expressed as a percentage of the target interval, using the same formula as in Experiment 1. Each participant's median estimate for each target interval (i.e., the median estimate over all of the two minute blocks, the median estimate over all of the four minute blocks, and the median estimate over all of the six minute blocks) was used in the following analyses. As described above, participants responded by key press to indicate when they thought the target interval had elapsed. In the event that participants did not respond and the program terminated the timing block, the program recorded a 0 for the participant's response time. Blocks with a response time of 0 were discarded, and the median of the remaining estimates was taken. In total, 5 trials were eliminated as a result of these program terminated trials. No participants were eliminated as a result of having more than one program-terminated trial.

Figure 3.1 shows the means of the median estimates for each interval and task frame. Participants underestimated the four and six minute intervals and overestimated the two minute intervals. Participants in the present frame condition produced longer estimates for all intervals. The results of an ANOVA with target interval as a within-subjects factor and task frame as a between-subjects factor confirmed this summary of the findings by revealing a significant main effect due to target interval, $F(1, 24) = 17.30$, $p < .001$, $\eta^2 = .34$, and task frame, $F(1, 34) = 4.83$, $p = .04$, $\eta^2 = .12$. No other effects were significant.

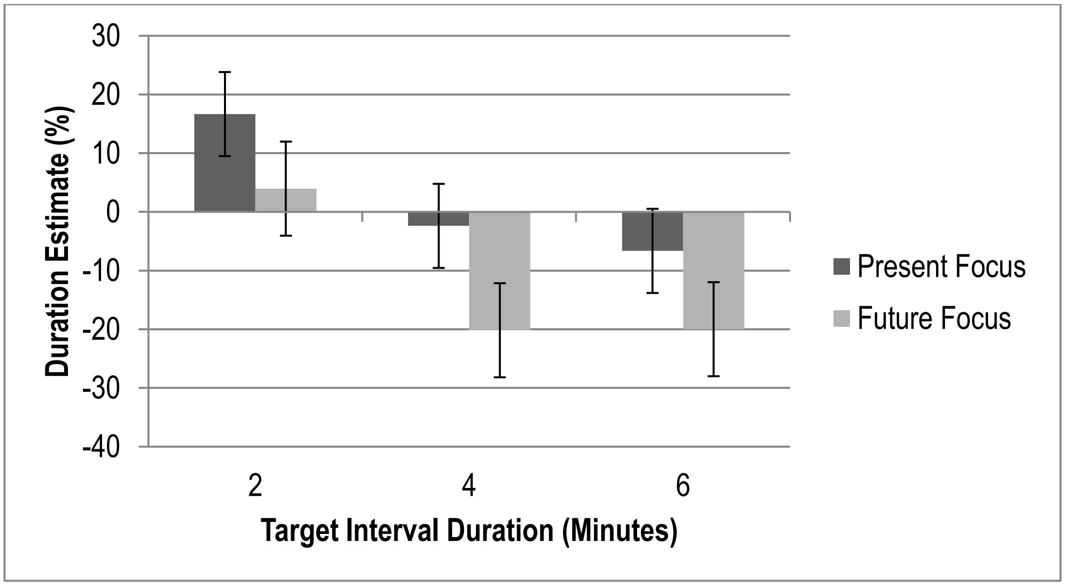

Figure 3.1. Duration estimates for 2, 4 and 6 minute target intervals in the future and present frame conditions. Negative values are underestimates; positive values are overestimates. Error bars represent the standard error of the mean.

Discussion

Experiment 2 investigated whether task framing could account for the fun effect in timing. Between conditions, I used an instructional manipulation to create a frame intended to focus attention either on the present, ongoing task or on a future task. Three critical findings emerged and need to be discussed: the longer estimates in the present frame condition compared to the future-frame condition, the longer estimates for the 2 minute interval compared to the 4- and 6- minute intervals, and the relationship between target duration and time estimates.

The predictions about the frame manipulations were supported. For each target interval, participants in the present frame condition produced estimates that were longer than those produced by participants in the future frame condition. This outcome is consistent with the assumption that focusing on the present may serve as an organizing structure for a series of events within an interval, resulting in fewer chunks within the interval. Participants may use the number of chunks as a method of estimating the time that has passed, with the assumption that more chunks is equivalent to a longer interval. When engaged in a prospective timing task, this results in longer productions, as more time has to pass in order to accumulate the target number of chunks.

The present research complements previous investigations which used different methods for manipulating the number of memory chunks that may be formed in the course of a target interval. A more typical method for manipulating chunking within an interval involves requiring participants to switch between two tasks or two types of stimuli. For example, Bueno Martinez (1994) had participants reproduce durations of 80 seconds while engaged in a structural processing task (i.e. attend to the font of words), a semantic processing task (i.e. attend to the category membership of words), or a mixture of processing tasks (i.e. attend to the font of words on some trials and the category membership of words on other trials). Bueno Martinez found that participants produced longer estimates when the task involved switching between structural and semantic processing than when it required only one type of processing. Consistent with this description, it might be argued that the method used by Bueno Martinez (1994) manipulated the number of memory chunks from the bottom-up. By contrast, with the type of framing manipulation that was used for Experiment 2, the number of memory chunks seems to be controlled by top-down influences.

The present research is also consistent with the work of Alards-Tomalin, Leboe-McGowan, Shaw, and Leboe-McGowan (2014) who used a radically different method for creating memory chunks. In a series of experiments, their participants were asked to judge the duration of visually presented stimuli. Specifically, in one study, participants were shown several series of 3 digits, using an inter-stimulus interval (ISI) between 620 ms to 785 ms. Each series consisted of either two digits close to each other on the number scale, followed by a third that was further along the scale (e.g., 1, 2, 9), or two digits far apart from each other on the number scale, and a third that was close to the second number in the series (e.g., 1, 7, 9). The three digits were presented in either a long-short ISI pattern (e.g., 785 ms between the first and second digits, and 680 ms between the second and third digits) or a short-long ISI pattern (e.g., 620 ms between the first and second digits and 752 ms between the second and third digits). Participants were asked to decide if the first or second ISI was longer. Alards-Tomalin et. al (2014) found that when two digits were further apart on the number scale (e.g., 1 and 7), participants were more likely to identify the ISI that separated them as the long ISI than when the two digits were close in numeric value (e.g., 1 and 2). Alards-Tomalin et al. concluded that the more similar two targets were to each other, the more likely they are to be perceived as a single unit, allowing them to be processed as one item or chunk rather than a series of unrelated items. The authors suggested that participants used the number of memory units for making time judgments, though they did not describe a mechanism for mapping memory units onto time estimates. The authors also did not describe how digits positioned on different points along the number scale might be used for creating different memory chunks. However, it is conceivable that the conceptual representation of the number scale acted as a top-down constraint to influence memory chunking in a manner that is analogous to the present/future frame manipulation used in Experiment 2.

In Experiment 1, participants underestimated all intervals, but this was not the case in Experiment 2 (Note: only the 4 minute interval was included and underestimated in both Experiments 1 and 2, and thus only it permits a direct comparison). In Experiment 1, participants underestimated the 4 minute interval by approximately 35%, whereas participants in the future frame condition of Experiment 2 underestimated the 4 minute interval by 20%. This difference in estimates may be due to differences in the ongoing tasks. Specifically, it is possible that the ongoing task used in Experiment 1, in which participants viewed grids filled with dots and had to respond with the number of dots they saw, was less attention demanding than the ongoing task used in Experiment 2, where participants had to generate novel words from cue words. A large body of evidence from research on timing of shorter intervals suggests that time interval productions are influenced by the amount of attention that is directed towards timing (e.g., Brown, 1985; Brown, 1997; Brown, Collier, & Night, 2013; Hemmes, Brown, & Kladopoulos, 2004; Hicks et al, 1976; Kladopoulos et al., 2004; Macar, Grondin, & Casini, 1994; McClain, 1983; Zakay & Tsal, 1989). Consistent with the AGM, if the attentional demands of the ongoing task in Experiment 1 were lower than those of the ongoing task used in Experiment 2, we would expect to see larger underestimates in Experiment 1 than in Experiment 2. Due to the differences in the ongoing tasks used in Experiments 1 and 2, this possibility is difficult to rule out, as there is no index available for verifying the attentional demands of the ongoing tasks used in Experiment 2.

An alternative possible reason for the larger underestimates in Experiment 1 than 2 is linked to differences in the methods. Participants were required to produce 27 intervals in Experiment 1 compared to only 9 in Experiment 2. In addition, in Experiment 2 but not 1, participants were given a reminder (always in view in the corner of the computer screen) about the duration of the to-be-produced target interval. As mentioned in the discussion of Experiment 1, the large number of intervals estimated in Experiment 1 might have produced fatigue and especially proactive interference, and thereby contributed to the larger underestimates in Experiment 1. Consistent with these speculations, it would seem likely that the time productions in Experiment 2 are more valid, as participants experienced less fatigue and proactive interference.

The finding that participants in Experiment 2 produced overestimates for the 2 minute condition, compared to underestimates in the 4 and 6 minute conditions was not anticipated. One possible explanation for the difference between the 2 minute condition versus the 4 and 6 minute conditions is that participants made estimates based on chunks stored in different memory systems. In the case of the 2 minute condition, participants may rely primarily on the chunks stored in short term or working memory, whereas in the case of 4 and 6 minute conditions participants may rely primarily on the chunks stored in long term or episodic memory. Consistent with this assumption, overestimates and underestimates might occur because short term versus episodic memory chunks differ in size or other properties. Many differences between short term and long term memory have been identified, such as the fact that short term memory is more influenced by phonemic similarity whereas long term memory is more influenced by semantic similarity (Craik & Levy, 1970; Shulman, 1971; Wickelgren, 1965). If phonemic chunks used in short term memory are larger, we would expect that estimation of shorter intervals would be more prone to overestimation than longer intervals, which might rely more on semantic chunks in episodic long term memory. Fewer chunks would be produced if participants were relying on short term memory, and participants would have to wait proportionally longer to accumulate the target number of chunks. This would be consistent with the findings in Experiment 2.

It is also possible that phonemic chunks used in estimating shorter intervals might be remembered less well or be confused with each other more frequently than semantic chunks used in estimating longer intervals. This would be consistent with evidence that phonemic encoding results in poorer recall than semantic encoding (e.g., Craik & Lockhart, 1972; Craik & Tulving, 1975). If this were the case, we would expect estimates to be longer for shorter intervals than for longer intervals, as participants would have to wait longer to accumulate the target number of chunks before deciding that the target interval had been reached.

