# Prospective and Retrospective Time Estimation:
# Investigating the Effects of Task Duration and Cognitive Load

By
Jesika Walker

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Arts
in
Psychology

Carleton University
Ottawa, Canada

## Abstract

This experiment investigated participants' ability to keep track of time during a visual and memory search task where task difficulty and duration were manipulated. Two hundred and ninety-two participants performed the task for eight or 58 minutes. Participants in the prospective time judgment condition were forewarned of an impending time estimate whereas those in the retrospective condition were not. Cognitive load was manipulated and assessed by assigning participants to either a consistent or a varied mapping condition. The results revealed overestimation and higher variability of estimates in the prospective condition compared to the retrospective one in the eight-minute task only. Moreover, participants significantly overestimated the duration of the eight-minute task and underestimated the 58-minute task. Finally, cognitive load had no effect on participants' time estimates. Thus, the well-known cross-over interaction between cognitive load and estimation paradigm (Block et al. 2010) does not seem to extend to longer durations.


*Keywords*: Time estimation, task duration, cognitive load, visual and memory search task

# Acknowledgements

I would like to express my gratitude to everyone who has supported me throughout this research. I am especially thankful for the guidance provided by my advisor, Dr. Guy Lacoix. I appreciate having had the opportunity to pursue my topic of interest and to develop a foundation for future time estimation research in the lab. Your comments on this thesis have helped me grow as a writer and a researcher.

I am sincerely grateful to everyone in the lab for their interest and advice on this project. Thank you very much to Mohammed Aswad and Calvin Findlay for helping run participants through the experiment and to Billal Ghadie for providing me with computer programming guidance. Your assistance made this experiment a success.

I would also like to thank my committee members Dr. Johanna Peetz and Dr. Olessia Jouravlev. I appreciate the time you took to read through this thesis and to provide me with constructive feedback. Thank you to Dr. Peetz for providing direction on the prospectus version of this document. Your feedback strengthened the quality of this experiment.

Finally and importantly, I am grateful for all the love and encouragement provided by my partner Arthur Braaten, our cats (Charlie and Baxter), and my closest friend Lori Watanabe. You have all been imperative to the completion of this thesis and of this Master's degree more generally.

List of Content

# List of Tables

# List of Figures

# List of Appendices

# Prospective and Retrospective Time Estimation: Investigating the Effects of Task Duration and Cognitive Load

A key feature of time perception is the awareness of time passing (Tobin, Bisson, & Grondin, 2010). There exists a long history of speculation about this awareness, specifically because time perception is crucial for the functioning of human beings (Roeckelein, 2000). To regulate daily activities such as cooking, driving, reading, or writing a thesis, people rely on an internal representation of time. The heightened pace of contemporary employment and family demands, however, have made people's ability to coordinate their activities in time extremely complex. Consequently, the relationship between time perception and task demands is an increasingly important area of interest in cognition (Block, Hancock, & Zakay, 2010).

To investigate human time perception, objective measurements of time (i.e., clock time) are compared with participants' subjective estimates. This is often accomplished by instructing participants to keep track of time while engaging in a concurrent task. Participants then provide an estimate on how long the task took to complete. When this method is employed, time estimates are called *prospective* (Hicks, Miller, & Kinsbourne, 1976). These instructions to pay attention to time result in the intentional encoding of temporal information and enhance estimate accuracy (Block & Zakay, 1997). In real life, however, people do not necessarily attend to the passage of time while engaging in their activities. Therefore, some have argued that a more realistic experimental setting involves one where participants are unaware that an impending time estimate will eventually be required (Tobin et al., 2010). When this method is employed, time estimates are called *retrospective* and they provide insight into the incidental encoding of temporal information because they exclude attention to time (Brown, 1985).

Direct comparisons between prospective and retrospective time estimates are quite informative because they help reveal the role of intentional and incidental encoding in human time perception. Yet, few such comparisons exist (Tobin et al., 2010) and they suggest that prospective time estimates are more accurate (closer to real time) and less variable than retrospective estimates (Block & Zakay, 1997). These differences are thought to result from separate underlying cognitive processes and some researchers have argued that different models for prospective and retrospective timing are needed (Zakay & Block, 1997). Nevertheless, there is a search for a unifying model in the literature (French, Addyman, Mareschal, & Thomas, 2014).

Moreover, most research on time estimation narrowly focuses on durations ranging from seconds up to two-minutes (e.g., Predebon, 1996; Avani-Babad & Ritov, 2003; Kurtz & Strube, 2003; Boltz, 2005). In fact, an eight-minute condition used in Tobin and Grondin (2009) is considered a long duration. Experiments that directly compare prospective and retrospective paradigms at long duration are extremely rare (Tobin & Grondin, 2009; and Tobin et al., 2010) and only one direct comparison exists for a task of one hour (Bakan, 1955). Whether prospective estimates remain more accurate and less variable at these very long durations requires further investigation.

Finally, not all activities are created equal. Primary tasks in this area of research differ greatly from one experiment to the next. For example, participants in past research have provided estimates of time spent listening to a piece of music (Brown & Stubbs, 1992), performing a Stroop task (Zakay & Fallach, 1984; Predebon, 1999), rehearsing a list of items (Miller, Hicks, & Willette, 1978), or even watching a pot of water boil (Block, George, & Reed, 1980). These tasks likely differ in the amount of mental effort required. This effort is typically conceptualized as

cognitive load and has been shown to interfere with people's perception of time (Block et al., 2010). In fact, it is understood that non-temporal aspects of a task sometimes compete for cognitive resources required in keeping track of time (Duzcu & Hohenberger, 2014).

Thus, the goal of the present thesis was to consider the impact of the time estimate paradigms (prospective vs. retrospective), task duration, and cognitive load on people's perception of time. Participants performed a visual and memory search task while either prospectively aware of an impending time estimate or not. This task was either easy (i.e., low cognitive load) or difficult (i.e., high cognitive load) and participants performed the task for either long or very long intervals of eight or 58 minutes, respectively.

To situate this present study within the broader time estimation literature, I will first provide background on prospective and retrospective time estimation paradigms, explain how these methods impact the accuracy and variability of time estimates, and present the contemporary cognitive models proposed to account their impact. Then, I will highlight the importance of considering the length of a task and its cognitive load demands when people are instructed to estimate its duration. Presumably, duration and cognitive load influence prospective and retrospective time estimates differently. Finally, I will outline how and why the present experiment investigated this hypothesis using a visual and memory search task.

## The Two Time Estimation Paradigms

Many factors have been shown to influence human time perception. Zakay and Block (1997) highlighted how complex the picture is and navigating this area requires careful consideration of complex interactions among numerous variables. At the very least, the most important factor to consider is whether a time estimate is made prospectively or retrospectively.

Prospective and retrospective time estimation paradigms differ in the amount of awareness participants have about an impending time estimate. Hicks et al. (1976) introduced this distinction when they asked two groups of participants to sort playing cards for 42 seconds. Only one group was told that they would also report on time elapsed at the end of the experiment. These two groups were further divided into different conditions based on how much information was required to sort the cards. Participants were told to stack all the cards into a single pile, two piles according to colour, or four piles according to each card's suit. After completing the task, participants were asked how much time had elapsed. When prospective estimates were made, there was a significant negative linear relationship between the amount of information to be processed on the cards and the accuracy of their time estimates. The prospective group with the least amount of information to process estimated 52.92 seconds whereas the one with most information estimated 31.00 seconds. However, there was no linear relationship between the card sorting conditions and the retrospective estimates. This study highlighted how experimental manipulations might influence estimates differently depending on whether people are in a prospective or retrospective paradigm.

Retrospective estimates have received much less experimental attention, yet some researchers argue that they are more ecologically valid (Brown & Stubbs, 1988) because people do complete a myriad of activities without fully attending the passage of time. Retrospective estimates thus offer insight into how people experience time when they are inattentive to it. In other words, participants in this paradigm process temporal information incidentally (Brown, 1985). In this way, these time estimates are useful for determining the role of attention in time perception. Nonetheless, retrospective estimates entail an important methodological challenge. As soon as they are requested, participants realize that keeping track of time is an important

experimental consideration (Tobin et al., 2010). Consequently, the retrospective paradigm becomes prospective after a single estimate, typically restricting this paradigm to between-participant designs.

Researchers have attempted to overcome the challenge of collecting multiple retrospective estimates by using a series of distinct tasks (e.g., Boltz, 2005; Grondin & Plourde, 2007; Tobin & Grondin, 2009). Participants perform all the tasks in sequence and then report on how long each one took to complete at the end of the entire experiment. By dividing the experiment into multiple units, all estimates made at the end of the experiment qualify as being made retrospectively. Although this strategy has been used in previous research, it is limited by an important confound called the time-order effect. That is, the first two tasks of equal duration are judged to be longer than subsequent tasks (Block & Zakay, 1997). This time-order effect is problematic because these estimates do not reflect memory for the length of individual intervals. Instead, they are constructed using memory for specific attributes of the entire sequence of intervals (Block, 1978). These estimates thus reflect the relative amount of time that one task took to complete in the context of all the others. Importantly, estimates can change depending on occurred task's position in the sequence.

It is therefore preferable to avoid time-order effects when investigating individual retrospective time estimates. To do this, researchers request a single time estimate per participant and recruit a large sample (Brown & Stubbs, 1988). Unfortunately, the logistical challenges posed by such experiments have led researchers to opt almost solely for prospective measures (Block & Zakay, 1997). Unlike retrospective time estimation, the already intentional nature of a prospective estimate permits the collection of multiple reports. Future time estimates remain part of a prospective paradigm regardless of how many estimates have been previously made. These

estimates can also be requested immediately after each task avoiding a time-order effect. As a result, it is possible to acquire more measurements with fewer participants using this design. Most of what is known about time perception thus involves the intentional encoding of temporal information.

The restrictive nature of retrospective estimates has also limited the number of direct comparisons made between retrospective and prospective estimation paradigms. Tobin et al. (2010) argue that direct comparisons between paradigms are only valid if estimates are made in the same experimental setting. This argument arises from studies outlining variables that influence time estimation other than prior knowledge of the time keeping task. For example, research suggests that "filled" and "empty" durations have differential effects. Filled durations contain content (e.g., short sounds) and this creates the perception that the interval is divided. When intervals are perceived as divided, participants make longer time estimates (Roeckelein, 2000). Consequently, tasks must contain the same number of breaks and the same organization of content to be fully comparable. Furthermore, to compare any time estimation task, the studies must have collected the estimates using the same recording method. How the participant produces a time estimate requires unique cognitive processes that can in turn shape the results of a study (Zakay, 1993; Beaudouin, Vanneste, Isingrini, & Pouthas, 2006).

**Recording a time estimate.** Three methods are typically employed to record participants' time perception: verbal estimation, production, and reproduction. Verbal estimates require the participant to translate an experienced duration into measurable units typically given in seconds (Zakay & Block, 1997). Production involves a similar translation, except that participants produce a verbally specified length of time (e.g., "please press the button after 20 seconds has passed"). By its very nature, production typically constitutes a prospective time

estimate. Finally, reproduction does not require a participant to translate any verbal information into conventional time units. Instead, participants experience a duration and then attempt to recreate it (Block, Zakay, & Hancock, 1998).