A starkly different possible explanation for the overestimate in the 2 minute condition versus the underestimate in the 4 and 6 minute conditions is that participants are using different benchmarks for tracking what constitutes a chunk in the 2 minute condition compared to the 4 and 6 minute conditions. If participants were relying on the number of cue words used, versus the number of words generated, we would expect fewer chunks to be generated. This in turn would lead to longer estimates if the participant were using the number of cue words, as it would take longer for the target number of chunks to be accumulated.

Experiment 2 also raises new questions about the relationship between target interval duration and estimate duration. In Experiment 1, there was a Weber fraction-like relationship between target interval and estimate duration for all but the shortest (i.e. 1 minute) interval. Underestimates were proportionally the same for the 1.5 to 5 minute intervals. This is in contrast with the findings from Experiment 2, where participants in both the present frame and future frame condition produced proportionally longer intervals for the 2 minute interval than for the 4 and 6 minute intervals. The small number of intervals tested in Experiment 2 does not permit conclusions about the relationship, as a trend analysis performed on only three intervals would be subject to bias from any random error in the data. Despite the inability to draw conclusions about the shape of the relationship between target interval and estimate duration in Experiment 2, it seems plausible that the results from Experiment 2 are a more valid representation of the relationship between target interval and estimate duration, due to the reduction in fatigue and interference.

If the results from Experiment 2 are more valid, several questions remain unanswered. Does this relationship consist of the pattern highlighted in Figure 3.1, with overestimates for short target intervals and underestimates for longer target intervals? If so, what is the border between short and long intervals? Or is the relationship more linear, with a more gradual decrease in estimate duration across increasingly longer target intervals? A hint at the latter possibility is evidenced in the data from the future frame condition in

Experiment 2, where there is a trend towards shorter productions in the 6 minute condition than the 4 minute condition. From a limited number of target intervals it is difficult to see the underlying pattern of the distribution. Furthermore, it is not clear whether the underlying pattern would be the same in the present and future frame conditions (i.e., a shift in the length of the productions, but the same pattern).

Chapter 4 – Replicating and Extending Experiment 2

Experiment 3 had three main goals: to investigate further the relationship between target interval duration and estimate duration, to determine whether the framing manipulation changes the relationship between target interval and estimate duration, and to ascertain whether the framing manipulation might be indirectly manipulating the allocation of attention.

Experiments 1 and 2 revealed different relationships between interval duration and estimate duration, but each of these experiments had limitations (e.g., too many intervals in Experiment 1; too few intervals in Experiment 2) that caution against drawing definitive conclusions about how target and estimated durations are related. For this reason, Experiment 3 used the methodological improvements introduced with Experiment 2 (e.g., display a reminder about the duration of the currently to-be-produced interval), and required participants to estimate the duration of a greater range of intervals than Experiment 2. In addition, Experiment 3 used the framing manipulation from Experiment 2 in order to investigate if or how this manipulation affects the relationship between actual and estimated durations.

Experiment 3 was also designed to examine whether the framing manipulation introduced in Experiment 2 might affect the attentional resources available for timing. Numerous studies using shorter intervals have demonstrated that when less attention is available for timing, either because of a more difficult ongoing task or because participants have been instructed to pay more attention to the ongoing task than the timing task, estimates are longer (cf. Brown, 1985; Brown, 1997; Brown et al., 2013; Hemmes et al., 2004; Hicks et al, 1976; Kladopoulos et al., 2004; Macar et al., 1994; McClain, 1983; Zakay & Tsal, 1989). One possible explanation for the difference in duration estimates in Experiments 1 and 2 was participants' allocation of attention to the different ongoing tasks. If the ongoing tasks had different attentional demands, we would expect to see a difference in estimates. The lack of accuracy and response time data from the ongoing task used in Experiment 2 makes it impossible to directly compare the two experiments. Additionally, differing attentional demands need to be ruled out as a possible explanation for the difference between the present frame and future frame conditions in Experiment 2. There is no reason to suspect that the future frame and present frame conditions required different amounts of attention, but the possibility could not be ruled out in Experiment 2. The dot counting task used in Experiments 1 and 3 provides both accuracy and response time data, allowing comparison of ongoing task performance in the present and future frame conditions. If it may be the case that the framing manipulation is simply an indirect manipulation of attentional resources, we would expect that performance on the ongoing task would be significantly different. If, on the other hand, the framing manipulation is not simply an indirect manipulation of attentional resources, we would not expect any significant difference in performance on the ongoing task.

Method

Participants and design. In Experiment 3, participants were forty-eight undergraduate student volunteers from the University of British Columbia, who participated in return for course credit. The experiment had one within-subjects factor, target interval duration, with six levels: 1, 2, 3, 4, 5, and 6 minutes, and one between-subjects factor, task framing, with two levels: present frame or future frame. Twenty-four participants were randomly assigned to the present frame condition and twenty-four participants were randomly assigned to the future frame condition.

Procedure. The procedure for Experiment 3 was a combination of the procedures used in Experiments 1 and 2. All participants were tested individually using e-Prime on a desktop computer. At the beginning of the session, participants' consent was obtained, and participants were asked to remove watches and place all time-tracking electronics out of sight. The experiment was conducted with the approval of the University of British Columbia Behavioural Research Ethics Board.

Participants had to carry out three tasks: Produce intervals with specified durations/length, count dots in visual displays, and answer trivia questions. The main or focal task was always the interval production task. This task had to be performed concurrently with the dot counting task. Participants were required to continue with the dot counting task for the duration of each specified target interval. In the present frame condition, they were instructed to press the spacebar to stop the dot-counting task after the specified target interval had elapsed. In the future frame condition, they were instructed to press the spacebar to begin the trivia task after the specified target interval had elapsed.

As part of the consent process, participants were told that they would be producing intervals of 1, 2, 3, 4, 5, or 6 minutes, and that that they should try to be as close as possible to the length of each interval (i.e. not too short; not too long). Each participant completed three productions per target interval of 1, 2, 3, 4, 5, and 6 minutes, for a total of 18 productions. The order of the target intervals was randomized for each participant. While engaged in the interval production task, participants completed the self-paced dot counting task described in Experiment 2. Participants were instructed to be as quick and accurate as possible on the dot counting task. When the participant believed the target interval had elapsed, they pressed the spacebar to terminate the interval production task and the dot counting task. If participants failed to make a response to the target interval production task after 8 minutes, the current target interval production task terminated and the trivia block began immediately. The participant then completed a trivia block, administered using the same method as Experiment 2.

As in Experiment 2, an instructional manipulation was used in order to assess the effect of framing of the timing task (present frame or future frame) on prospective interval production. The instructions used in Experiment 3 were the same as those used in Experiment 2, aside from the change in ongoing task. Participants in the present frame condition were told that their main task was to continue the dot counting block for the duration of the target interval (e.g., for 1, 2, 3, 4, 5, or 6 minutes). Participants in the future frame condition were told that their main task was to start the trivia block on time, after the target interval had elapsed (e.g., 1, 2, 3, 4, 5, or 6 minutes). Participants in both the present frame and future frame conditions completed exactly the same tasks. As in Experiment 2, the trivia block was included to ensure that participants in the future frame condition had a distinct task to look forward to that was unrelated to the dot counting task. The experiment took approximately one hour to complete

Results

Dot counting task. For each participant, the mean accuracy and response time on the dot counting task were calculated. The mean accuracy on the dot counting task was 70.54% (*SD* = 16.07) in the present frame condition and 75.08% (*SD* = 14.07) in the future frame condition. The mean response time was 1710.76 ms (*SD* = 233.65 ms) in the present frame condition and 1939.20 ms (*SD* = 238.66 ms) in the future frame condition. None of these differences were significant in *t*-tests with condition as the between-subjects factor. The largest obtained *t*-value was for mean accuracy, t (46) = 1.04, $p = .30$.

Interval production performance. The key dependent variable of interest was the closeness of the time estimate produced by each participant. Estimates were converted to relative difference scores, expressed as a percentage of the target interval, using the same formula used in Experiment 1. Each participant's median estimate for each target interval (i.e., the median estimate over all of the one minute blocks, the median estimate over all of the two minute blocks, etc.) was used in the following analyses. As described above, participants responded by key press to indicate when they thought the target interval had elapsed. In the event that participants did not respond and the program terminated the timing block, the program recorded a 0 for the participant's response time. No participants were eliminated as a result of having more than one terminated block.

Figure 4.1 shows the means of the median estimates for each interval and task frame. Participants underestimated all of the intervals with the exception of the 1 minute interval. A repeated measures ANOVA with target interval (1, 2, 3, 4, 5, and 6 minutes) as a within-subjects factor and task frame (future vs. present) as a between-subjects factor revealed significant main effects of target interval, $F(1, 46) = 242.08$, $p < .001$, $\eta^2 = .25$, and task frame, $F(1, 46) = 184.387$, $p < .01$, $\eta^2 = .24$. The interaction between target interval and task frame was not significant, *F*(1, 46) = .006, $p > .10$, $\eta^2 = .00$. Separate trend analyses were conducted on the estimates in each frame condition. In the future frame condition, there was a significant linear trend, *F*(1, 23) = 11.41, $p < .01$, $\eta^2 = .33$. In the present frame condition, there was a significant linear trend, *F*(1, 23) = 11.70, $p < .01$, $\eta^2 = .33$, and significant quadratic trend, *F*(1, 23) = 4.65, $p < .05$, $\eta^2 = .17$. No other effects were significant.

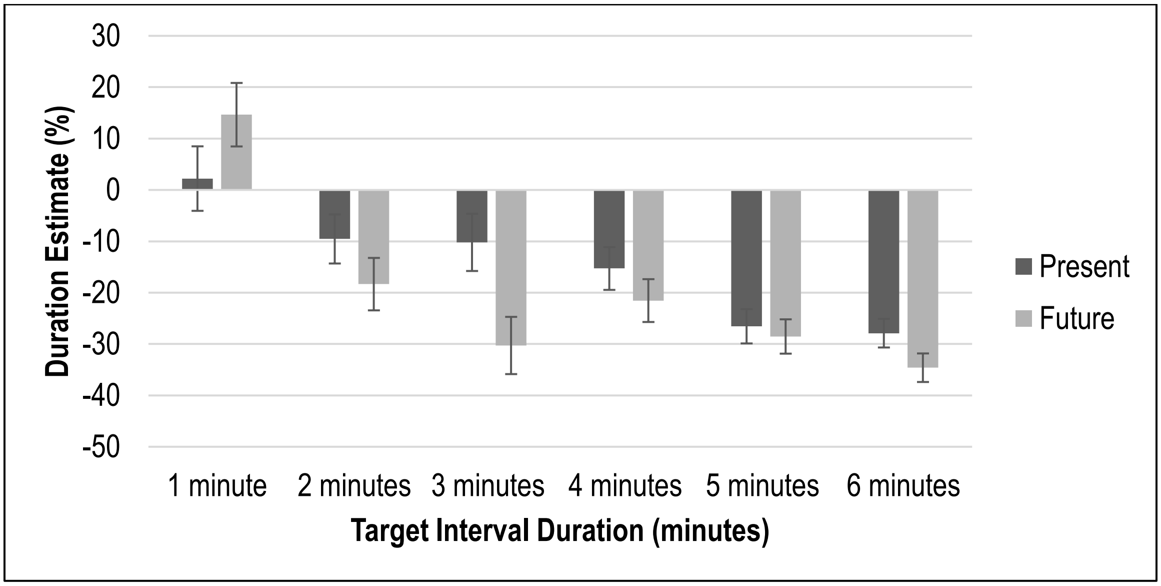

Figure 4.1. Duration estimates for 1, 2, 3, 4, 5, and 6 minute target intervals in the future and present frame conditions. Negative values are underestimates; positive values are overestimates. Error bars represent the standard error of the mean.

Discussion

Experiment 3 was designed to replicate and extend the framing effect found in Experiment 2, to determine whether the framing manipulation might simply be an indirect attention manipulation, to investigate the relationship between target interval and estimate duration, and to determine whether the framing manipulation used in Experiment 2 changes the relationship between target interval and estimate duration. Three findings require further discussion: the ongoing task performance in the present and future frame conditions, the effect of the framing manipulation on estimates produced in the present frame and future frame conditions, and the difference between estimates produced for different target intervals in the present and future frame conditions.

One possibility raised in the Discussion of Experiment 2 was that the framing manipulation simply changed the amount of attention available for timing, that is, that it acts as an indirect manipulation of attention. To address this possibility, participants in Experiment 3 completed a simultaneous ongoing task that provided both accuracy and response time data. No significant differences were found in either performance accuracy or response time data between the present and future frame conditions. This lack of significant differences appears to indicate that the framing manipulation is not simply a proxy for an attentional manipulation.

A framing effect was observed in Experiment 3. Participants produced shorter estimates in the future frame condition than in the present frame condition, though this was not true for all intervals. Shorter estimates in the future versus the present frame condition are consistent with the assumption that focusing on the present may serve as a more effective organizing structure for a series of events within an interval, resulting in fewer chunks within the interval.

The framing effect was substantially smaller in Experiment 3 (i.e. across all target intervals, estimates were 5.23% larger in the present than future frame condition) than in Experiment 2 (i.e., across all target intervals, estimates were 14.60% larger in the present than future frame condition). One possible explanation for this difference in the frame effect size is random variation, due to the use of different participants and comparison across experiments.

Another and more interesting possibility is that the size of the frame effect depends on the processing requirements of the ongoing task. For Experiment 2, the ongoing task required participants to generate novel words on the basis of a cue word. For Experiment 3, the ongoing task required participants to count the number of dots displayed on a computer screen. The latter task is more structured and more data driven then the word generation task, which is more subject-initiated and guided. This distinction may result in different processing and encoding of events within an interval (Craik, 1986). In the dot counting task, the bottom-up constraints imposed by this task might prevent or reduce the ability to chunk effectively in the present frame condition, reducing the size of the framing effect. In the word generation task, participants can use the top-down structure of the task to organize the information based on the cue words viewed or the target words generated, resulting in more effective chunking or memory for the interval that had elapsed. Additionally, the subject-initiated nature of the word generation task may result in better memory for the cues and words generated during the interval, reducing the attentional resources necessary for completion.

The framing effect was not consistent across all intervals. For the shortest interval (1 minute), participants in the present frame condition produced estimates that were shorter than those produced by participants in the future frame condition. This pattern is the opposite of that observed with the longer intervals in Experiments 2 and 3. This is not the first time that estimates of 1 minute target intervals have not fit the pattern observed in longer intervals: in Experiment 1, estimates of the 1 minute target interval were different than estimates of longer intervals. Taken together, the differences between the 1 minute intervals and longer intervals in Experiments 1 and 3 may indicate that estimating the duration of 2 or more minutes differs qualitatively from estimating the duration of shorter intervals. If that is the case, it may be that separate models best fit the data from timing short intervals (perhaps <2 minutes) and longer intervals (>2 minutes), though follow up research is necessary to pinpoint where the differences emerge.

In Experiment 3, participants' estimates became proportionally shorter as the target interval lengthened. This is consistent with the pattern observed in Experiment 2, though the small number of intervals tested in Experiment 2 made the relationship difficult to test. Neither the AGM nor memory storage models provide an explanation for why estimates are proportionally shorter for longer intervals. One possibility is that participants have less experience timing and receiving feedback on longer intervals. We are less likely to rely on external timing devices for shorter intervals and are more confident in our ability to time them, resulting in greater experience timing these intervals and receiving feedback on our performance. When timing longer intervals, we may be more likely to reach for external devices or to subdivide them into shorter intervals in order to increase accuracy. This lack of precise stored values in reference memory would result in inaccurate estimates, although it is not clear why they would be underestimates.

Although estimates in both the present and future frame conditions became proportionally shorter as the target interval lengthened, the relationship between target interval and estimate duration was not the same in both framing conditions. In the future frame condition, the relationship between target interval and estimate duration was linear. In the present frame condition, both the linear and quadratic trend were significant. This difference in the relationship between target interval and estimate duration may reflect participants' differing organization of the intervals under present and future frame conditions. If the present frame condition acts as a more effective organizing structure, one possibility is that each chunk is larger than it would be in the future frame condition, and that the size of the difference is consistent across the entire length of the target interval. This would result in estimates that were longer, but we would expect the same underlying relationship between target interval and estimate duration in the present frame and future frame condition. If, on the other hand, the present frame results in chunks that are larger and that the size or organization of the chunks changes as the interval lengthens, then we would expect to see a different relationship between target interval and estimate duration in the present and future frame conditions. The results of Experiment 3 seem to be pointing towards the latter explanation, but further research with longer target intervals and possibly different stimuli that lend themselves to different chunking methods may help resolve this question.

Two critical questions remain unanswered: why do participants consistently underestimate longer intervals, and do estimates differ based on the amount of attention available for timing? One possible explanation for participants' underestimates is that the reference memory component of memory storage models and of the AGM contains inaccurate stored values for longer target intervals. If this is the case, then giving feedback may result in more accurate productions. Attention may also play a role in timing longer intervals. Previous research in prospective timing has found that productions are shorter when participants have more attention available for timing, but the effect of attentional manipulations on timing of everyday intervals is unknown. Experiment 4 was conducted in order to more thoroughly investigate the role of reference memory and attention in timing everyday intervals.

Chapter 5 – Learning to Estimate the Duration of Brief Tasks

Experiment 4 was designed to answer three main questions: does the amount of attention available for timing influence the duration of estimates; does feedback change the estimates produced by participants; and does the availability of attention resources influence the benefits derived from receiving feedback?

Experiments 1, 2, and 3, reported in previous chapters, revealed that participants often underestimate everyday intervals. The exact size of participants' underestimates appears to depend on the duration of the interval and on the framing of the task. What is responsible for the production of underestimates? One possibility is that underestimates are caused by a lack of precise stored values in reference memory. Less experience timing longer intervals may result in an imprecise reference memory for what constitutes the "correct" interval. If this is the case, and the lack of accurate reference memory is because of a lack of experience rather than an inability to encode these stored values, it follows that providing feedback to participants should result in estimates that are more accurate.

The effect of feedback on the duration of interval estimates has been studied previously, though largely in the range of seconds rather than minutes. Philbin and Seidenstadt (1983) found that participants timing intervals from 3 to 23 seconds improved their interval estimation accuracy significantly when given feedback about their accuracy to the nearest 0.1 second. Brown (2008) found that participants engaged in a reproduction task of intervals from 6 to 14 seconds who were given feedback about the magnitude and direction of their timing error (e.g., "your estimate was 2.1 seconds too long") showed a decrease in timing error compared to participants who did not receive feedback.

Feedback results in more accurate estimates, but does the type of feedback matter? Ryan and Fritz (2007) presented participants in a production and reproduction timing task with feedback about their timing accuracy. Participants produced or reproduced intervals from 3.5 to 14.5 seconds. Prior to receiving feedback, participants in the production condition produced intervals that were longer than the target interval (i.e., were overestimates), and participants in the reproduction condition reproduced intervals that were shorter than the target interval (i.e., were underestimates). In order to investigate the role of feedback and varying feedback types, some participants received feedback that accurately represented their performance, while others received feedback that minimized or exaggerated their estimate errors. Following each timing trial, participants in the feedback groups were presented with a rectangle that represented the actual duration of the target interval, and a second rectangle that represented their production or reproduction. In the 80% feedback group, the second rectangle was 80% of the size that would correspond to the participants' actual estimate (i.e., in the production condition, where participants typically overestimated the interval, the 80% feedback group saw a rectangle was closer to the target interval than their actual production, minimizing their error). In the 120% feedback group, the second rectangle was 120% the size that would correspond to the participants' actual estimate (i.e., in the production condition, the rectangle was further from the target interval than their actual production, exaggerating their error). A fourth group received no feedback. The group that received no feedback did not produce more accurate intervals after repeated trials. All feedback groups produced and reproduced more accurately with practice. Participants who received feedback that exaggerated their error gained the most accuracy following feedback, followed by participants who received accurate feedback, followed by participants who received feedback that minimized their error. Ryan and Fritz interpreted this finding to suggest that feedback is used to refine reference memory in the prospective timing task, and to shift the decision threshold in reproduction tasks.

Feedback may help participants refine the stored values in reference memory for the target intervals. Feedback has been demonstrated to result in more accurate productions and reproductions in short intervals (Brown, 2008; Philbin & Seidenstadt, 1983; Ryan & Fritz, 2007), but not in longer, everyday intervals. Stored values in reference memory may be imprecise for longer intervals, either because we have less experience timing longer intervals or because there is a wide range of acceptable response times for most everyday timing tasks (i.e., it is not critical whether your tea stops steeping at exactly 4 minutes since both 3.5 and 4.5 minutes result in acceptable tea). It may be possible to refine them by giving feedback immediately following participants' productions of these intervals. While it seems intuitive that feedback would help refine reference memory, it may be the case that with long intervals we are less able or less likely to use it to refine our reference memory.

Attention may influence our ability to use feedback to refine our reference memory for longer intervals. Attention has been shown to affect encoding in episodic memory (Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Naveh-Benjamin, Ishaik, & Anderson, 2000; Naveh-Benjamin & Guez, 2000; Naveh-Benjamin, Guez, & Marom, 2003). When participants' attention is divided between an episodic memory task and a secondary, attention-demanding task, encoding in the episodic memory task is impaired. Reference memory may follow a similar pattern, in which case we would expect that feedback would be less effective when paired with a difficult ongoing task than when paired with an easy ongoing task.