Compared to production and reproduction, verbal estimates tend to be more variable. This method relies on the assumption that participants can reliably translate experience into conventional units of time. Thus, variability likely stems from individual differences in self-report (Block et al., 1998). Participants also tend to underestimate durations when they are asked to report verbally (Hornstein & Rotter, 1969) and these estimates increase with poorer access to semantic memory (Ogden, Wearden, & Montgomery, 2014).

Despite their variability, verbal estimates are versatile in that they can be applied to lengthy tasks. In comparison, reproduction tasks quickly become unfeasible for time intervals spanning more than several minutes and the production method is restrictive in that participants can only do so prospectively. Consequently, most direct comparisons between paradigms have used a verbal estimation method (e.g., Avni-Babad & Ritov, 2003; Kurtz & Strube, 2003; Hicks et al., 1976; Tobin et al., 2010) as did the present experiment.

**Differences between paradigms.** Direct comparisons are essential for clarifying the magnitude of estimate differences between prospective and retrospective paradigms which are thought to employ different cognitive processes. A common finding is that prospective estimates tend to be greater on average compared to retrospective estimates (Brown, 1985; Kikkawa, 1983; Kurtz & Strube, 2003; Zakay, 1992). This difference between estimation paradigms is provided by Block and Zakay (1997) where they presented mean estimation ratios for 20 experiments. These ratios were calculated by dividing the participants' estimated time by objective time in each experiment. Ratios closer to one indicate closer estimate accuracy. The mean ratio for

prospective estimates was found to be .89 compared to .77 for retrospective estimates. As such, prospective estimates are more accurate. Although estimates in both paradigms were fairly accurate (a ratio of over .76), the prospective to retrospective estimation ratio was 1.16. This indicated that prospective estimates in this sample were 16% greater on average than retrospective estimates.

Furthermore, prospective estimates tend to contain less variability than retrospective estimates (Brown, 1985; Brown & Stubbs, 1992; Predebon, 1995). In Block and Zakay (1997), the average coefficient of variation (i.e., the standard deviation divided by the mean) was .33 for prospective estimates and .37 for retrospective estimates. The average retrospective to prospective coefficient of variation ratio was 1.15. Retrospective estimates thus contained 15% more inter-subject variability on average than prospective estimates.

Ostensibly, attention to time explains participants' larger, but less variable estimates when they make prospective compared to retrospective estimates (Hicks et al., 1976; Klapproth, 2007). This attentional effect is evidenced by studies that have adopted a dual-task paradigm in which a secondary task diverts attention away from the timing task (Brown & Stubbs, 1992; Predebon, 1996). Under this condition, people perceive time as passing more quickly because they are not fully attending to the passage of time (Casini & Macar, 1997).

In comparison, the instructions that are given to participants in retrospective time estimation tasks do not explicitly require them to pay attention to time. Hence, memory is thought to be the dominant factor under this condition. In fact, before Hicks et al. (1967), retrospective estimates were known as *remembered durations* (Wearden, 2016). This expression implies that participants use their memory of events to estimate the passage of time. That is, they use previous experiences as temporal references to help them keep track of time. Calling

retrospective estimates “remembered” was eventually deemed to be confusing, however, because memory is also an essential part of prospective estimation models (e.g., Ogden, Wearden, & Jones, 2008).

Despite potential differences in the underlying processes, some researchers have chosen to emphasize the similarities between time estimation paradigms. For example, French et al. (2014) have argued that both paradigms rely on reference memory, suggesting that time estimation is learned through experience. Moreover, some research suggests that attention to the task at hand does play a role in retrospective estimates (Cahoon & Edmonds 1980; Underwood & Swain, 1973). Although the experimental instructions in retrospective paradigms might not explicitly require it of participants, some attention is necessary to complete certain tasks. Indeed, for incidental encoding of temporal information to occur, some attention must have been payed to time throughout the task. Consequently, cognitive models are being developed to clarify the roles of attention and memory in time estimation. Currently, prospective and retrospective time estimates are often explained using two different frameworks.

## Contemporary Models of Time Estimation

**The Attentional-Gate model.** Traditionally, attentional demands in time estimation have been discussed using the Attentional Gate Model (AGM; Zakay & Block, 1997). The AGM is thus the most prevalent model used to explain prospective time estimation. In this model, attention is thought to act as a gate for pulses (i.e., internal clock time units) to be counted accurately by an accumulator. The accumulator uses working memory and long-term memory to report on the accumulation of pulses previously experienced (Block, 2003; Zakay & Block, 1997).

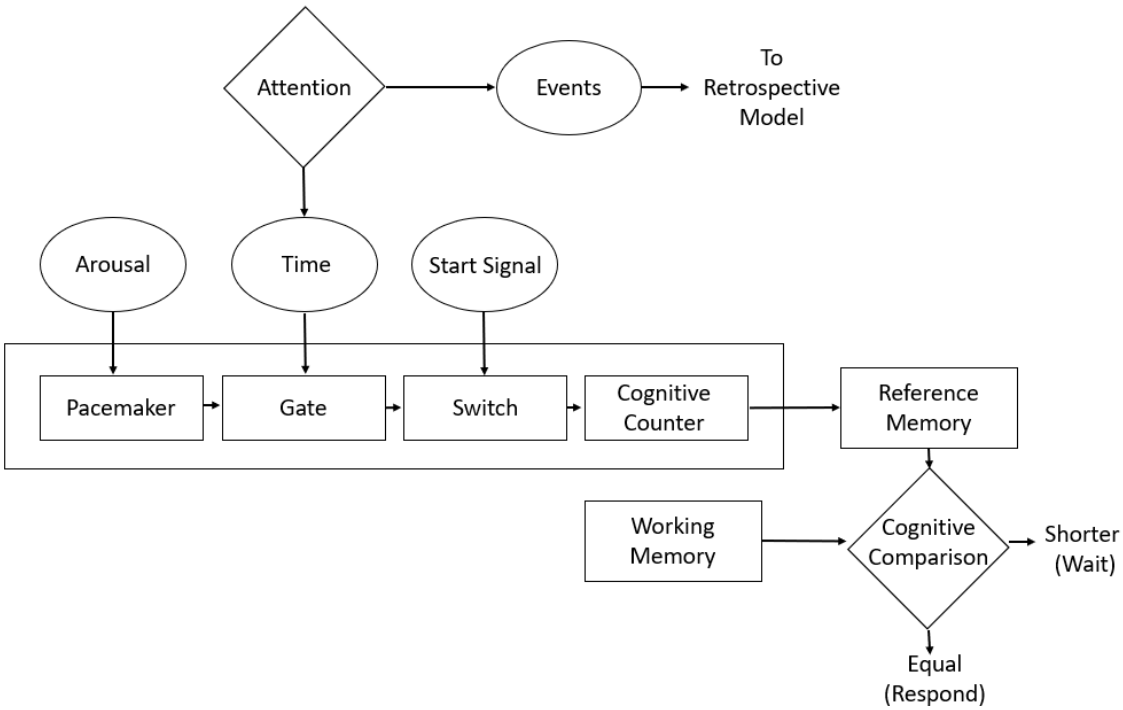


*Figure 1.* Attentional-Gate Model based on Zakay and Block (1997).

When attention is allocated to the passage of time, more pulses are accumulated, and a person is more likely to overestimate the experienced duration. Conversely, when attention is diverted away from a timing task, it is assumed that fewer pulses reach the accumulator (Gibbon, Church, & Meck, 1984) and subjective duration is shortened. This results in feeling as though time passed more quickly. Therefore, distractions decrease perceived duration as the gate fails to accurately record time (Burle & Casini, 2001). This model assumes that attentional resources are needed to perform a time estimation task and that these resources are distributed from the same pool as other attentional tasks.

Although prospective estimates involve paying attention to time, the AGM does not include a mechanism that can account for it. Accordingly, there is no way for it to assess whether at a given moment pulses are freely flowing through the gate or not, and whether these pulses are being accurately accumulated. Furthermore, the accumulation of signals let in by the attentional

gate is assumed to be kept in memory without decay and without interference by subsequent accumulated signals (Taatgen, Rijn, & Anderson, 2007). Consequently, it is unknown how requesting multiple time estimates might interact with this model. Specifically, it does not explain the time-order effect, where the first two tasks of equal duration are judged to be longer than subsequent tasks when performed sequentially (Block & Zakay, 1997). Finally, although attention is outlined as affecting the gate component, it is unknown if attention affects internal clock speed, the switch activating the accumulator, or the gate facilitating the flow of pulses (Burle & Casini, 2001).

These weaknesses highlight the fact that the AGM does not completely explain how time estimates are produced. In fact, the AGM have been the focus of debate in the literature (see Lejeune, 2000; Buhusi & Meck, 2006) and new attention models have been suggested (e.g., e.g., ACT-R; Taatgen et al., 2007). Some models attempt to explain why differences in time estimation accuracy occur between paradigms (e.g., GAMIT model; French et al., 2014). This is necessary because the AGM fails to account for how attentional demands interact with retrospective estimates. Rather than leading to underestimation of a perceived duration like in prospective paradigms, retrospective estimates increase when more attention is directed toward non-temporal aspects of a task (Block et al., 2010). In contrast to prospective paradigms, this finding suggests that time is perceived as passing more slowly under distracting conditions. Findings such as this highlight how estimate differences between prospective and retrospective paradigms are central in that they challenge clock models and encourage their evolution (French et al., 2014).

**Contextual change.** To address the shortcomings of the AGM with regards to retrospective time estimates, theories that focus on the role of memory have emerged. Like the

AGM, models of retrospective estimation are not free of assumptions. Estimates in retrospective paradigms are thought to rely heavily on contextual cues to tag moments in memory. For instance, Block and Reed (1978) provided an overview of retrospective timing models that emphasize this assumption. The *event-memory hypothesis* considers the nature of events encoded during a target interval. For example, a participant might remember stimuli presented during the interval based on position, complexity, or sensory modality. These idiosyncrasies act as events containing contextual cues and provide information about how much time has passed. Contextual cues are assumed to be retrieved automatically when a person is asked how much time has elapsed (Block & Zakay, 2008).

The event-memory hypothesis often used to explain retrospective estimation is called the *contextual change hypothesis*. This theory asserts that change in context throughout an interval accounts for change in retrospective estimates. Block and Reed (1978) suggested that this model provides insight into the time-order effect. Because context changes rapidly at the onset of an experiment, stimulus intervals are judged to be longer at the beginning of a given duration (Hintzman, Block, & Summers, 1973).

Furthermore, the effect of segmented time when intervals are filled versus empty can also be interpreted by a change in context. Lim, Kum, and Lee's (2015) recent study on contextual change lends support to this hypothesis. In this experiment, participants were randomly assigned to a prospective or retrospective time estimation paradigm. All participants watched a video of people dining at a restaurant. In one condition, however, participants observed a dining scene with four components: waiting in line, arriving at the table, browsing the menu, and giving the order. In another condition, participants observed only the first two components of the restaurant scene. Both video conditions were manipulated to have the same duration. Participants who

made prospective time estimates underestimated the four-segment condition relative to the two-segment condition (a mean estimation ratio of 1.14 compared to .93). In contrast, participants who made retrospective estimates overestimated the four-segment condition relative to the two-segment condition (a mean estimation ratio of .92 compared to 1.04). Lim et al. concluded that changes in the scenes diverted attention away from the prospective timing task. As predicted by the AGM, this lead to underestimation. By contrast, a contextual change model can account for the retrospective results. An increased number of segments lead to increased contextual cues, lengthening retrospective estimates in the four-segment condition.

Although it is often argued that separate models are needed to explain prospective and retrospective estimation of time (Zakay & Block, 1997), both can be examined within a contextual change framework. Instructing participants to keep track of time in prospective paradigms could be directing their attention toward smaller contextual changes. This would then prompt the retrieval of a greater number of contextual cues when reporting on time elapsed, creating a tendency for prospective estimates to be greater than retrospective ones. As such, the instructions given to participants at the start of an experiment create a biasing effect on the reported estimate.

**Contextual change, duration, and task demands.** The contextual change hypothesis makes a critical prediction about the length of a target duration. As the duration increases, it predicts that fewer changes occur per unit of time (Wearden, 2016). In other words, the continuation of a task makes it less memorable over time. Increasing the length of the interval to be estimated while holding all else constant should lead to underestimation, especially in retrospective paradigms where memory primarily constructs the time estimate.

Additionally, the *scalar property of variance* in time estimation should be observed as the target duration increases: The variability of time estimates increases with longer durations as participants become less sensitive to a unit change in time. For example, it is easier to notice that one second has passed when the interval is six seconds long compared to when it is 106 seconds long. The scalar property of variance is an important characteristic of human time perception because it is attributed to people's sensitivity to time (Wearden, 2016). Thus, estimates vary less for shorter durations because the unit changes are easier to detect and encode. In longer tasks, fewer cues are retrieved to provide clues as to how much time has passed. Estimates thus become more variable in retrospect.

Although more variable, retrospective time estimates increase as time elapses. That is, as time increases, so do estimates. These two properties (mean accuracy and the scalar property of variance) provide the basic data that must be explained by any model of time estimation. However, mean accuracy is subject to increasing task demands. From a contextual change perspective, task demands manipulate the number of contextual changes that occur throughout the interval. This increase results in greater retrospective time estimates (e.g., McClain, 1983).

The demands of a task are sometimes manipulated by increasing the number of stimuli presented during the interval. For example, increasing the number of stimuli presented has been shown to increase retrospective estimates regardless of whether participants actively or passively processed information. For instance, Predebon (1996) assigned his participants to make either prospective or retrospective time estimates. They performed either a same-different task which required active processing of stimulus information or they passively attended to stimuli presented on a screen. Furthermore, the number of stimuli presented throughout the interval was either small or large. Under both active and passive processing conditions, retrospective estimates

increased with stimulus quantity. In contrast, prospective estimates decreased with stimulus quantity, but this occurred in the active processing condition only. Predebon concluded that retrospective estimates are influenced most by the number of events occurring throughout the task, regardless of processing demands.

The number of stimuli presented in Predebon's (1996) tasks likely acted as individual events or contextual changes, which were then used to construct a retrospective time estimate. In this respect, the results of any time estimation study must consider both the length of time being estimated and the number of contextual changes that occur throughout the interval when interpreting experimental results. The effects of task duration on time estimation will be reviewed first.

## Task Duration

**Short durations.** Tobin et al. (2010) highlighted how very few direct comparisons between prospective and retrospective time estimation have investigated durations over 121 milliseconds. This finding is depicted in Table 1. Accordingly, durations under 121 ms are considered "short" in the present review. Block and Zakay's (1997) meta-analysis of prospective and retrospective time estimation consists mainly of experiments that employed short durations. Although individual experiments rarely compared the two paradigms directly, it was found that when tasks in this meta-analysis lasted longer than 15 seconds, time estimation accuracy was more likely to be influenced by the paradigm employed. Specifically, prospective estimates were less variable and larger on average relative to retrospective estimates. The shortest intervals in this meta-analysis (5.0-14.9 sec) showed no such paradigm effect.

Table 1.

*Past direct comparisons between prospective and retrospective time estimation*

| Experiment | Duration |
|---|---|
| ***120 seconds or less*** | |
| Avni-babad and Ritov (2003) | 120 s |
| Boltz (2005) | 7-10 s |
| Brown (1985) | 16 and 32 s |
| Bueno Martinez (1990) | 80 s |
| Gruber and Block (2003) | 15 s |
| Hicks, Miller, and Kinsbourne (1976) | 42 s |
| Kalpproth (2007) | 15-40 s |
| Kurtz and Strube (2003) | 30 and 60 s |
| McClain (1983) | 120 s |
| Miller, Hicks, and Wilette (1987) | 32-54 s |
| Predebon (1995) | 10-50 s |
| Predebon (1996) | 48 s |
| Predebon (1999) | 12.5-50 s |
| Zakay (1992) | 3 and 6 s |
| Zakay (1993) | 12 and 15 s |
| Zakay and Fallach (1984) | 10 s |
| ***121 seconds to 20 minutes*** | |
| Block (1992) | 165 s (Exp. 1) |
| | 160 s (Exp. 2) |
| Block, George, and Reed (1980) | 270 s |
| Brown and Stubbs (1988) | 466 and 836 s |
| Brown and Stubbs (1992) | 14.45-19.18 (Exp.1) |
| | 7.7-19.6 min (Exp. 2) |
| Bisson, Tobin, and Grondin (2012) | 14 min |
| Lim, Kum, and Lee (2015) | 60 s – 20 min |
| Kikkawa (1983) | 20 min |
| ***Over 20 minutes*** | |
| Tobin and Grondin (2009) | 8 and 24 min |
| Tobin, Bisson, and Grondin (2010) | 12, 35, and 58 min |
| Bakan (1955) | 60 min |

*Note.* This table is an updated version of Table 1 presented in Tobin et al. (2010).

Block and Zakay (1997) argued that a contextual change framework could appropriately explain the lack of paradigm effect at durations spanning 5 to 14.9 seconds. As time intervals increase, the continuation of a single task becomes less salient in memory. In fact, timing very short durations such as these likely relies heavily on short-term memory. Given that only a small

amount of information can be stored in short-term memory and only for a brief amount of time (Baddeley, Hitch, & Allen, 2009), durations over two minutes require the engagement of long-term memory processes. As such, models such as the AGM might be suited well enough for very short tasks like the ones typically employed, yet it breaks down when longer durations are introduced. A framework that emphasizes the contribution of long-term memory in time estimation, such as a contextual change framework, might then be more appropriate to predict estimation behaviour at longer durations.

Indeed, the length of a task influences retrospective estimates more strongly than prospective estimates because they rely more heavily on memory and the retrieval of contextual cues (Block & Zakay, 2008). These facts lead to the following prediction. As the duration of time to be estimated increases, retrospective estimates become shorter and the difference in accuracy between prospective and retrospective paradigms becomes greater. Additionally, unlike the retrospective paradigm where longer tasks are thought to decrease the number of contextual changes remembered, participants who are attentive to time during the task might strategize by attending to smaller context changes, lengthening their prospective time estimates.

Although a contextual change framework predicts an interaction between the length of an interval to be estimated and the estimate paradigm in which it occurs, the durations itself is predicted to influence time estimation regardless of which paradigm is employed. When a range of short intervals are investigated in a single experiment, there is a general tendency for the briefer intervals to be overestimated and the longer intervals tend to be underestimated. This phenomenon is known as *Vierordt's law* (Brown & Strubbs, 1988). The point at which subjective time estimates and physical duration are equal is termed the *indifference point*, which is also dependent on the range of target intervals being measured. For example, when participants

estimated durations that ranged from 0.5 to 8.5 seconds, the indifference point was found at 3.5 seconds (Eisler, Eisler, & Hellstrom, 2008).

Unfortunately, there is a limited amount of research on Vierordt's law for durations beyond seconds. Roy, Christenfeld, and McKenzie (2005) suggested that Vierordt's law reflects a bias in memory. People tend to remember longer tasks as being shorter than they were because their memory of the interval lacks detail and is unlikely to be a direct reflection of reality. If this is true, underestimation of longer tasks should remain robust across paradigms because memory is an important component in each. Currently, it remains unclear if the type of paradigm interacts with task duration at longer intervals. Differences in the experimental methods used, the findings, and the interpretation of the findings in relation to cognitive processes all contribute to this ambiguity.

**Long durations.** Surprisingly, a review of the literature suggests that durations over two minutes may be seen as "long durations". As was demonstrated by Table 1, the relative paucity of data directly comparing prospective and retrospective estimation paradigms for these durations also makes it difficult to establish how accurate this variety of time estimation is. Moreover, the few experiments directly comparing paradigms that have employed durations over two minutes have sometimes failed to find a difference between prospective and retrospective time estimates. What follows is a review of experiments directly comparing prospective and retrospective estimation paradigms at long durations.

In an early experiment, Kikkawa (1983) asked participants watch a film for 20 minutes. Half the participants were aware that they should keep track of the duration of the film while the other half were not. Participants were relatively accurate at estimating the length of the film and the experiment failed to find an effect of estimation paradigm.

More recently, Tobin and Grondin (2009) asked participants to perform two video game tasks (for eight and 24 minutes each) and a reading task (for eight minutes). The experiment lasted a total of 40 minutes. Upon completing all three tasks, they were asked to judge the length of each task individually as well as estimate the entirety of the experiment. Once more, no significant difference between the paradigms was found. However, all participants were relatively inaccurate with less than 50% providing an estimate within range of the actual duration. Moreover, significant overestimation was observed in their eight-minute conditions (mean estimation ratio of 1.44) and underestimation was observed in the 24-minute condition (mean estimation ratio of 0.80). This result suggests that estimates are influenced by the duration of the task akin to Vierordt's law in shorter tasks. It should be noted, however, that when participants are asked to make multiple retrospective time estimates regarding a series of tasks, the ones that occur toward the beginning of the experiment tend to be judged as longer than subsequent tasks (Block & Zakay, 1997). As a result, a possible time-order effect made it difficult to interpret this experiment unambiguously.

Not all experiments that have compared prospective and retrospective estimates at longer durations have failed to produce an effect, however. For instance, Tobin et al. (2010) asked their participants to play a video game in one of three duration conditions (12, 35, or 58 minutes). The interaction between task duration and paradigm was not significant, but regardless of task duration, prospective estimates were larger than retrospective estimates. In fact, prospective estimates were higher by 23% (mean ratio of 1.38 compared to a retrospective mean ratio of 1.15; Tobin et al., 2010). This is an even stronger effect than what has been typically observed (e.g., 16%; Block & Zakay, 1997). In this experiment, the most pronounced overestimation happened in the 12-minute condition, as Vierordt's law would predict.

Bisson, Tobin, and Grondin (2012) uncovered a similar effect in an experiment conducted with children. They asked children to estimate the passage of time during normal classroom activities which lasted 14 minutes. Children read or engaged in a computer game activity. For both tasks, prospective estimates were significantly greater (mean ratios over 1.20) than retrospective estimates (mean ratios under .80). Although retrospective estimates were more variable than prospective estimates, this variability was not statistically significant. Studies such as Tobin et al. (2010) and Bisson et al. (2012) thus provide evidence for overestimation in prospective paradigms, yet potentially comparable variability for prolonged tasks. However, in the longest experiment directly comparing estimation paradigms, this was not the case.

Bakan (1955) asked participants to perform a visual search task using booklets for one hour. They searched for two-digit numbers in a set of ninety numbers. Once participants completed the visual search on one page, they proceeded to the next page where they continued this task with a new set of digits for one hour. Some participants were not privy to the time and thus made retrospective estimates upon completing the task. In another condition, participants were asked to make prospective time estimates. Both groups were fairly accurate, estimating 60.05 minutes in the retrospective condition and 54.09 minutes in the retrospective condition. This was not a significant difference, however.