Experiment 4 was also designed to investigate the role of attention in producing everyday intervals. The AGM predicts that when more attention is available for and directed towards timing, participants will respond earlier on a production task, resulting in larger underestimates (Zakay & Block, 1998). This prediction is consistent with a large body of experimental work on timing shorter intervals that focuses on manipulating the attentional resources available for or directed towards timing (cf. Brown, 1985; Brown, 1997; Brown, Collier, & Night, 2013; Hemmes et al., 2004; Hicks et al, 1976; Kladopoulos et al., 2004; Macar et al., 1994; McClain, 1983; Zakay & Tsal, 1989). Such experiments usually rely on a directional manipulation (i.e., emphasizing that the timing task is more or less important than the ongoing task) or a load manipulation (i.e., manipulating the difficulty of the ongoing task). The role of attention in timing longer, everyday intervals has not been determined. If it is the case that the amount of attention directed towards timing influences the duration of estimates, then it may be necessary to modify the AGM to account for the effect of the framing manipulation used in Experiments 2 and 3. Conversely, if there is no effect of the amount of attention directed towards timing, it may be the case that memory storage models best account for timing in the range of everyday intervals. In order to investigate the effect of attention, Experiment 4 used a variation on the dot counting tasks used in Experiments 1 and 3 which allowed for the manipulation of the amount of attention available for timing.

Experiment 4 was conducted to investigate the effect of attention and feedback on the duration of prospective time productions in the range of everyday tasks. Participants completed a timing task and received feedback for each interval produced. In order to determine whether attention plays a role in timing these longer intervals, participants completed either an easy or a difficult version of the dot counting task. In order to determine whether error exaggerating or error minimizing feedback is more likely to result in learning, three feedback types were used. Some participants were given feedback that exaggerated their estimate error, some participants were given feedback that minimized their estimate error, and some were given accurate feedback. One possibility is that receiving feedback that exaggerates estimate errors will result in a larger shift in the length of productions than feedback that minimizes estimate errors. Another possibility is that receiving feedback that exaggerates estimate errors is discouraging and will result in a smaller shift in the length of productions than feedback that minimizes estimate errors.

Method

Participants and design. In Experiment 4, participants were fifty-four undergraduate student volunteers from the University of British Columbia, who participated in return for course credit. The experiment had one within-subjects factor, target interval duration, with three levels: 2, 4, and 6 minutes, and two between-subjects factors, ongoing task difficulty (easy vs. difficult) and feedback type (minimized estimate errors, exaggerated estimate errors, accurately represented estimate errors).

Procedure. All participants were tested individually using e-Prime on a desktop computer. At the beginning of the session, participants' consent was obtained, and participants were asked to remove watches and place all time-tracking electronics out of sight. The experiment was conducted with the approval of the University of British Columbia Behavioural Research Ethics Board.

Participants had to carry out two tasks: Produce intervals with specified durations/length, and count dots in visual displays. The main or focal task was always the interval production task. This task had to be performed concurrently with the dot counting task. Participants were required to continue with the dot counting task for the duration of each specified target interval. In all conditions, participants were instructed to press the spacebar to indicate when they thought the target interval had elapsed.

As part of the consent process, participants were told that they would be producing intervals of 2, 4, or 6 minutes, and that that they should try to be as close as possible to the length of each interval (i.e. not to short; not too long). Each participant completed three productions per target interval of 2, 4, and 6 minutes, for a total of 9 productions. The order of the target intervals was randomized for each participant. While engaged in the interval production task, participants completed the self-paced dot counting task. For each trial of the dot counting task, participants were presented with a 6 x 6 grid filled with between 10 and 15 dots. Some dots were red, and some were blue. Participants were instructed to determine whether there were more blue or more red dots in the display and press the right arrow key to indicate that there were more red dots than blue dots and the left arrow key to indicate that there were more blue dots than red dots. In order to create easy and difficult conditions, the difference between the number of red and blue dots was varied. For participants in the easy condition, the difference between the number of blue dots and the number of red dots was 3-4 (e.g., 7 blue dots and 3 red dots). For participants in the difficult condition, the difference between the number of blue dots and number of red dots was 1-2 (e.g., 7 blue dots and 6 red dots). Participants were instructed to be as quick and accurate as possible on the dot counting task. When the participant believed the target interval had elapsed, they pressed the spacebar to terminate the interval production task and the dot counting task. If participants failed to make a response to the target interval production task after 8 minutes, the current target interval production task terminated. After the target interval production task was terminated, either automatically or by the participant, the participant then received feedback on their timing performance.

The feedback given to participants following each target interval production task appeared in the center of the computer monitor. For participants in the accurate feedback condition, the feedback informed them of the number of seconds by which they had under- or overestimated the interval (e.g., “You responded 30 seconds before the correct time”). For participants in the estimate error minimizing condition, the feedback informed them that they under- or overestimated by half of the number of seconds they actually under- or overestimated by (e.g., someone who responded 30 seconds early would receive feedback that said “You responded 15 seconds before the correct time”). For participants in the estimate error exaggerating condition, the feedback informed them that they under- or overestimated by 50% more than the number of seconds they actually under- or overestimated by (e.g., someone who responded 30 seconds early would receive feedback that said “You responded 45 seconds before the correct time”). After viewing the feedback display, participants began the next target interval production immediately. The experiment took approximately 1 hour to complete.

Results

Dot counting task. For each participant, the mean accuracy and response time on the dot counting task was calculated. The mean accuracy on the dot counting task was 97.04% (*SD* = 2.57) in the easy condition and 80.37% (*SD* = 10.63) in the difficult condition. The mean response time was 1123.76 ms (*SD* = 346.42 ms) in the easy condition and 3528.62 ms (*SD* = 277.58 ms) in the difficult condition.

Participants in the easy condition were significantly more accurate [*t* (51) = 7.78, *p* <.001] and significantly faster to respond [*t* (51) = -4.38, *p* <.001] than participants in the difficult condition, confirming that the difficulty manipulation was successful.

Interval production performance. The key dependent variables of interest were the length of productions in the easy and difficult conditions and the change in productions after feedback. Estimates were converted to a difference score, expressed as a percentage of the target interval, using the same formula used in Experiment 1. As described above, participants responded by key press to indicate when they thought the target interval had elapsed. In the event that participants did not respond and the program terminated the timing block, the program recorded a 0 for the participant's response time. One participant was excluded for having a terminated block.

Figure 5.1 shows the means of the median estimates for each interval and task difficulty. Participants underestimated all of the intervals. A repeated measures ANOVA with target interval (2, 4, and 6 minutes) as a within-subjects factor and task difficulty (easy vs. difficult) as a between-subjects factor revealed a significant main effect of target interval, $F(2, 104) = 16.09$, $p < .001$, $\eta^2 = .21$. The main effect of task difficulty was not significant, $F(1, 52) = .56$, $p > .10$, $\eta^2 = .01$, and the interaction between target interval and task difficulty was not significant, $F(2, 104) = .45$, $p > .10$, $\eta^2 = .02$.

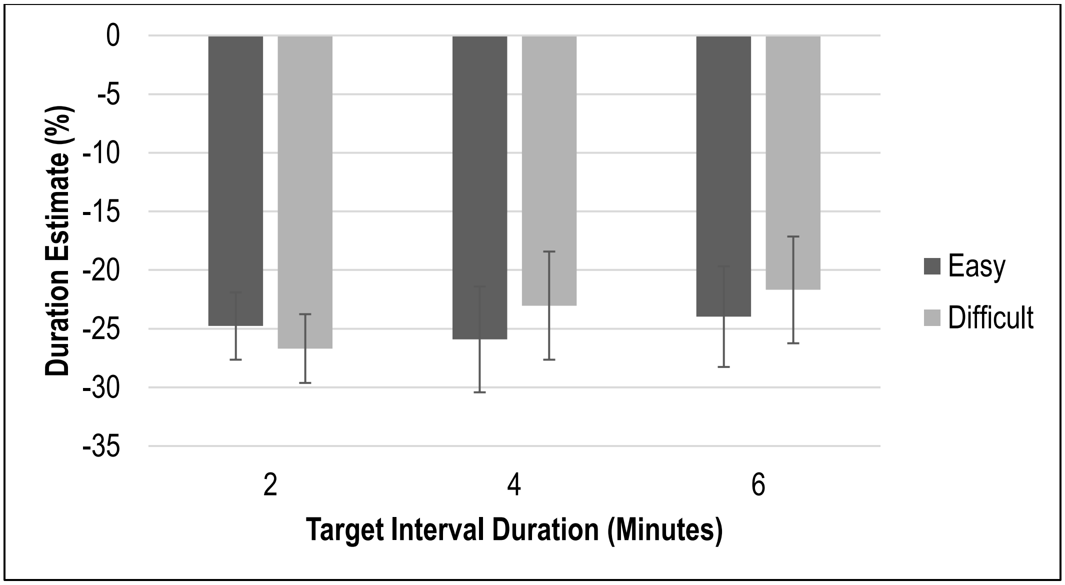

Figure 5.1. Duration estimates of the 2, 4, and 6 minute target intervals by difficulty of the ongoing task. Negative values are underestimates; positive values are overestimates. Error bars represent the standard error of the mean.

Figure 5.2 shows the average of estimates for the first and the last estimate each interval, collapsed across feedback type and ongoing task difficulty. Participants underestimated all of the intervals. A repeated measures ANOVA with target interval (2, 4, and 6 minutes) and trial (first and last) revealed an significant interaction between target interval and trial, $F(2, 104) = 7.07$, $p < .01$, $\eta^2 = .12$. Main effects for target interval [$F(2, 104) = 10.86$, $p < .01$, $\eta^2 = .17$] and trial [$F(1, 52) = 11.23$, $p < .01$, $\eta^2 = .18$] were also significant. Participants produced significantly longer estimates following feedback for the 2 minute and 4 minute intervals, but not the 6 minute intervals.