In summary, most experiments that have directly compared retrospective and prospective time estimation tasks have used durations that were less than two minutes. These experiments have generated the most robust evidence for a paradigm effect where prospective estimates were larger and thus more accurate than retrospective estimates (Block & Zakay, 1997). Experiments that compared eight to 58-minute durations support the existence of Vierordt's law (Tobin & Grondin, 2009), and an overestimation of prospective estimates (Tobin et al., 2010; Bisson et al.,

2012). However, some research suggests that there is no such difference between estimation paradigms (Kikkawa, 1983; Tobin & Grondin, 2009). Finally, the only experiment lasting one hour yielded no effect of estimation paradigm (Bakan, 1955). Evidently, more research is needed investigating the influence of long durations on time estimation. Moreover, as predicted by contemporary models in this area, task duration might interact differently with estimation paradigm depending on the difficulty of the task employed.

**Task Demands**

Most time estimates are obtained after participants engage in some sort of task. For instance, past experiments have employed a range activities including a Stroop task (Zakay & Fallach, 1984; Predebon, 1999), memorizing and rehearsing lists of items (Miller, Hicks, & Willette, 1978), tracing a star figure through a mirror (Kurtz & Strube, 2003), identifying the number of angles on a tactile stimulus (Zakay, 1993), listening to a piece of music (Brown & Stubbs, 1988; 1992), and even passively watching a boiling pot of water (Block et al., 1980) or a flickering lightbulb (Zakay, 1992). The mental effort required to do these tasks obviously varies greatly.

The range of tasks used in time estimation research makes it difficult to compare the results of one experiment to the next. Specifically, task demands can interfere with time estimation. For instance, rehearsing lists of items (e.g., Miller et al., 1978) is known to produce an underestimation of time passed (Franssen, Vandierendonck, & Hiel, 2006; Rattat & Droit-Voilet, 2012). Tasks that use rehearsal thus interfere with timing. In fact, interference in timing by a concurrent task is a well-documented phenomenon in prospective paradigms (see Brown, 1997 for a review). The amount of effort demanded of a task is defined as *cognitive load* (Block et al., 2010).

**Prospective dual-task conditions.** In prospective paradigms, people must divide their attention between keeping track of time and a concurrent task. Therefore, they may be viewed as dual-task paradigms where participants are engaging in both a temporal task and a non-temporal task (Zakay & Block, 2004). Interference happens when a concurrent, non-temporal task disrupts timing performance. Almost all concurrent tasks in a review by Brown (1997) led to underestimation of a given duration when a prospective paradigm was used. These studies included tasks that require perceptual motor-coordination (e.g., mirror drawing; Brown 1985), perceptual discrimination (e.g., loudness discrimination; Grondin & Marcar, 1992), visual or spatial processing (e.g., visual search; Fortin, Rousseau, Bourque, & Kirouac, 1993), and verbal tasks (e.g., proofreading; Brown & Stubbs, 1992). Furthermore, underestimation is a function of task difficulty: tasks that demand high cognitive effort produced shorter estimated durations. This effect was found for rotor tracking, visual search, mental arithmetic (Brown, 1997), and a rhyming task (Zakay, Nitzan, & Glicksohn, 1983).

Arguably, keeping track of time in prospective paradigms demands limited attentional resources that can be diverted away by a concurrent task (Brown, Collier, & Night, 2013). This is also sometimes referred to as the *attention-allocation model* (Brown 1997, 2006) because both the timing task and the concurrent task are thought to draw on the same limited attentional resources. Within an attention allocation model, the aforementioned experiments provide plenty of evidence in support of the AGM (Block, 2003). When attention is diverted away from a timing task in the AGM, fewer pulses reach the accumulator (Gibbon et al., 1984). This creates a subjective shortening of the duration because each pulse is counted toward a time estimate. Fewer pulses equate to fewer perceived moments in attentionally demanding tasks.

To build on the attention-allocation model, time estimation research is currently concerned with identifying the cognitive resources shared between time keeping and concurrent tasks. This is done specifically by testing interference effects where a concurrent task disrupts time estimation and where time estimation disrupts task performance (Ogden et al., 2014). Dual-task interference has been found in tasks that require memory updating (Ogden, Salominaite, Jones, Fisk, & Montgomery, 2011), choice reaction time (Rattat, 2010), sequence reasoning, and sequence monitoring (Brown & Merchant, 2007). However, studies measuring the effects of a concurrent task on timing performance suffer from the same limitations as the ones reviewed in relation to prospective and retrospective time estimation comparisons. First, interference effects vary according to whether verbal estimation, production, or reproduction is used (Ogden et al., 2014; Droit-Volet, Wearden, & Zelanti, 2015). Second, these studies have employed a range of durations. Both the estimation method employed, and the duration of the task can interact with the effect of task difficulty on time estimation.

Given that the present experiment employed a verbal estimation method, predictions were fueled by previous literature outlining how task difficulty influences verbal estimates. The results have been mixed. In Zakay and Shub's (1998) Experiment 1, participants performed either a Stroop task or a card-sorting task. Each of these tasks contained two difficulty levels (easy or difficult) and participants were randomly assigned to one of these four conditions. They were also either instructed to press a button after 32 seconds elapsed (temporal production) or to verbally report how long the task was in seconds when they heard a "beep" sound (verbal estimation). Time estimates recorded via the production method were longer under high cognitive load, regardless of whether participants were doing the Stroop task or the card-sorting

task. This effect was not found when estimates were provided verbally. Instead, time estimates were insensitive to changes in cognitive load.

In contrast, Hemmes, Brown, and Kladopoulos (2004) found no difference between a reproduction method and verbal estimation in a number-reading task. Participants in this experiment either performed a concurrent task while keeping track of time or they did not. The concurrent task condition yielded underestimations in both reproductions and verbal estimations of time.

To summarize, Zakay and Shub's (1998) experiment suggests that it is perhaps more difficult to detect an interference effect between task demands and time perception when verbal estimates are made. Block et al. (1998) argued that this occurs because verbal estimates are more variable. However, Hemmes et al.'s (2004) experiment suggests that interference can be detected even with a verbal estimation method. As such, additional research was needed to clarify how task demands interact with time perception when estimates are made verbally.

**Duration and task demand in prospective conditions.** Notwithstanding the effort that people must put into a given task while keeping track of time, short intervals should still be overestimated, and long intervals should still be underestimated according to *Vierordt's law* (Brown & Strubbs, 1988). As a result, resource-allocation models and *Vierordt's law* make the critical prediction that longer durations containing a difficult task should yield the most dramatic underestimations. When fewer attentional resources are available in conditions of high task demands, fewer pulses are accumulated according to the AGM. Therefore, attentional processes and *Vierordt's law* should work toward the same shortening effect on time estimation in lengthy high load tasks and prospective paradigms.

Evidence in support of this prediction was found by Duzcu and Hohenberger (2014). Participants in one experiment performed a minimally demanding task where they pressed a key with their left or right hand depending on if a coloured rectangle appeared on the right or left of the screen. They were then asked to reproduce the duration of each rectangle by pressing a key. Participants in second experiment performed a more demanding Simon task and were also asked to reproduce the duration in the same manner as the first experiment. Although participants in this second experiment also responded to the location of coloured stimuli, they were instructed to press the left or right key depending on which colour was presented, rather than the location it was presented in. When the location of the stimuli and the hand requiring the motor response were incongruent, participants had to ignore interfering information about location to accurately produce response, making the task more difficult. In each experiment, participants performed their respective task for 15, 30, or 45 seconds. Duzcu and Hohenberger (2014) found that the shorter task conditions (i.e., 15 seconds) were not underestimated, even for the most difficult task. However, task difficulty was found to shorten time estimation when the task was longest (i.e., 45 seconds).

More evidence was needed to confirm whether such an interaction exists. This was thus done in the present experiment by manipulating task difficulty and duration to observe how increased cognitive load interferes with timing ability for tasks of different lengths. Investigating the effects of cognitive load was meant to help clarify why some research has failed to find a paradigm effect at long durations (Bakan, 1955; Kikkawa, 1983; Tobin & Grondin, 2009). Consequently, it was also important to consider how cognitive load interferes with task duration when retrospective paradigms are employed.

**Duration, task demand, and retrospective conditions.** As evidenced by interference effects, more mental effort is required of participants in prospective paradigms because they must divide their attention between the primary task and a time keeping task (Zakay, 1998). Unlike prospective paradigms, however, retrospective paradigms do not involve intentional effort to encode temporal information. Participants are instead fully attending to the non-temporal task and estimation relies on incidental encoding of time (Brown, 1985). Therefore, estimates in retrospective paradigms are not made under dual-task conditions. Still, the contextual change model predicts that retrospective estimates are influenced by cognitive load under certain circumstances. Higher cognitive load tasks that increase the number of contextual changes throughout the interval are predicted to lengthen perceived duration (Block, Hancock, & Zakay, 2010).

In an experiment where participants switched back and forth between two different versions of the Stroop task (colour-word vs. word-colour), the number of switches directly influenced retrospective estimates (Zakay & Block, 2004). Time estimates became longer with increased task switching. This result conformed to the contextual change hypothesis and provided insight about perceived time in situations where people engage in multitasking behaviour. Although task switching demands mental effort, it also divides the interval into multiple segments. This segmentation had been shown to increase time estimation in retrospective conditions (Zakay, Tsal, Moses & Shahar, 1994; Lim et al., 2015). An overestimation effect is likely not a function of cognitive load per se, but the perceived segmentation of the duration brought about by an increased number of memorable events under high load.

The relationship between cognitive load and retrospective time estimation was a focus in a fairly recent meta-analysis by Block et al. (2010). This meta-analysis found a small but positive interaction between retrospective paradigms and cognitive load. Under high cognitive load, time was overestimated relative to low cognitive load conditions. As predicted by a contextual change framework, when experiments featured increased interruptions or multiple high priority events (when the tasks featured more segmentation), estimates increased. Notably, this effect was only present in tasks over 60 seconds. Shorter intervals did not increase retrospective estimates under high cognitive load.

Recall that both retrospective conditions (Block & Zakay, 1997) and increased task duration (e.g., Tobin & Grondin, 2009) produce underestimation of time. For estimates to increase with duration when they are made retrospectively, cognitive load must then counteract both the effect of duration and estimate paradigm. Otherwise, the results in Block et al. (2010) would not be observed. Evidently, the influence of cognitive load is quite strong. In fact, Block et al. (2010) provide a striking picture of how cognitive load interacts with both prospective and retrospective estimate paradigms. This is outlined in Figure 2. It is understood that prospective time estimates tend to be greater than retrospective estimates (Block & Zakay, 1997). However, they become shorter under high load conditions (Brown, 1997). Taken together with the fact that people make greater retrospective estimates under high load conditions, there is a complete reversal of effects typically observed in each paradigm. Remarkably, retrospective estimates become longer and thus more accurate than prospective estimates under high cognitive load conditions.

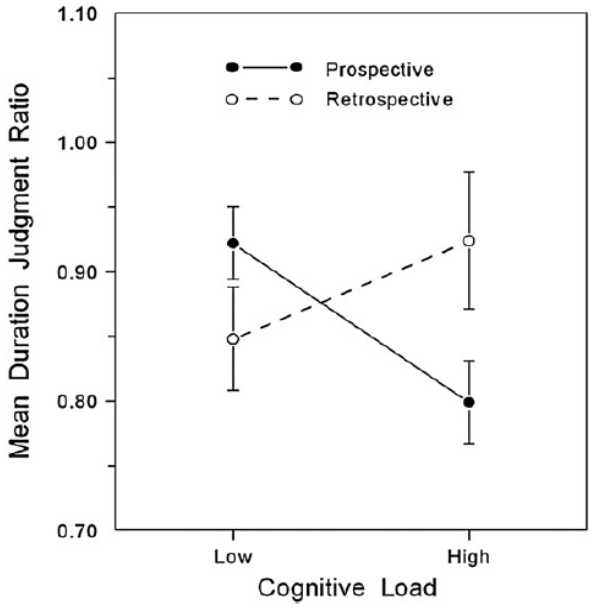


*Figure 2.* Interaction between estimation paradigm (prospective and retrospective) and cognitive load (high and low; Block, Hancock & Zakay, 2010).

Contrary to retrospective paradigms, Block et al. (2010) found that duration did not moderate time estimates in prospective paradigms. They also found that stimulus modality (in line with Boltz, 2005), number of trials (contrary to McClain, 1983), effect of age (contrary to Block et al., 1998) all failed to produce an effect. These findings have important implications for arguments that the two estimate paradigms rely on the same cognitive mechanisms (Brown & Stubbs, 1992). Given that this meta-analytic review comprised over 100 studies on time estimation, it provides strong evidence in support of different estimation processes underlying each paradigm. Reconciling this evidence under one cognitive model is an important challenge and requires directly comparing time estimates made under prospective and retrospective conditions.

The cross-over interaction between paradigm and cognitive load is especially problematic for clock models. As French et al. (2014) pointed out, there is no a priori reason for this effect to

occur in the AGM. Retrospective estimates simply do not fit this model. Considering retrospective time estimation receives little experimental attention in comparison to its prospective counter-part and that contextual change frameworks require formal modelling (Block & Zakay, 2008), more evidence is needed regarding this cross-over effect. Understanding it could help advance current cognitive models for time perception.

Additionally, there exists some confusion as to the role of task duration in cognitive load research. Block et al. (2010) found an effect of duration only in retrospective paradigms. However, cognitive load shortened time estimates regardless of task duration in prospective paradigms. This contradicts results in Duzcu and Hohenberger (2014). Notably, 60 seconds in this meta-analysis was considered "long" thus making it unclear how cognitive load and paradigm interact at very long durations from minutes to near an hour. To investigate this area further and provide a clearer picture of how accurately people can keep track of time, it was necessary to study the effect of estimation paradigm, task duration, and cognitive load simultaneously.

## The Present Experiment

The purpose of the present experiment was to directly compare prospective and retrospective time estimates for different durations and under low and high cognitive load. Participants thus performed a visual and memory search task and were then asked to provide a verbal time estimate. Prior to beginning the task, participants were randomly assigned to a prospective or retrospective time estimation condition. Within these two conditions, the task took either eight minutes to complete (i.e., a long duration; as per Tobin & Bisson, 2009) or 58 minutes to complete (i.e., a very long duration; as per Tobin et al., 2010). Finally, within these four conditions, the task involved a low or high cognitive load. The time estimates were

requested from participants in all conditions at the end of their task. This 2 (paradigm: prospective vs. retrospective) x 2 (duration: long vs. very long) x 2 (cognitive load: high vs. low) design was used to help reveal the way these factors interact to influence verbal time estimation.

This design was meant to primarily help fill current gaps in the time estimation literature. Not only has the prospective paradigm dominated the research, especially with respect to cognitive load and task interference (Brown, 1997), but very few direct comparisons have been done between prospective and retrospective paradigms (Tobin et al., 2010). Moreover, what is known about cognitive load and time estimation comes primarily from short durations (Brown, 1997). There was evidently a need to investigate effects of estimation paradigm, task duration, and cognitive load. Doing so helped determine the effects of these variables on time estimation and helped determine how they interact with each other.

Second, understanding conditions under which each paradigm yields the most accurate perception of time was intended to reveal fundamental processes underlying intentionally or incidentally encoding of temporal information. This design was used to help inform cognitive models in time perception research. Specifically, clock models such as the Attentional-Gate Model are challenged by growing research involving retrospective paradigms (French et al., 2014), making this experiment especially pertinent. Furthermore, the contextual change hypothesis for retrospective paradigms is vague and requires the collection of more data for formal models to develop. To achieve these goals, it was necessary to use a reliable method for manipulating cognitive load.

**An alternative approach to measuring cognitive load.** If prospective time estimation demands attentional resources that can be diverted away by a concurrent task (Brown et al., 2013), then there must be a capacity limit on the amount of resources available. A critical

component of this limited capacity prediction is a practice effect. That is, practice within a given task increases available resources within this limited capacity system, resulting in better dual-task performance. This is known as *automaticity* (Logan, 1988). In other words, when interference effects are reduced with practice on a task, this practice is interpreted as having induced automaticity (Posner & Snyder, 1975). Whether it be attention, response selection, motor execution, or some other component, the cognitive processes necessary to carry out a task become automatic and no longer require conscious awareness.

When participants are instructed to keep track of time while engaging in a well-practiced concurrent task, the amount of interference should be significantly lower than in a non-practiced task (Brown & Bennet, 2002). This would signify the reallocation of resources from the concurrent non-temporal task to time keeping. What is most critical about automaticity is that it occurs even though tasks might require a range of cognitive resources. For example, Brown and Bennet (2002) found that practice in both a rotor-tracking task and a mirror-drawing task reduced timing interference. Considering the wide range of tasks employed in time perception research (Block et al., 2010), automaticity provides a generalizable measure of cognitive load. More specifically, with increased duration, practice should increase automaticity and reduce the cognitive load of any task.

Indeed, the relationship between automaticity and cognitive load is well-known. Schneider and Shiffrin (1977) established the link clearly when they formulated their theory of controlled versus automatic information processing. Controlled information processing requires attentional resources, is capacity limited, and is dependent on task demands. In contrast, automatic processing is largely unaffected by task demands because it relies on long-term memory and requires attention only when a well-learned response is needed. To demonstrate

controlled and automatic processing, Schneider and Shiffrin (1977) used a visual and memory search task. Participants were presented with either digits or consonants to be remembered (i.e., the memory set) and then searched for these characters among distractors in a visual array. If the stimuli presented in the memory set also appeared in the visual array, participants were to generate a "target-present" response. Otherwise, they responded that the target was absent.

The key manipulation in Schneider and Shiffrin's (1977) experiment was whether the alphanumeric characters in the memory set on a given trial could act as distractors on subsequent trials. This manipulation was meant to distinguish between the effect of consistent versus varied mapping of targets to responses. *Consistent mapping* is obtained when a given response is always associated with a specific stimulus. In Schneider and Shiffrin's experiments, this meant that memory set items could never be distractors. Mapping a stimulus to a response like this allows for automaticity to build. On the contrary, *varied mapping* is obtained when there is no one-to-one correspondence between a given stimulus and its associated response. For instance, specific digits or consonants could appear in either the memory set or the distractor set so that participants' responses are not mapped to the same stimuli across trials.

The impact of response mapping and its relation to automaticity has been examined recently in the same-different task (Walker & Cousineau, submitted) and extensively in visual search (e.g., Bargh,1992; Cousineau & Shiffrin, 2004; Czerwinski, Lightfoot, & Shiffrin, 1992; Hélie & Cousineau, 2011; Cousineau & Larochelle, 2004; Lefebvre, Cousineau, & Larochelle, 2008; Moors & De Houwer, 2006; Schuster, Rivera, Sellers, Fiore, & Jentsch, 2013). Collectively, these studies highlight that any condition featuring consistent mapping should lead to automaticity and thus lower cognitive load.

Applying this manipulation to an experiment featuring a concurrent time keeping task enables direct observation of cognitive load effects on time estimation. Most importantly, the concept of automaticity directly corresponds to the amount of time spent on a cognitive task (Schneider & Shiffrin, 1977). As such, cognitive load should decrease more drastically with time when stimuli are consistently mapped to responses and automaticity sets in. Accurate time estimation performance should then increase. This makes it possible to test an interaction between task duration and cognitive load on human time perception.

**Visual and memory search.** Past research on time estimation in visual and memory search tasks provide some insight into what might be expected to occur in the proposed experiment. As mentioned above, participants in a visual and memory search task scan an array of characters in search for specific targets. This process is typically repeated over many trials. Targets come from a memory set shown prior to the array and this memory set usually contains digits or letters that can vary in number. The same is true for the array. Any number of characters can be used as distractors. This is called *set size* and increasing set size makes the visual search process more challenging for the participant (Schneider & Shiffrin, 1977).

Fortin and Rousseau (1987) investigated the effect of performing visual and memory search tasks while making time estimations. In this task, participants were shown a memory set consisting of one to six digits ranging from zero to nine. The search array consisted of one digit and participants had to quickly identify whether this digit was part of the memory set. In line with a varied mapping procedure, targets changed on each trial so that specific digits were not consistently associated with a given response. Concurrently, participants produced two-second intervals by pressing a button. Fortin and Rousseau found that for each one digit memory set size

increase, the average interference in temporal production was 22 milliseconds. They concluded that set size interferes with time perception.

In a series of follow-up experiments, Fortin, Rousseau, Bourque, and Kirouac (1993) replicated Fortin and Rousseau (1987), and then created another visual and memory search task where the target differed from all the distractors in the set based on only one feature. This way, participants never had to remember a new target with each trial. Using this consistent mapping method, time estimates were unaffected even with a large set size. That is, no interference occurred in temporal productions. These results outline how response mapping in a concurrent task impacts time perception.

In Fortin et al.'s (1993) experiments, interference in timing performance depended on whether targets were consistently mapped to responses or not. Although memory and attention were necessary to perform the visual and memory search in all the experiments outlined, automatic processing likely developed in consistent mapping conditions, easing task difficulty. Responding to stimuli thus required fewer cognitive resources needed to perform the concurrent timing task so that interference effects were not found with increasing set size. Given that automaticity increases with practice, the longer participants perform the visual and memory search task, the smaller the predicted interference effect in prospective conditions. Response mapping and prolonged duration in the present study thus offered additional insights into the effect of cognitive load in time estimation.

Notably, the experiments outlined above featured prospective time estimation only. Retrospective time estimation in visual search has not been investigated since Bakan (1955). Intuitively, varied mapping should result in an increased number of contextual changes and more changes (or segments) results in the perception of the interval being longer (Lim et al., 2015).

The present experiment thus investigated whether varied mapping in a visual and memory search task does lead to longer retrospective estimates.

## Hypotheses

The first and most crucial prediction of the present experiment was that the effects of both estimation paradigm and task duration depend on cognitive load. This prediction is based on the significant interaction found by Block et al. (2010) in which prospective and retrospective paradigms underwent a complete reversal of their typical effects under high cognitive load. To clarify, a cross-over interaction between the estimation paradigm (prospective vs. retrospective) and cognitive load (high vs. low) was expected to occur. However, it was unknown whether this effect is confined to durations under two minutes, which constitutes most of the studies included in Block et al.'s (2010) meta-analysis.