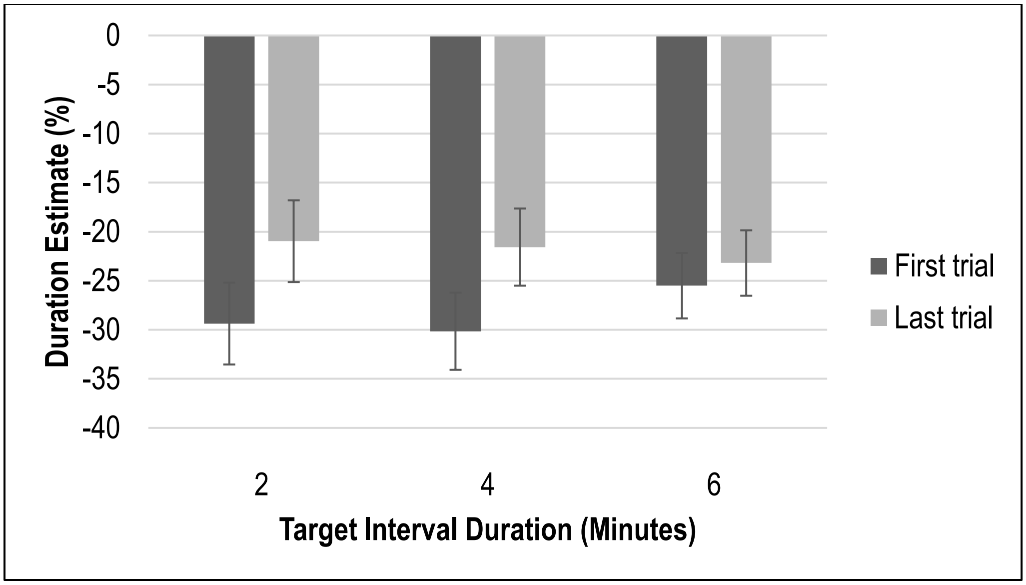

Figure 5.2. Duration estimates for the first and last trials of the 2, 4, and 6 minute target intervals. Negative values are underestimates; positive values are overestimates. Error bars represent the standard error of the mean

In Experiment 4, it was hypothesized that the type of feedback given (minimized estimate errors, exaggerated estimate errors, accurately represented estimate errors) would influence the magnitude of the difference between the first and last trial. In order to conduct these analyses, a change score was calculated by subtracting the absolute value of the percentage participants under- or over-estimated by on the last trial from the absolute value of the percentage participants under- or over-estimated by on the last trial. Doing so results in a positive value if participants' estimates were further from accurate on the first trial than the last, and a negative value if participants' estimates were closer to accurate on the first trial than the last. Neither ongoing task difficulty nor feedback type resulted in any significant main effects or a significant interaction.

Discussion

Experiment 4 was conducted in order to investigate the role of the amount of attention directed towards timing and feedback in timing everyday intervals. Three findings require further discussion: the lack of a difference in productions between the easy and difficult ongoing task conditions, the lack of a difference in participants' ability to use feedback in the easy and difficult ongoing task conditions, and the longer productions following feedback for the 2 and 4 minute intervals.

The lack of a difference in productions between the easy and difficult conditions of the ongoing task is inconsistent with the AGM. Analyses of the response time and accuracy data on the ongoing task confirmed that the difficult version of the task was indeed more difficult than the easy version of the task. Similar attentional load manipulations in shorter intervals have demonstrated that participants produce shorter intervals when engaged in the easy version of a task, and longer intervals when engaged in the difficult version of a task (e.g., Brown, 1985; Brown, 1997; Hicks et al., 1976; McClain, 1983). In Experiment 4, participants in the easy condition produced intervals that were not significantly different from those produced by participants in the difficult version of the task. Taken together, these findings suggest that, contrary to expectations, the amount of attention available for timing does not seem to influence the duration of productions in the range of minutes as it does with intervals in the seconds range. Such a finding is inconsistent with the AGM and may indicate that the AGM is not a good fit for explaining time estimation in this range.

Attention also had no effect on participants' ability to use feedback, as there was no main effect of ongoing task condition (easy vs. difficult) on participants' change from the first to the last trial. Participants in both ongoing task conditions produced longer estimates following feedback. This was contrary to the expectation that participants in the easy condition would have more attentional resources available for encoding feedback. This expectation drew from previous work demonstrating that encoding in episodic memory is better when more attentional resources are available (Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik, Naveh-Benjamin, Ishaik, & Anderson, 2000; Naveh-Benjamin & Guez, 2000; Naveh-Benjamin, Guez, & Marom, 2003). In contrast with these findings, frequency information (e.g., whether "book" or “journal” appears more frequently in English) seems to be encoded with little or no effect of additional attentional demands (Hasher & Zacks, 1979; Hasher & Zacks, 1984; Zacks, Hasher, & Sanft, 1982). It is possible that encoding in reference memory follows a similar pattern, in which case, we would not expect to see a difference in the effectiveness of feedback when paired with easy and difficult ongoing tasks.

Feedback resulted in longer productions on the last trial than on the first for the 2 and 4 minute conditions. This finding suggests that reference memory for these intervals can benefit from precise feedback. When participants receive feedback, they are able to refine their reference memory for the target interval and perform more accurately on a subsequent trial. This feedback effect was not present for the 6 minute target interval. It may be the case that there is no benefit from feedback because these intervals are too long for us to reflect on performance in the same way as we can for shorter intervals, or that more feedback trials are required to achieve a benefit when the task requires estimating longer intervals.

If feedback results in more accurate productions, and participants have prior experience estimating the duration of everyday tasks, why are they so inaccurate in timing everyday intervals? One possibility is that the feedback given in Experiment 4 differs significantly from the kind of feedback we receive in daily life. The feedback given in Experiment 4 referenced the direction (e.g., early or late) and magnitude of the error (e.g., 34 seconds). In daily life we may be more likely to receive feedback that tells us we were early or late without reference to a precise magnitude (e.g., “very early” vs. “58 seconds early”). Another possibility is that reference memory values are stored with contextual information, so feedback in one context is not generalized to a different context. We receive feedback on our accuracy in timing a 3 minute egg, but the same reference memory counts are not accessed when we are told to meet someone in 3 minutes. We also may lack experience timing these longer intervals precisely and receiving precise feedback about our performance. When cooking your morning eggs, it is rarely critical that you cook them for exactly 3 minutes and not a second over or under. A slightly overcooked fried egg is not, after all, inedible. This may result in a range of stored values for everyday intervals that are considered “good enough”. In addition, many factors outside of the precise time influence whether your egg is cooked to your liking – the temperature of the pan when you put the egg in, the temperature of the egg itself, how many eggs you put in the pan on a particular day. The combination of these factors, as well as the flexibility of what is considered properly cooked, may result in a lack of precise reference counts for everyday intervals.

This lack of stored reference counts may explain why participants are inaccurate, but it cannot explain why participants produce underestimates instead of overestimates. One possibility is that, when the reference memory for a particular interval is imprecise or contains a range of acceptable values, we are biased towards underestimation because of reinforcement and punishment. If you respond significantly too early to your frying egg, the worst that happens is that your egg is undercooked and requires a bit more time. If you respond significantly too late, your egg is burnt. Compared to responding late, responding early is rarely punished. You cook your egg a little longer, put your tea back in to steep, or wait a few seconds for a colleague to show up. If you respond too late, your food is inedible and your colleague is angry. Given that participants produced longer intervals after receiving corrective feedback, it is possible that the lack of precise reference counts is to blame for the consistent underestimates in Experiments 1-3.

Chapter 6 – General Discussion

Making my perfect cup of tea requires three minutes of steeping, and I have high confidence in my ability to keep track of the required steeping time without using any timing device. Similarly, I feel confident about my ability to return a phone call in 5 minutes without having to monitor the flow of time with external devices or alarms. These types of tasks require what is commonly referred to as prospective time production. For a prospective time production task, we are aware of the need to track time and to produce a response after a given interval has elapsed. Very little research has been done on the timing of everyday tasks in the range of minutes. Critically, it is not known whether models developed to explain timing of very short duration tasks can be extended to longer intervals. In order to determine what cognitive machinery might be required for prospective timing in the range of everyday tasks, a number of questions need to be answered. Is our confidence in our ability to estimate these intervals justified, that is, is it supported by evidence that people are able to estimate accurately the duration of everyday tasks? What factors might influence our ability to estimate the duration of intervals in this range? When we are inaccurate about estimating the duration of such tasks, do we tend to respond too early or too late? Can we learn from feedback about our timing accuracy? What cognitive mechanism(s) permit us to make accurate and reliable time estimates? This last question was the overall goal of my research.

Although a great deal of research has focused on the cognitive machinery required for prospective timing [i.e. for estimating when to stop an ongoing activity (take the tea bag out of the pot), or when to begin an intended activity (return a phone call)], most of this work has been on durations that are much shorter (e.g. milliseconds and seconds) than those that pertain to everyday task. Nevertheless, the models that do exist to account for prospective timing highlight the importance of memory processes and of attention in timing. The roles of these factors are highlighted in Chapter 1 of this document, in my detailed review of the most common existing timing models. The experiments reported in the preceding chapters were all conducted with the overarching goal of determining whether the most frequently cited of all models, the AGM and the memory storage models might be extended for explaining our ability to estimate the duration of everyday tasks. In this chapter, I will focus on general findings that emerged from this research; the specific findings from each study were discussed in the chapter where each study was reported.

Research Contributions

The overall goal of my dissertation was to investigate timing of everyday tasks, with a particular focus on determining whether the AGM and the memory storage models might explain our ability to estimate the duration of everyday tasks. In the sections that follow, I will briefly summarize the main findings from my research and examine their theoretical implications. In addition, where appropriate, I will suggest avenues for future research.

Underestimates are common. Over the course of all four experiments, participants on average underestimated intervals from 1-6 minutes, though specific combinations of ongoing tasks and the shorter intervals occasionally resulted in overestimates. In particular, participants overestimated the shortest interval in Experiments 2 and 3. Over the course of all four experiments, the size of underestimates varied, from approximately 2.5% to approximately 40.5%. In Experiments 2 and 3, underestimates became proportionally larger as the target interval lengthened. In Experiments 1 and 4, there was a Weber fraction-like relationship between target interval and estimate duration. Given the methodological shortcomings of Experiment 1 and the presence of feedback in Experiment 4, it seems likely that the results of Experiments 2 and 3 are more reliable, though follow up research is required to verify this conclusion.

One possibility is that underestimates are common because our reference memory for these intervals is imprecise. This may occur either because we do not have sufficient experience to have a reliable stored value or because the range of acceptable performance on everyday tasks results in a large range of possible pulse counts. If, however, our reference counts were imprecise we would expect greater variability in participants' estimates when they were unsure of the precise reference count, rather than a mean shift. In Experiments 2 and 3, underestimates became proportionally larger for longer intervals, but the variability did not increase. In Experiment 4 participants' estimates improved following feedback, but variability did not decrease significantly. It seems unlikely, therefore, that a lack of precise reference counts is solely to blame for underestimates.

A different possibility is that the underestimates revealed by the present research are attributable to demand characteristics, and more specifically, to participants' tendency to respond early in order to complete the experiment more quickly. If participants believed that they would be able to leave early if they completed the tasks more quickly, they might have responded before they believed a to-be-estimated interval had elapsed. Although participants were not told that they would be estimating a particular number of intervals (i.e., I implied that they would continue for the 1 or 2 hours they had signed up for), they may have assumed that if they responded earlier they would get to leave earlier. Many studies using the UBC Human Subjects Pool require participants to complete a pre-determined set of tasks, after which they are free to go.

This explanation for the underestimates produced by participants cannot be completely ruled out. However, if it were the case that participants were responding early in order to reduce their time in the lab, it would seem likely that they would underestimate the longest intervals by the greatest proportion, as this would lead to the greatest time savings and the earliest finishing time. In Experiments 1 and 4, participants underestimated all target intervals by approximately the same proportion, which conflicts with this explanation.

Future research is needed to fully rule out the possibility that underestimates are the result of participants' motivation to leave early. One possible avenue for research would be to motivate participants to achieve accurate estimates. For example, one could instruct participants that the experiment will continue and they will only be able to leave early if estimates are within 10% (or some other small margin of error) of the target interval. Such a requirement would encourage participants to produce intervals as accurately as possible in order to leave early. In this case, responding early would be counterproductive, as it would actually result in the experiment lengthening until the scheduled time had elapsed.

Relationship between target interval and estimate duration. In Experiments 2 and 3, underestimates became proportionally larger as the target interval lengthened, though this pattern was not observed in Experiments 1 and 4. Given the methodological limitations of Experiment 1 and the potential influence of feedback on the intervals in Experiment 4, it seems that the results of Experiments 2 and 3 might reveal the underlying pattern more accurately. Larger underestimates for longer intervals seem to indicate that the relationship between estimate duration and target interval is not a constant proportion – as the interval lengthens, our predictions become proportionally shorter, and since most productions are underestimates, less accurate.

It is possible that the proportionally larger underestimates for longer target intervals seen in Experiments 2 and 3 is due to inexperience in timing longer intervals, combined with a bias to respond early if we are unsure. We may have less reference memory information available for long duration tasks because we are more likely to subdivide them into shorter intervals. We may also tend to reach for external devices as we get less confident in our ability to time these longer intervals, resulting in less experience and less reference memory. The lack of reference memory for intervals in this range cannot fully explain why we might underestimate by a larger proportion for longer durations (e.g., 6 minutes) versus shorter durations (e.g., 2 minutes). As described above, if reference memory values were lacking, we would expect more variability rather than a change in the mean. Learning and reinforcement may provide an explanation for this mean shift. In many everyday tasks, such as frying an egg or leaving your office to meet a colleague, responding too late results in greater negative consequences (e.g., your egg is overcooked or burnt, your colleague is frustrated by waiting for you) than responding too early (e.g., you can keep cooking your egg, you have to wait a few seconds). It may be the case that responding too early is reinforced and responding too late is punished, resulting in a bias towards responding early. This explanation requires further research. One possible avenue for research would be to look at the relationship between individual differences such as risk averseness and conscientiousness and interval productions.

The size of the underestimate is not dependent on attention available for timing. The overarching goal of my dissertation was to identify the model or models that best fit timing in the range of everyday tasks. One of the models identified as a possible fit in Chapter 1 was the Attentional Gate Model. The AGM claims that when attentional resources are directed towards or available for timing, the attentional gate (see Figure 6.1) is widened, and pulses pass through the gate and into the accumulator more quickly than when attentional resources are unavailable or directed towards other tasks. The AGM predicts that productions will be shorter when more attention is allocated towards timing. The model assumes that attention is a divisible resource that can be consciously and deliberately allocated to timing or to other tasks.

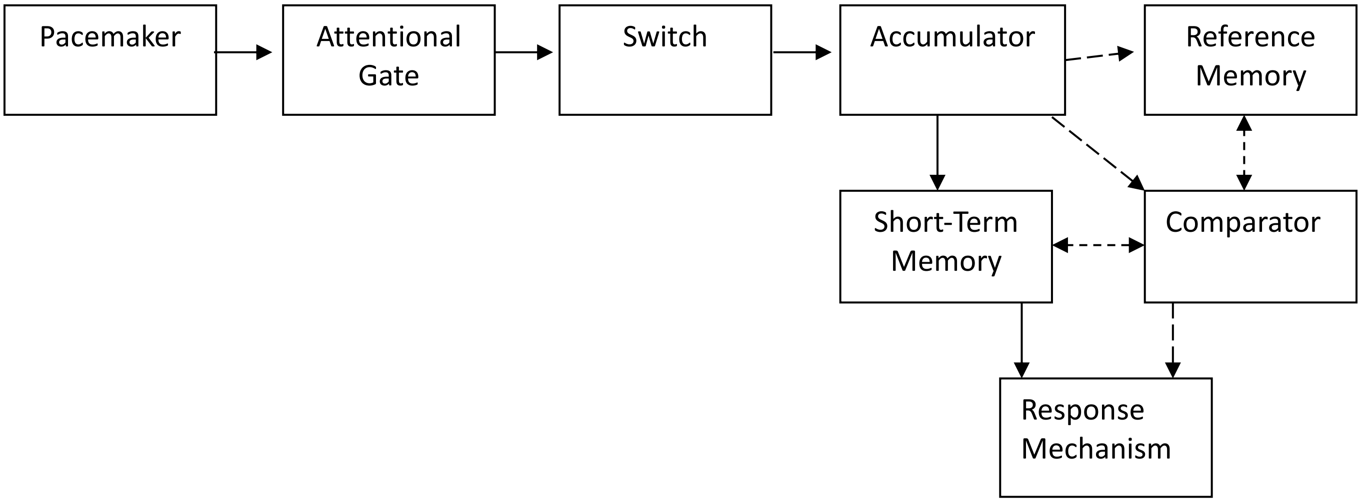

Figure 6.1. The attentional gate model (AGM)

Experiment 4 demonstrated that the AGM might not be a good fit for timing in the range of everyday tasks. In this experiment, I manipulated the difficulty or attentional resource demands of the tasks to be performed concurrently with making prospective time estimates. Although the difficulty manipulation was successful (i.e., participants in the easy condition responded more accurately and more quickly on the ongoing task than participants in the difficult condition), there was no difference in the duration of estimates produced by participants in the easy and difficult ongoing task conditions. In addition, in Experiments 1 and 3 there was no correlation between participants' performance on the ongoing task (response time and accuracy) and the duration of their estimates. Taken together, these results indicate that the AGM may not extend to timing of longer intervals.

The size of the underestimate is dependent on memory processing. Memory storage models were also identified as a possible fit for timing in this range. Memory storage models predict that productions will be shorter when the activity or information within an interval is more segmented (Block, 1978; Block & Reed, 1978; Ornstein, 1969), which may occur because of the number of stimuli or the organization of the stimuli. Manipulating the number of segments during an interval has typically been accomplished by imposing an organizing structure on the stimuli (e.g., Poynter, 1983) or by having participants switch back and forth between task types (e.g., Block and Reed, 1987). These manipulations, while effective for manipulating the segmentation of an interval, carry the possible confound of changing the attentional or memory demands of the ongoing task. In Experiments 2 and 3, a novel framing method for manipulating chunking within an interval was used. Memory storage models would predict that when the information within an interval is organized into fewer chunks, the estimate will be shorter as it will take longer for the participant to accumulate the target number of chunks. Participants were instructed to focus either on the present task or on a future task, but the underlying tasks were identical in both groups. In both experiments, participants in the future frame condition produced longer intervals than participants in the present frame condition. These findings are consistent with the idea that the present frame results in a better organizational structure than the future frame. Experiments 2 and 3 demonstrated that memory storage models may be a good fit for interval timing in the range of everyday tasks.

When we are focused on the present task, we view the events within each interval as hanging together, resulting in fewer chunks. When we are focused on a future task, the events within each interval as disconnected, resulting in more chunks. These chunks are then compared to the reference memory values for the target interval. If an interval has been stored as fewer chunks, it takes more time for the target number of chunks to be accumulated (see Figure 6.2). If an interval has been stored as more chunks, it takes less time for the target number of chunks to be accumulated. Taken together with the results of previous research in this range (Block, 1978; Block & Reed, 1978; Poynter, 1983), it appears that memory storage models are a better fit for timing in the range of everyday intervals than the AGM.

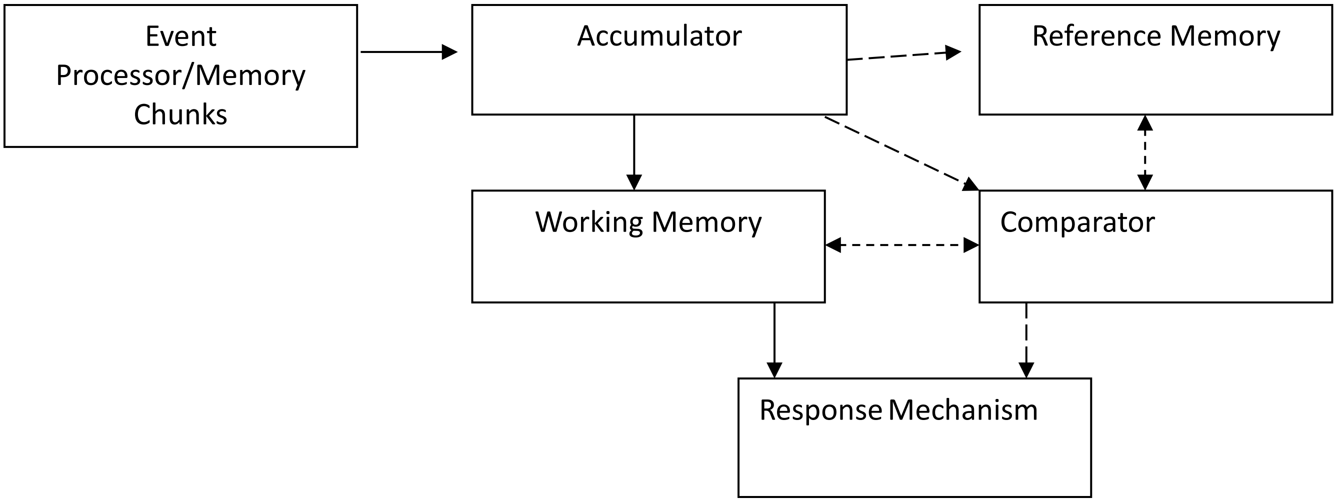

Figure 6.2. Possible organization of memory storage models

It is necessary to follow up with future research that verifies that chunking is indeed the key difference between the present and future focus conditions. One way to test this assumption would be by using an ongoing task that lends itself to a free recall test. If participants were given words from various categories to study during the timing task, we might expect that their free recall would be more highly organized in the present focus than the future focus condition, with items grouped into larger and more related chunks (Bousfield, 1953; Buschke, 1976). If input from participants was collected using E-Prime, rather than pencil and paper, it would also be possible to analyze their free recall responses for breaks in output time. This type of analysis has been used to investigate chunking in motor sequence learning (c.f. Sakai, Hikosaka, & Nakamura, 2004; van Vugt, Jabusch, & Altenmüller, 2012), but may prove useful in analyzing the organization and chunking of free recall in this context.