When cognitive load was low, it was predicted that this experiment would replicate Tobin et al.'s (2010) finding that prospective time estimates are greater than retrospective estimates on average. Additionally, all time estimates were expected to be more variable in the 58-minute condition because of the scalar property of variance (Wearden, 2016). It was also hypothesized that there would be a significant interaction between estimate paradigm and task duration under low cognitive load. Specifically, at 58 minutes, a mean difference between prospective and retrospective estimates would be greater than at eight minutes. This hypothesis was driven by two fundamental features of the task design: Vierordt's law and the use of a verbal estimation method.

Notably, verbal estimation has been shown to interact with estimate paradigm differently. Verbal and retrospective estimation both rely heavily on access to memory (Block et al., 2010; Ogden et al., 2014). Consequently, both verbal estimation and retrospective paradigms were

expected to compete for this limited cognitive resource, causing interference in time perception and pronounced underestimation of time when employed together. Moreover, under low cognitive load, there should be fewer contextual changes as task stimuli register with ease. This was expected to lead to significant underestimation in the retrospective condition. Considering both these contributors to underestimation, the greatest mean difference between paradigms was expected in the 58-minute low cognitive load condition.

When cognitive load was high, retrospective estimates were expected to be greater than prospective estimates. This is the opposite pattern than was predicted in the low cognitive load condition and was drawn from the interaction seen in Block et al. (2010). However, this was hypothesized to be true only in the eight-minute condition when cognitive load would be highest due to less time spent practicing the task (i.e., the lowest automaticity). After 58-minutes of practice, interference effects were predicted to decline with automaticity in the prospective paradigm, increasing accuracy and lengthening estimates. In the retrospective paradigm, practice was predicted to lessen the number of contextual changes remembered throughout the interval. Therefore, underestimation was expected to occur for retrospective estimates, and variability would increase, becoming more like the effects typical of low cognitive load conditions. It was thus hypothesized that estimates would converge, making it difficult to detect an effect of estimation paradigm under high cognitive load at 58 minutes.

## Method

### Participants

Three hundred and eighteen participants with normal or corrected to normal vision were recruited for this study. All participants were students from Carleton University and recruited through a research participation system consisting of students enrolled in an introductory to

psychology course (SONA). As compensation for their time, participants received 1.5% toward research participation credits. This study was approved by Carleton University's Research Ethics Board.

Twenty-six participants did not complete the experiment due to scheduling error (one participant), experimenter error (three participants), technical difficulties (12 participants), or voluntary withdrawal (10 participants). Participants who voluntarily withdrew reported being too tired to complete the experiment or needed to use the washroom. One participant fell asleep during the visual and memory search task. Thus 292 participants between the ages of 17 and 64 (*Mean* = 19.61, *Median* = 19) successfully completed the experiment. This resulted in a total of 37 participants per condition, with one condition (the prospective, high cognitive load, 58-minutes) consisting of 38 participants. Overall, 214 of the participants were female, 74 participants were male, one participant was transgender, and three participants did not indicate a gender.

## Experimental Design

The proposed experiment consisted of a 2 (paradigm: prospective vs. retrospective) x 2 (duration: eight minutes vs. 58-minutes) x 2 (cognitive load: high vs. low) between-subjects design. As such, four out of the eight groups received instructions to keep track of time throughout the task. The other four groups did not know to keep track of time. Participants were randomly assigned to one of these eight experimental conditions.

## Visual and Memory Search Task

The experiment was computer programmed with E-prime (Psychology Software Tools, Pittsburgh, PA) on a PC computer. To ensure that the low and high cognitive load conditions of

the task were sufficiently different from each other, the visual and memory search task mirrored that of Cousineau and Larochelle (2004). All memory set items (i.e., the targets) and non-target items (i.e., distractors) consisted of alphanumeric characters chosen from a set of eight (L, R, S, H, 1, 3, 6, 7). The characters were always arranged in a square formation (two lines containing two digits) and located directly in the center of the screen. The characters were presented in white Times New Roman 18-point font. The background was black.

The use of digits in the task also helped suppress counting strategies when time estimates are made prospectively. When asked to keep track of time, participants might have attempted to use various techniques such as counting or singing (Grondin & Killeen, 2009). This behaviour happens automatically without the participants being instructed to do so (Fetterman, Killeen, & Hall, 1998). Remembering target digits thus inhibited counting because these are competing processes.

The memory set and the search array always contained four alphanumeric characters. On target-present trials, one of the characters in the search array was a target while the other three were distractors. On target-absent trials, none of the four characters in the search array contained a target. For each participant in the low cognitive load conditions, stimuli were consistently mapped to responses. In these consistent mapping conditions, four characters were chosen to remain as targets across all trials. Furthermore, these characters were either all digits or all letters. The other four characters were used as the distractors. As such, participants in the low load condition who saw letters as their targets throughout the task always looked for a letter among digits in the search array. Those who saw digits as targets always looked for digits among letters.

In the high cognitive load conditions, stimuli were variably mapped to responses. In these varied mapping conditions, the targets on one trial could act as a distractor on a subsequent trial. Therefore, on each trial, four characters were randomly selected to serve as the targets and the remaining four characters were randomly selected to be distractors. The targets could thus consist of both digits and letters that change on every trial making these conditions much more difficult.

### Visual and Memory Search Procedure

All participants received a detailed explanation of the visual and memory search task procedure before beginning the experiment. They were asked to maintain an accuracy of 90% or more. For participants in the prospective time-keeping condition, they were told that both their accuracy in the visual and memory search task and their ability to keep track of the time were equally important for the experiment.

The procedure for the visual and memory search is largely based on Cousineau and Larochelle (2004). A typical trial is illustrated in Figure 3 and proceeded as follows. First, a fixation star appeared for 500 ms. This was followed by a memory set containing four target digits for 1000 ms, and then a fixation star for 500 ms. The visual search array containing distractors (and one of the targets on target-present trials) was then be presented for a maximum of 3,000 ms or until a response was made. Finally, feedback on their response was displayed for 900 ms. The inter-stimulus interval (ISI) lasted 100 ms plus the remaining time carried over from the 3,000 ms maximum of the visual search display. Consequently, every trial took 6,000 ms regardless of the participants' RTs. There were 80 trials in the eight-minute task and 580 trials in the 58-minute task. There were also an equal number of target-present and target-absent trials for a total of 40 target-present trials in the eight-minute task and 290 in the 58-minute task.

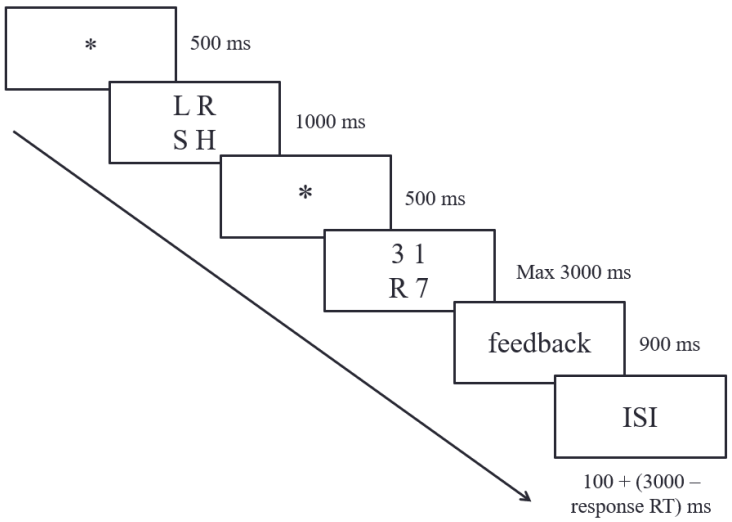


*Figure 3.* Visual and memory search task procedure for a target-present, low cognitive load condition.

Participants responded using the "Z" and "M" keys. These response keys were counter-balanced so that half the participants pressed "M" when they perceived that the target as present and half pressed "M" when they perceived the target as absent. Visual feedback informed the participant whether they correctly identified that the target was present or absent on that trial ("CORRECT" or "INCORRECT"). Feedback on mean accuracy was also be provided here as a percentage (e.g., "96.00% accurate") to keep participants informed on whether they were reaching the performance goal of over 90% accuracy. If no response was made, the trial was recorded as null and "INCORRECT" appeared on the screen, prompting the participant to increase their attention on subsequent trials. As per Fortin et al. (1993), each participant had 20 practice trails before beginning the experiment.

## General Procedure

Participants completed the experiment alone in a quiet room. At the start, they were asked to leave their phone, music devices, headphones, bracelets, watches, and any other distracting

items in a container located inside the testing room. They were also asked to turn off the electronic devices. The container was placed in a position unreachable by the participant in session yet inside the testing room to ensure they are not preoccupied with the location of their valuables. Participants were told that the rationale for removing these items was distractions like wrist jewelry can sometimes interfere with quick motor responses. Finally, there were no clocks or other distractions that could interfere with their search performance.

All participants were informed that the task required each person to complete a randomly selected number of experimental trials chosen by the experimental software. They were informed that regardless their respective task length, they would be guaranteed the full 1.5% toward their research participation credits. As such, participants were assured that the task would not take more than 90-minutes. Prior research has found that estimates are distorted by the time limit set for the task. For example, if participants are told that the task cannot exceed three hours, they overestimate the duration of the task (Thomas & Handley, 2008). Therefore, this 90-minute cap was strategically chosen as to not pull the estimates too far away from the actual time elapsed in the 58-minute conditions. Yet, this cap was still large enough to allow for estimate variation. Moreover, the participants were informed that the software was programmed to generate trials for at least two minutes. Thus, they were made aware that the task could last anywhere from two minutes to 90 minutes in duration.

Participants in the retrospective conditions were told that the purpose of the experiment was to study performance in visual and memory search. This is in line with research that has investigated retrospective time estimation in which the task was disguised as testing perceptual processes (Brown, 1985). They were told that having a random number of trials helps determine the effect of training. Participants in prospective conditions knew that their perception of time

was being studied. However, they were also informed that an equally important experimental purpose is the investigation of accuracy in visual and memory search tasks.

All participants saw the same prompts after completing the visual and memory search task: "*How long was this task from the moment you pressed the 'Start' key to your last response? Your estimate should not include the practice trials completed prior to pressing the start key.* ***Please do not check the time****. Please type your answer in* ***minutes*** *the box below and press 'Enter' to submit a time estimate*". Located underneath this message, there was a small box for their response. Next participants were prompted to estimate the maximum and minimum duration of the task: "*Please fill in the blanks: At the very least, I think the task took me ____ minutes to complete. At very most, I think the task took me ____ minutes to complete. Press 'Enter' to submit your response."* They were then asked to how long they thought the task took to complete including the practice trials at the beginning of the experiment.

Next, participants were asked to indicate how fast or slow time seemed to pass throughout the task on a Likert scale (1 = very slow; 7 = very fast). These constitute passage of time judgements (PoTJ) (Wearden, 2015). Finally, participants were asked to indicate how engaged they were throughout the task on a Likert scale (1 = not very engaged; 7 = very engaged) and how bored they were throughout the task on a Likert scale (1 = not bored at all; 7 = very bored). This was followed by demographic questions about age and gender. They then saw a message on-screen informing them that the task was complete and to check-in with the experimenter.