Feedback influences the duration of estimates. In Experiment 4, participants produced longer, more accurate intervals in the 2 and 4 minute condition following 2 feedback trials. The type of feedback (minimizing error, accurate, exaggerating error) and difficulty of the ongoing task (easy vs. difficult) did not influence participants' ability to improve following feedback. Why is it that two feedback trials are sufficient to result in improved performance? If only two feedback trials are required to improve, why are participants inaccurate in timing everyday durations, given that they have prior experience timing these intervals? This surprising finding requires follow up research.

How much training is required for perfect performance? In Experiment 4, participants' estimates improved by approximately 10% following feedback. One question that remains unanswered is whether this the relationship between number of feedback trials and the amount of improvement is linear or follows some other pattern. Is the reduction in error linear (e.g., every 2 feedback trials results in getting 10% closer to accuracy, in which case we would expect perfect performance after 6 feedback trials), or is the reduction in error proportional (e.g., error was reduced by 1/3, so after 2 feedback trials performance has shifted from -30% to -20%, after 2 more feedback trials performance shifts by 1/3 of 20%). At what point is the amount of error small enough that participants stop benefiting from precise feedback? The present study cannot answer these questions, and a future experiment with a larger number of feedback trials is required.

How long does the effect of feedback last? One possible explanation for participants' poor performance despite their experience with these intervals and the improvements seen following feedback is that the effect of feedback is not long-lasting. In Experiment 4, participants received feedback and immediately began the next trial. Memory storage models and the AGM make reference to long-term memory storage of reference values, which would suggest that feedback would have a long-term effect, but this would be dependent on the quality of participants' encoding of the feedback. If the feedback was not encoded well, it may not be stored in long-term memory and accessed when timing in the future. Future experiments investigating how long the effect of feedback lasts are required to address this possibility.

Why are improvements following feedback not seen for the 6 minute intervals? Participants in Experiment 4 showed improvement following feedback on the 2 and 4 minute intervals, but not the 6 minute intervals. Given the unexpected finding that only 2 feedback trials were required to improve performance for the shorter intervals, it may simply be the case that more feedback trials are required for the longer 6 minute interval. It may be the case that there is more interference between the beginning of the interval and receiving feedback for the 6 minute interval than for the shorter intervals, which may result in participants being less able to apply the feedback for the longer interval.

Memory storage models provide a possible explanation for the seemingly conflicting findings that feedback improves performance in as few as two trials, but large underestimates are common. It may be the case that reference memory involves both temporal information and contextual information. If this is the case, improvements following feedback may not transfer to a new context, which may explain why participants are able to improve from feedback but are generally poor at estimating intervals in this range. The AGM makes no claim about reference memory counts including contextual information, suggesting that they are stored in long-term semantic memory. According to the AGM, pulses are produced by the pacemaker at regular intervals, and reference memory counts are based on past experience. Given that pulses do not change in size or rate, there is no reason to expect that counts stored in reference memory would include contextual information.

In memory storage models, episodic memory is the source of the to-be-counted chunks. Given that reference memory is the result of experience in timing longer durations, and episodic memory is the source of the chunks used in timing, contextual information may be stored along with the number of chunks in reference memory. Feedback could then result in improvement that would not transfer to a new context. This brings up several avenues for future research – does improvement after feedback transfer to a new context? Do people with semantic memory impairments show improvement? What about people with episodic memory impairments?

Limitations

There are several limitations in Experiments 1-4. All experiments were conducted using students from the UBC human subjects pool and may not generalize to a broader population. Compared to an age-matched non-student population, students may engage in more frequent timing in this range. They are required to estimate how much more time they have when approaching the end of an exam, how long it will take to get to their next class, and whether they have enough time to complete an assignment, This experience may result in different timing behaviour than seen outside of a student population. Most of the participants in Experiments 1-4 were between 18-22 years old, with a few exceptions. This may result in different timing behaviour than seen in an older population. Compared to older groups of adults, participants in this study likely have less developed frontal lobes (Sowell et al., 1999), which may result in poorer executive functions such as planning and decision making. It is not known to what extent these factors might influence timing behaviour. They may also be more likely to use external devices for timing short duration tasks than older adults, as they are used to having easy access to smart phones with timers. There is often no need to consciously time everyday tasks that require precise responses (e.g., timing a 3 minute egg) because setting a timer is so easy.

All experiments were also conducted in a lab setting, using artificial tasks that bear little similarity to real-world everyday timing tasks. In real-world tasks, there are often many additional external cues to aid us in performing the task, such as the colour or smell of your steeping tea. There are also more potential distractions from the primary timing task in the world outside of the lab. If it is the case that reference memory counts involve contextual information, participants' large underestimates may have been influenced by a lack of familiarity with the tasks being used. Practice trials of the ongoing task in each experiment were employed to increase task familiarity, but participants did not receive practice trials of the primary timing tasks.

The artificial nature of the timing tasks may have resulted in low investment in performance. In Experiments 1-3, there were no consequences for poor performance. In Experiment 4, participants received feedback, but it is unlikely that they were highly invested in performing accurately. It is possible that participants may perform more accurately if there are consequences to poor performance. In many everyday timing tasks, there are social consequences for poor performance – if you show up late to meet a colleague, they are frustrated. If you arrive too early to pick somebody up, they are stressed and anxious. Future research should investigate whether the addition of consequences for poor performance results in more accurate estimates.

Conclusions

The overall goal of my dissertation research was to identify the cognitive mechanism that underlies our ability to time intervals in the range of everyday tasks. We complete many tasks every day that require us to respond when a given interval has elapsed, such as making a cup of tea or meeting someone in 5 minutes. Previously, very little work focused on investigating timing in this range. Overall, my dissertation research shows that people tend to underestimate these intervals, and that memory storage models might provide a better explanation than the AGM for performance on timing tasks in this range.

The results of the experiments reported in the preceding chapters indicate that memory storage models (Block, 1978; Block & Reed, 1978; Ornstein, 1969) are a better fit for prospective timing of everyday intervals than the AGM (Zakay & Block, 1996, 1998). When engaged in timing longer durations, participants use the organization of chunks in episodic memory in order to determine how much time has elapsed and when to respond that the interval is over. Further research is required in order to determine whether performance on lab-based prospective timing tasks is similar to performance on timing tasks in the context of everyday activities, where we have more experience (and are potentially better able to use context in judging time), may be more motivated to respond accurately, and receive more feedback.

References

Abel, S.M. (1972). Discrimination of temporal gaps. *The Journal of the Acoustical Society of America, 52(2),* 519-524. doi: 10.1121/1.1913139

Alards-Tomalin, D., Leboe-McGowan, J.P., Shaw, J.D.M., & Leboe-McGowan, L.C., (2014). The effects of numerical magnitude, size, and color saturation on perceived interval duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition 40(2),* 555-566. doi: 10.1037/a0035031

Allan, L.G., (1979). The perception of time. *Perception & Psychophysics 26(5),* 340-354. doi: 10.3758/BF03204158

Allan, L.G., & Kristofferson, A.B. (1974). Judgments about the duration of brief stimuli. *Perception & Psychophysics, 15(3),* 434-440. doi: 10.3758/BF03199282

Allan, L.G., Kristofferson, A.B., & Wiens, E.W. (1971). Duration discrimination of brief light flashes. *Perception & Psychophysics, 9(3b),* 327-334. doi: 10.3758/BF03212659

Allman, M.J., Teki, S., Griffiths, T.D., & Meck, W.H., (2014). Properties of the internal clock: First- and second-order principles of subjective time. *Annual Review of Psychology, 2014, 65,* 743-771. doi: 10.1146/annurev-psych-010213-115117

Allport, D.A. (1968). Phenomenal simultaneity and the perceptual moment hypothesis. *British Journal of Psychology, 59(4),* 395-406. doi: 10.1111/j.2044-8295.1968.tb01154.x

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113(4),* 518-540. doi: 10.1037/0096-3445.113.4.518

Bilgin, B., & LeBoeuf, R.A. (2010). Looming losses in future time perception. *Journal of Marketing Research, 47,* 520-530. doi: 10.1509/jmkr.47.3.520

Block, R.A. (1978). Remembered duration: Effects of event and sequence complexity. *Memory & Cognition, 6(3),* 320-326. doi:10.3758/BF03197462

Block, R.A. (1990). Models of psychological time. In R.A. Block (Ed.), *Cognitive Models of Psychological Time* (pp. 1-35). New Jersey: Lawrence Erlbaum Associates, Inc..

Block, R.A., (1992). Prospective and retrospective duration judgment: The role of information processing and memory. In F. Macar, V. Pouthas, & W.J. Friedman (Eds.), *Time, Action, and Cognition* (pp.141-152). Netherlands: Kluwer Academic Publishers.

Block, R.A. & Gruber, R.P., (2014). Time perception, attention, and memory: A selective review. *Acta Psychologica, 149,* 129-133. doi: 10.1016/j.actpsy.2013.11.003

Block, R.A., Hancock, P.A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica, 134,* 330-343. doi:10.1016/j.actpsy.2010.03.006

Block, R.A., & Reed, M.A., (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 4(6),* 656-665. doi: 10.1037/0278-7393.4.6.656

Block, R.A., & Zakay, D. (1997). Prospective and retrospective judgments: A meta-analytic review. *Psychonomic Bulletin & Review, 4,* 184-197. doi: 10.3758/BF03209393

Boles, D.B., & Law, M.B. (1998). A simultaneous task comparison of differentiated and undifferentiated hemispheric resource theories. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 204-215. doi: 10.1037/0096-1523.24.1.204

Bousfield, W.A., (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology, 49*, 229-240. doi: 10.1080/00221309.1953.9710088

Brown, S.W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics, 38,* 115-124. doi: 10.3758/BF03198848

Brown, S.W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception & Psychophysics, 59,* 1118-1140. doi: 10.3758/BF03205526

Brown, S.W. (2006), Timing and executive function: Bidirectional interference between concurrent temporal production and randomization tasks. *Memory & Cognition, 34(7),* 1464-1471. doi: 10.3758/BF03195911

Brown, S.W. (2008). The attenuation effect in timing: Conteracting dual-task interference with time-judgment skill training. *Perception, 37,* 712-724. doi: 10.1068/p5698