Participants were asked if they had complied with the instructions not to use personal devices during the experiment. They were then debriefed, and special care was taken to ensure that participants in the retrospective condition understood the true nature of the experiment. They

were also asked whether it was clear before beginning the experiment that they would be asked to provide time estimates. Finally, they were asked to not disclose the time estimation aspect of the study to friends who might eventually participate in the experiment. Finally, participants in the retrospective conditions then signed a secondary consent form allowing the experimenter to use their data.

# Results

## Participant Screening

Of the 292 participants who completed the experiment, 15 were excluded from the analyses. Five were excluded because they reported checking the time throughout the task. All five of these participants belonged to a retrospective, 58-minute condition (four from the low cognitive load task and one from the high cognitive load task). Additionally, one participant from a retrospective condition was excluded because she guessed the true purpose of the experiment. This participant reported mentally keeping track of time throughout the task. Furthermore, four participants were mistakenly provided with prospective consent forms yet were not told by the experimenter to keep track of time. It was therefore unclear to these participants how much attention they should have dedicated to time-keeping throughout the task. These participants were excluded from further analyses. Finally, five participants performed the visual and memory search task at chance. These participants were removed from further analyses because they were likely inattentive to the task and were thus unaffected by the cognitive load manipulation.

## Visual and Memory Search Results

**Data cleaning.** The remaining 277 participants' visual and memory search data were examined for outliers. Each participant's accuracy was first analyzed with respect to the average

performance of their group. Those who had an overall accuracy exceeding three standard deviations above or below their group's mean were excluded from the analyses. This resulted in the removal of an additional four participants' data. All four participants belonged to a prospective, 58-minute condition (one from the high load and three from the low load task). The final sample for this experiment thus consisted of 273 participants after data cleaning. This resulted in a range of 30 to 36 participants per group.

**Visual and memory search accuracy.** In this experiment, variable mapping was meant to implement high cognitive load and consistent mapping was meant to implement low cognitive load. To investigate whether this manipulation was successful, a 2(estimation paradigm: prospective vs. retrospective) x 2(task duration: eight vs. 58 minutes) x 2(response mapping: consistent vs. varied mapping) ANOVA was first performed on participants' mean visual and memory search task accuracy. Unfortunately, Levine's test was significant, $F(7,265) = 23.75$, $p < .001$, suggesting that the group variances were not equal. This was expected, however, considering there should be less variability in an easy task (i.e., consistent mapping) than a difficult one (i.e., varied mapping). Thus, the results should be interpreted cautiously. The results of the ANOVA are shown in Figure 4.

No three-way interaction was found between time estimation paradigm, task duration, and response mapping, $F(1,265) = .05$, $p = .83$. However, a significant two-way interaction was found between task duration and response mapping, $F(1,265) = 7.16$, $p = .008$, $\eta_p^2 = .03$. The duration of the task significantly impacted accuracy in the varied mapping condition only, $F(1,265) = 18.68$, $p < .001$, $\eta_p^2 = .06$. Specifically, mean accuracy in the varied mapping, eight-minute condition ($M = 81.77\%$, $SD = 8.48\%$, *95%* CI [80.37, 83.17]) was significantly lower than mean accuracy in the varied mapping, 58-minute condition ($M = 86.12\%$, $SD = 7.54\%$, *95%*

CI [84.71, 87.53]). In other words, higher accuracy was observed when participants had more practice on the task. However, there was no significant difference in accuracy between the eight and 58-minute tasks when response mapping was consistent. Additionally, the remaining two-way interactions were not significant (*p*s > .43).

Significant main effects were also found for task duration, $F(1,265) = 11.06$, $p = .001$, $\eta_p^2 = .04$, and of response mapping, $F(1,265) = 389.53$, $p < .001$, $\eta_p^2 = .60$. Specifically, participants had higher accuracy overall if they were in the 58-minute task (*M* = 92.31%, *SD* = 8.32%, *95%* CI [91.28, 93.34]) compared to the eight-minute task (*M* = 89.89%, *SD* = 10.27%, *95%* CI [88.91, 90.88]). Participants were also more accurate overall when the task featured consistent mapping (*M* = 98.25%, *SD* = 2.08%, *95%* CI [97.23, 99.28]) compared to varied mapping (*M* = 83.96%, *SD* = 8.29%, *95%* CI [82.95, 84.94]). No main effect of time estimation paradigm was found, $F(1,265) = .68$, $p = .41$.

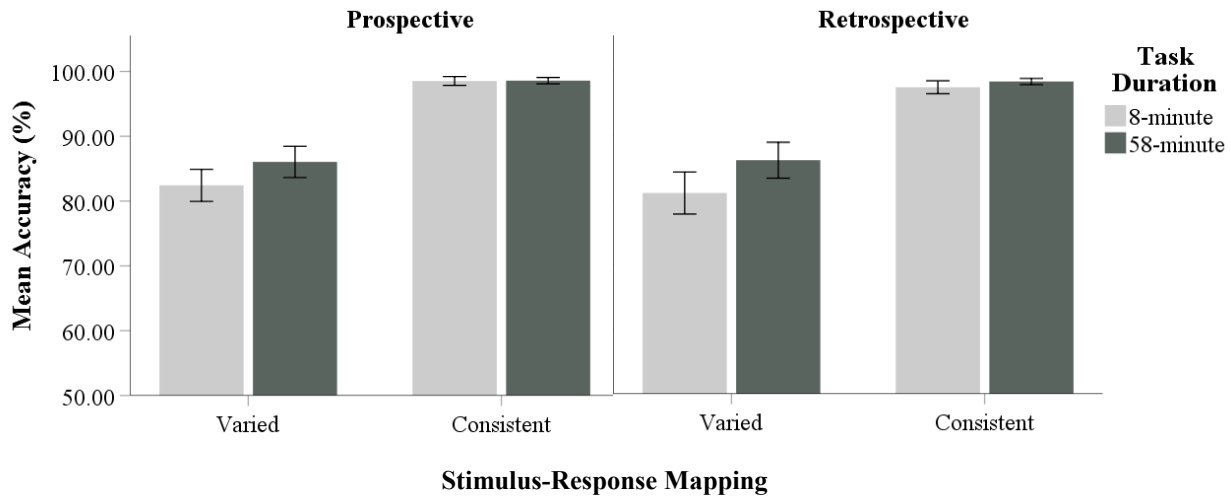


*Figure 4*. Mean percentage of correct responses on the visual and memory search task. The left panel contains the results of participants in the prospective time estimation paradigm. The right

panel contains the results of participants in the retrospective time estimation paradigm. Error bars show the 95% confidence intervals of the mean.

Thus, the response accuracy results suggest that the cognitive load manipulation was successful. Even with minimal practice (the eight-minute condition), it was much easier for the participants to reach high levels of accuracy if they were in a consistent mapping task compared to the varied mapping one.

**Visual and memory search RTs.** In line with previous research, response times (RTs) below 200 ms were removed from analysis because they were assumed to represent response anticipation (Wolfe, Palmer, & Horowitz, 2010). This resulted in the removal of 1,544 trials (1.5% of trials). Finally, only the RTs of accurate trials were used in analysis. On average, target present trials ($M$ = 968.55, $SD$ = 222.91) were significantly faster than target absent trials ($M$ = 1156.47, $SD$ = 407.67), $t(272) = 12.68$, $p < .001$. Considering target present trials were also less accurate than target absent trials, this is the typical speed-accuracy trade-off seen in simple decision tasks. As such, the pattern of RTs was as expected.

A 2(estimation paradigm: prospective vs. retrospective) x 2(task duration: eight vs. 58 minutes) x 2(response mapping: consistent vs. varied mapping) ANOVA was next performed on participants' response times. Once more, Levine's test was significant, $F(7,265) = 3.92$, $p < .001$. The ANOVA revealed a three-way interaction between time estimation paradigm, task duration, and response mapping, $F(1,265) = 6.43$, $p = .01$, $\eta_p^2 = .02$. This is presented in Figure 5. To understand this interaction, the ANOVA was decomposed. A significant simple interaction was found between estimation paradigm and response mapping in the eight-minute condition, $F(1,265) = 13.06$, $p < .001$. However, there was no paradigm x mapping simple interaction effect in the 58-minute condition, $F(1,265) = .002$, $p = .96$. Second order simple effects were then

analysed for the eight-minute condition. The effect of estimation paradigm on RTs was significant for participants who performed the varied mapping task, $F(1,265) = 14.11$, $p < .001$. That is, they had faster RTs in this task if they were also in the prospective condition compared to when they were in the retrospective condition (*Mean Difference* = 175.53 ms, *95%* CI [83.51, 267.55]). In the eight minute consistent mapping condition, no second-order simple effect of paradigm was found, $F(1,265) = 1.84$, $p = .18$. Participants' RTs were similar in this condition regardless of whether they were in a prospective or retrospective timing condition. These results further confirm that the cognitive load manipulation was successful. RTs were slower when participants were in a varied mapping condition. However, this result was moderated by task duration and estimation paradigm.

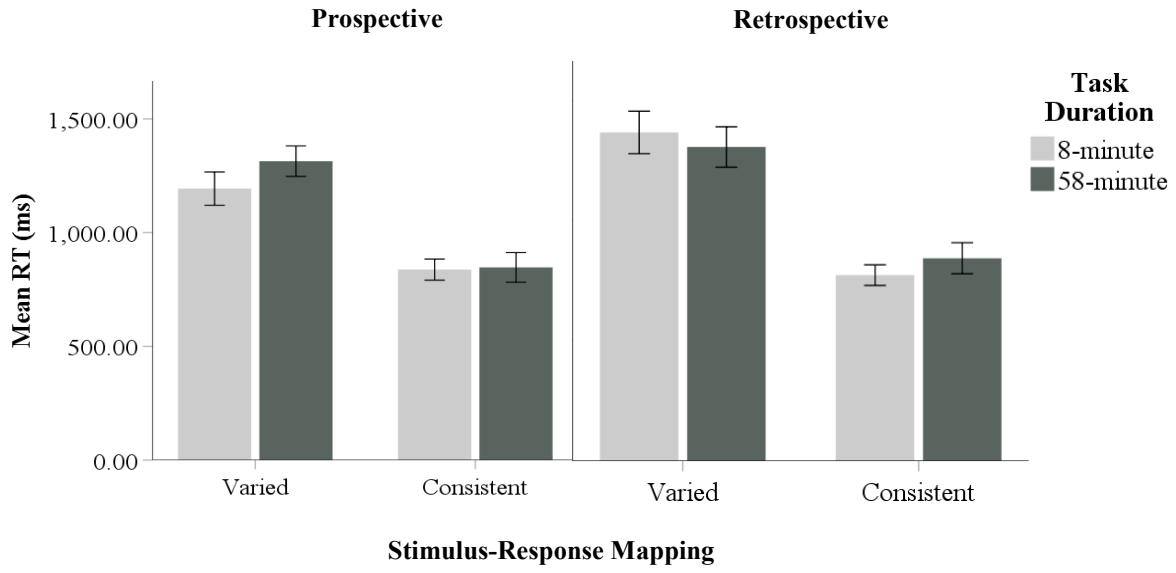


*Figure 5*. Mean response time (RT) in milliseconds (ms) on the visual and memory search task. The left panel contains the results of participants in the prospective time estimation paradigm. The right panel contains the results of participants in the retrospective time estimation paradigm. Error bars show the 95% confidence intervals of the mean.