Brown, S.W., Collier, S.A., & Night, J.C., (2013). Timing and executive resources: Dual-task interference patterns between temporal production and shifting, updating, and inhibition tasks. *Journal of Experimental Psychology: Human Perception and Performance, 39(4),* 947-963. doi: 10.1037/a0030484Bueno Martinez, B., (1994). The role of cognitive changes in immediate and remote prospective time estimations. *Acta Psychologica, 85,* 99-121. doi: 10.1016/0001-6918(94)90027-2

Buschke, H., (1976). Learning is organized by chunking. *Journal of Verbal Learning and Verbal Behavior*, *15,* 313-324. doi: 10.1016/0022-5371(76)90028-1

Church, R.M., & Broadbent, H.A., (1990). Alternative representations of time, number, and rate. *Cognition, 37,* 55-81. doi: 10.1016/0010-0277(90)90018-F

Church, R.M., & Gibbon, J.G., (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behaviour Processes, 8(2),* 165-186. doi: 10.1037/0097-7403.8.2.165

Craig, A.D., (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews: Neuroscience, 3,* 659-666. doi:10.1038/nrn894

Craig, A.D., (2003). Interoception: The sense of the physiological condition of the body. *Current Opinion in Neurobiology, 13,* 500-505. doi: 10.1016/S0959-4388(03)00090-4

Craig, A.D., (2009). Emotional moments across time: A possible neural basis for time perception in the anterior insula. *Philosophical Transactions of The Royal Society B, 364,* 1933-1942. doi:10.1098/rstb.2009.0008

Craik, F.I.M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), Human memory and cognitive capabilities: Mechanisms and performances (pp. 409-422). Amsterdam: Elsevier

Craik, F.I.M., Naveh-Benjamin, M., Ishaik, G., & Anderson, N.D., (2000). Divided attention during encoding and retrieval: Differential control effects? *Journal of Experimental Psychology; Learning, Memory, and Cognition, 26(6),* 1744-1749. doi: 10.1037/0278-7393.26.6.1744

Creelman, C.D., (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America, 34(5),* 582-593. doi: 10.1121/1.1918172

Divenyi, P.L., & Danner, W.F., (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Perception & Psychophysics, 21(2),* 125-142. doi: 10.3758/BF03198716

Druckman, J.N., (2001). Evaluating framing effects. *Journal of Economic Psychology, 22(1),* 91-101. doi: 10.1016/S0167-4870(00)00032-5

Dunegan, K., (1993). Framing, cognitive modes, and image theory: Toward and understanding of a glass half full. *Journal of Applied Psychology, 78(3),* 491-503. doi: 10.1037/0021-9010.78.3.491

Esposito, M.J., Natale, V., Martoni, M., Occhionero, M., Fabbri, M., & Cicogna, P., (2007). Prospective time estimation over a night without sleep. *Biological Rhythm Research, 38,* 443-450. doi: 10.1080/09291010601068776

Gibbon, J. (1991). Origins of scalar timing. *Learning and Motivation, 22,* 3-38. doi: 10.1016/0023-9690(91)90015-Z

Gibbon, J., & Church, R.M., (1990). Representation of time. *Cognition, 37,* 23-54. doi: 10.1016/0010-0277(90)90017-E

Gibbon, J., Church, R.M., & Meck, W.H., (1984). Scalar timing in memory. In J. Gibbon & L. G. Allan (Eds.), Annals of the New York Academy of Sciences: Timing and time perception. New York: New York Academy of Sciences.

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

Grondin, S. & Macar, F. (1992). Dividing attention between temporal and nontemporal tasks: A performance operating characteristic –POC- analysis. In F. Macar, V. Pouthas, & W.J. Friedman (Eds.), *Time, Action, and Cognition* (pp.119-128). Netherlands: Kluwer Academic Publishers.

Hasher, L. & Zacks, R.T., (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108(3),* 356-383. doi: 10.1037/0096-3445.108.3.356

Hasher, L. & Zacks, R.T., (1984). Automatic processing of fundamental information: The case of frequency occurrence. *American Psychologist, 39(12),* 1372-1388. doi: 10.1037/0003-066X.39.12.1372

Hemmes, N.S., Brown, B.L., Kladopoulos, C.N. (2004). Time perception with and without a concurrent nontemporal task. *Perception & Psychophysics, 66,* 328-341. doi: 10.1016/0010-0277(90)90017-E

Hicks, R.E., Miller, G.W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *The American Journal of Psychology, 89,* 719-730. Recovered from http://www.jstor.org/pss/1421469

Hoagland, H., (1933). The physiological control of judgments of duration: Evidence for a chemical clock. *Journal of General Psychology, 9,* 267-287. doi: 10.1080/00221309.1933.9920937

Jones, B., & Huang, Y.L., (1982). Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effects. *Psychological Bulletin, 91(1),* 128-142. doi: 10.1037/0033-2909.91.1.128

Kahneman, D. *Attention and effort*. New York: Prentice-Hall, 1973.

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

Kinchla, J. (1972). Duration discrimination of acoustically defined intervals in the 1- to 8-sec range. *Perception & Psychophysics, 12(3),* 318-320. doi: 10.3758/BF03207213

Kladopoulos, C.N., Hemmes, N.S., & Brown, B.L. (2004). Prospective timing under dual-task paradigms: Attentional and contextual-change mechanisms. *Behavioural Processes, 67,* 221-233. doi: 10.1016/j.beproc.2003.12.004

LaBerge, D., (1983). Spatial extent of attention to letters and words. *Journal of Experimental Psychology: Human Perception and Performance, 9(3),* 371-379. doi: 10.1037/0096-1523.9.3.371

Lewis, P.A., & Miall, R.C., (2009). The precision of temporal judgment: Milliseconds, many minutes, and beyond. *Philosophical Transactions of the Royal Society B, 364,* 1897-1905. doi:10.1098/rstb.2009.0020

Liu, Y., Li, S., & Sun, Y., (2014). Unpacking a time interval lengthens its perceived temporal distance. *Frontiers in Psychology, 5,* doi: 10.3389/fpsyg.2014.01345

Macar, F., Grondin, S., & Casini, L. (1994). Controlled attention sharing influences time estimation. *Memory & Cognition, 22,* 673-686. doi: 10.3758/BF03209252

McClain, L. (1983). Interval estimation: Effect of processing demands on prospective and retrospective reports. *Perception & Psychophysics, 34,* 185-189. doi: 10.3758/BF03211347

Naveh-Benjamin, M., & Guez, J., (2000). Effects of divided attention on encoding and retrieval processes: Assessment of attentional costs and a componential analysis. *Journal of Experimental Psychology; Learning, Memory, and Cognition, 26(6),* 1461-1482. doi: 10.1037/0278-7393.26.6.1461

Naveh-Benjamin, M., Guez, J., & Marom, M., (2003). The effects of divided attention at encoding on item and associative memory. *Memory and Cognition, 31(7),* 1021-1035. doi: 10.3758/BF03196123

Ornstein, R.E., (1969). *On the experience of time.* Middlesex, England: Penguin books.

Philbin, T., & Seidenstadt, R.M., (1983). Feedback and time perception. *Perceptual and Motor Skills, 57,* 308. doi: 10.2466/pms.1983.57.1.308

Poynter, W.D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition, 11(1),* 77-82. doi: 10.3758/BF03197664

Ryan, L.J., & Fritz, M.S., (2007). Erroneous knowledge of results affects decision and memory processes on timing tasks. *Journal of Experimental Psychology: Human Perception and Performance, 33(6),* 1468-1482. doi: 10.1037/0096-1523.33.6.1468

Sakai, K., Hikosaka, O., & Nakamura, K., (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Science, 8(12),* 547-553. doi: 10.1016/j.tics.2004.10.005

Shouval, H.Z., Shuler, M.G.H., Agarwal, A., & Gavornik, J.P., (2014). What does scalar timing tell us about neural dynamics? *Frontiers in Human Neuroscience, 8(19),* 1-6. doi: 10.3389/fnhum.**2014**.00438

Shulman, H.G., (1971). Similarity effects in short-term memory. *Psychological Bulletin, 75(6), 399-415.* doi: 10.1037/h0031257

Stroud, J.M. (1967). The fine structure of psychological time. *Annals of the New York Academy of Sciences, 138(2)*, 623-631. doi: 10.1111/j.1749-6632.1967.tb55012.x

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs, 77.* doi*:* 10.1037/h0093864

Treisman, M. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19,* 705-743. doi: 10.1068/p190705

Treisman, M., (1993). On the structure of the temporary sensory system. *Psychologica Belgica, 33(2),* 271-283.

Treisman, M., & Brogan, D., (1992). Time perception and the internal clock: Effects of visual flicker on the temporal oscillator. *European Journal of Cognitive Psychology, 4(1),* 41-70. doi: 10.1080/09541449208406242

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,* 705-743. doi: 10.1068/p190705

Tversky, A. & Kahneman, D., (1981). The framing of decisions and the psychology of choice. *Science, 211(4481), 453-458.* doi: 10.1126/science.7455683

van Vugt, F.T., Jabusch, H.C., & Altenmüller, E., (2012). Fingers phrase music differently: Trial-to-trial variability in piano scale playing and auditory perception reveal motor chunking. *Frontiers in Psychology, 3,* 1-9. doi: 10.3389/fpsyg.2012.00495

Wickelgren, W.A., (1965). Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology, 70(1),* 102-108. doi: 10.1037/h0022015

Zacks, R.T., Hasher, L., & Sanft, H. (1982). Automatic encoding of event frequency: Further findings. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 8(2), 105-116. doi: 10.1037/0278-7393.8.2.106

Zakay, D. (1990). The evasive art of subjective time measurement: Some methodological dilemmas. In R.A. Block (Ed.), *Cognitive Models of Psychological Time* (pp. 59-84). New Jersey: Lawrence Erlbaum Associates, Inc..

Zakay, D., & Block, R. A. (1996). The role of attention in time estimation processes. In M. Pastor & J. Artieda (Eds.), *Time, Internal Clocks, and Movement* (pp. 143-164). New York: Elsevier Science.

Zakay, D., & Block, R.A., (1997). Temporal cognition. *Current Directions in Psychological Science, 6(1),* 12-16. doi: 10.1111/1467-8721.ep11512604

Zakay, D., & Block, R., (1998). New perspective on prospective time estimation. In V. De Keyser, G. d'Ydewalle, & A. Vandierendonck (Eds.), *Time and the Dynamic Control of Behaviour* (pp. 129-142). Bern: Hogrefe & Huber Publishers.

Zakay, D., Nitzan, D., & Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. *Perception & Psychophysics, 34,* 451-456. doi: 10.3758/BF03203060

Zakay, D., & Tsal, Y. (1989). Awareness of attention allocation and time estimation accuracy. *Bulletin of the Psychonomic Society, 27,* 209-210. doi: 10.3758/BF03334586