## Time Estimation Results

Like Tobin et al. (2010), 2 (paradigm: prospective vs. retrospective) x 2 (duration: eight minutes vs. 58 minutes) x 2 (cognitive load: high vs. low) between-subjects ANOVA were conducted on duration estimate ratio results, absolute standard error results, and the Weber's fraction variant results.

**Duration estimate ratio results.** The duration-estimate ratio (RATIO) was calculated following Tobin et al. (2010) by dividing each participant's time estimate (not including the practice trials) ($E_D$) by the actual task duration ($T_D$). If the task was overestimated, this value was greater than 1. If the task was underestimated, this ratio was less than 1.

$$\text{RATIO} = E_D / T_D$$

Four participants in this study did not answer the time estimation question. Additionally, participants' time estimates were screened for outliers. Typically, time estimates are removed if they are three times the standard deviation below or above the mean for their group (Tobin & Grondin, 2009). Two participants provided outlying estimates. One participant belonged to the prospective, high load, 58-minute condition, and the other belonged to the retrospective, low load, eight-minute condition. In both cases, the participants severely overestimated the duration of their task. These participants were removed from analysis of RATIO. On average, participants in this experiment were accurate at estimating the task's duration ($M = 1.05$, $SD = .51$).

Levine's test for homogeneity of variance between groups was again significant, $F(7,259) = 8.91$, $p < .001$. Rather than unequal variance resulting from differences in task difficulty between the groups, the task duration drove this significant result. Variations in responses were significantly smaller in the 58-minute task compared to the eight-minute task. Differences in variance are thus due to the large discrepancy in task durations used in this

experiment. Despite violating the assumption of homogeneity of variance, the time estimation literature typically uses the ANOVA procedure without transforming the data (Tobin et al., 2010). The same was done for the present results.

No significant three-way interaction was found, $F(1,259) = .44$, $p = .51$. However, the ANOVA revealed a significant two-way interaction between estimation paradigm and task duration, $F(1,259) = 4.73$, $p = .03$, $\eta_p^2 = .02$. Following examination of the simple main effects, estimation paradigm had a significant effect on participants' RATIO, but only in the eight-minute task, $F(1,259) = 5.38$, $p = .02$, $\eta_p^2 = .02$. As seen in Figure 6, participants overestimated the eight-minute task's duration more if they were part of the prospective condition ($M = 1.40$, $SD = .51$, *95%* CI [1.25, 1.52]) as opposed to the retrospective one ($M = 1.29$, $SD = .50$, *95%* CI [1.16, 1.43]). The other two-way interactions were not significant (*p*s > .45).

The ANOVA also revealed a significant main effect of duration, $F(1,259) = 142.33$, $p < .001$, $\eta_p^2 = .36$. On average, participants overestimated the duration of the eight-minute task ($M = 1.33$, $SD = .52$, *95%* CI [1.26, 1.40]) and underestimated the duration of the 58-minute task ($M = .73$, $SD = .24$, *95%* CI [.66, .80]). This result is in line with Tobin and Grondin (2009), who found a significant overestimation in their eight-minute conditions ($M = 1.44$) and underestimation their 24-minute condition ($M = .80$). In the present experiment, no other significant main effects were found for RATIO (*p*s > .37).

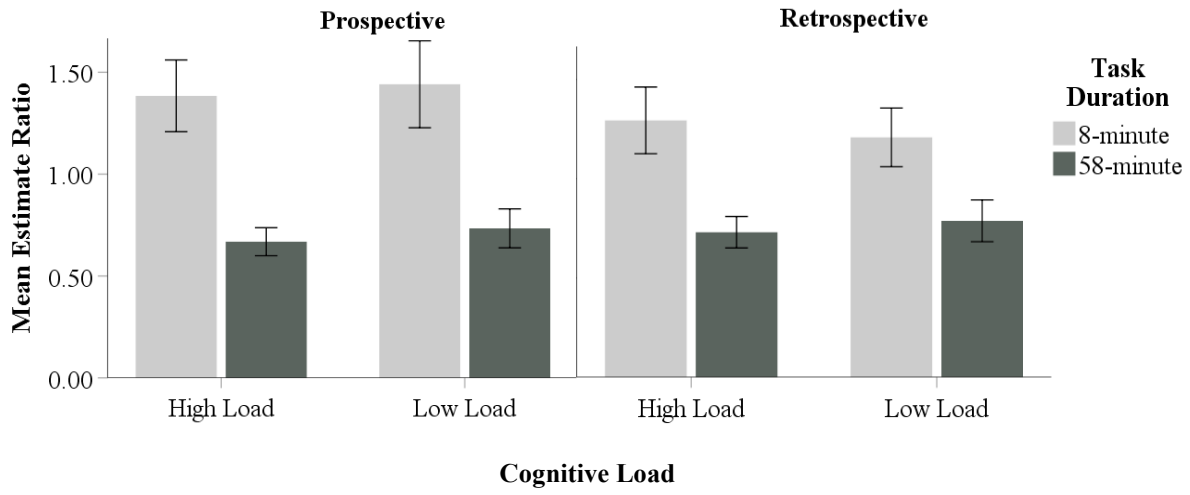


*Figure 6*. Mean time estimation RATIOs (participants' duration estimate over the actual task duration). Left panel = participants in the prospective time estimation paradigm; right panel = participants in the retrospective time estimation paradigm. Error bars show the 95% confidence intervals of the mean.

**Absolute standard error results**. To determine how far the estimate deviated from real time, the second dependent variable was the absolute standard error (ASE). A larger ASE means that the estimate was further from the actual task duration. ASE was calculated by taking the absolute difference between a time estimate and the task duration ($E_D - T_D$) and dividing this by the task duration ($T_D$).

$$\text{ASE} = \text{absolute}\ (E_D - T_D) / T_D$$

Levine's test was significant once again, $F(7,259) = 10.33$, $p < .001$. There was no three-way interaction, nor were there any two-way interactions found, *p*s > .18. These results are outlined in Figure 7. Contrary to Tobin et al. (2010), there was a main effect of estimation paradigm, $F(1,259) = 4.58$, $p = 0.33$, $\eta_p^2 = .02$. On average, participants in a prospective paradigm provided estimates that were further away from the task's actual duration ($M = .43$, $SD = .29$, *95%* CI [.37, .48]) compared to participants who were in a retrospective paradigm ($M =$

.34, *SD* = .37, *95%* CI [.29, .40]). The paradigm effect was only marginally significant in Tobin et al. ($p$ = .067, $\eta_p^2$ = .03). They too found that ASE was higher in the prospective paradigm ($M$ = .49) than it was in the retrospective paradigm ($M$ =.36).

Like Tobin et al. (2010), a main effect of duration was also found in the present experiment, $F(1,259) = 12.11$, $p = .001$, $\eta_p^2 = .04$. Participants who were in an eight-minute condition provided estimates further away from the actual duration (M = .45, SD = .42, *95%* CI [.40, .51]) than participants who were in a 58-minute condition (M = .32, SD = .18, *95%* CI [.26, .37]). Like Tobin et al., the shortest task duration was the one with the highest ASE.

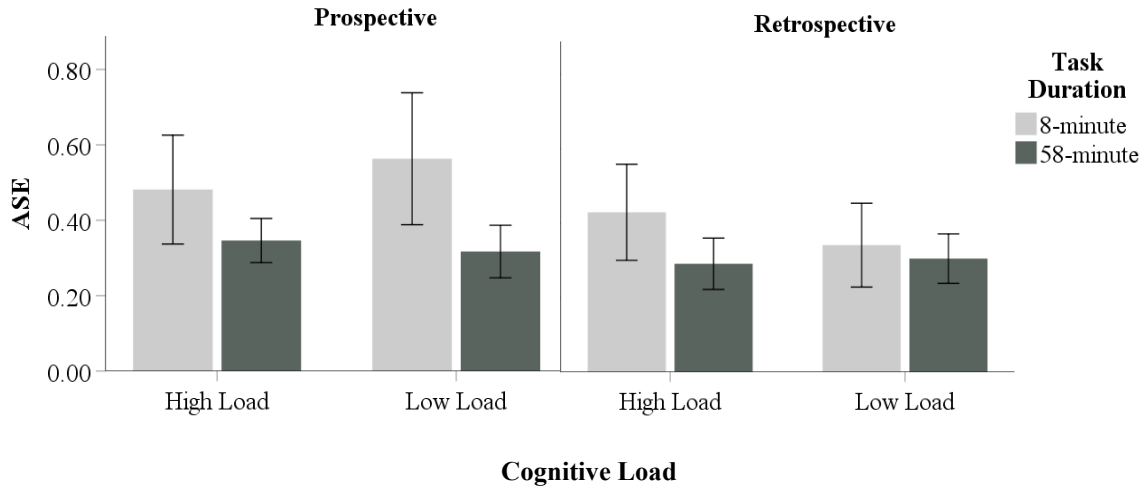


*Figure 7*. Absolute standard errors (the absolute difference between the participants' duration estimate and the actual task duration, divided by the actual task duration). Left panel = participants in the prospective time estimation paradigm; right panel = participants in the retrospective time estimation paradigm. Error bars show the 95% confidence intervals of the mean.

**Weber's fraction variant results**. Sensitivity to time was measured using a variant of a Weber's Fraction (WF). A hallmark of time estimation is that the variability of people's time estimates increases with the length of the duration. However, this variability is typically

proportional to the amount of time being estimated (Wearden, 2016; Wearden & Lejeune, 2008). This is the scalar property of variance: The higher the WF, the higher the time estimation variability. WF was calculated by taking maximum ($Max_T$) and minimum ($Min_T$) time estimates from each participant, finding the difference and dividing this difference by the actual task duration ($T_D$).

$$WF = (Max_T - Min_T) / T_D$$

After providing their time estimate, participants were then asked to provide a maximum and a minimum time estimate. Because these estimates were used to derive the WF, they too were screened for null and outlying responses. All participants provided a minimum time estimate. However, two participants did not provide a maximum. Moreover, one participant provided "6 minutes" as their maximum after providing a minimum estimate of "55 minutes". As such, his response was removed from further analysis. Finally, three participants provided either a maximum or minimum estimate that exceeded three standard deviations the mean of their respective groups. Thus, their responses were also removed. Considering that WF is a measure of variability differences between groups, it was no surprise to find the assumption of homogeneity of variance was violated, $F(7,261) = 10.26$, $p < .001$.

No three-way interaction was found, $F(1,261) = .41$, $p = .52$. However, like RATIO, a two-way interaction was found between estimation paradigm and task duration, $F(1,261) = 7.16$, $p = .008$, $\eta_p^2 = .03$. Analysis of simple main effects revealed that estimation paradigm had a significant effect on WF in the eight-minute condition only, $F(1,261) = 10.36$, $p = .001$, $\eta_p^2 = .04$. Specifically, participants in the eight-minute condition had a higher WF if they were also in the prospective condition ($M = .80$, $SD = .57$, *95%* CI [.71,.90]) as opposed to the retrospective one ($M = .59$, $SD = .43$, *95%* CI [.49,.68]). Furthermore, there was a significant main effect of

duration, $F(1,261) = 77.56$, $p < .001$, $\eta_p^2 = .23$. The average WF of the eight-minute condition (*M* = .69, *SD* = .52, *95%* CI [.63,.80]) was higher than the 58-minute condition (*M* = .26, *SD* = .23, *95%* CI [.20,.33]). These results are available in Figure 8.

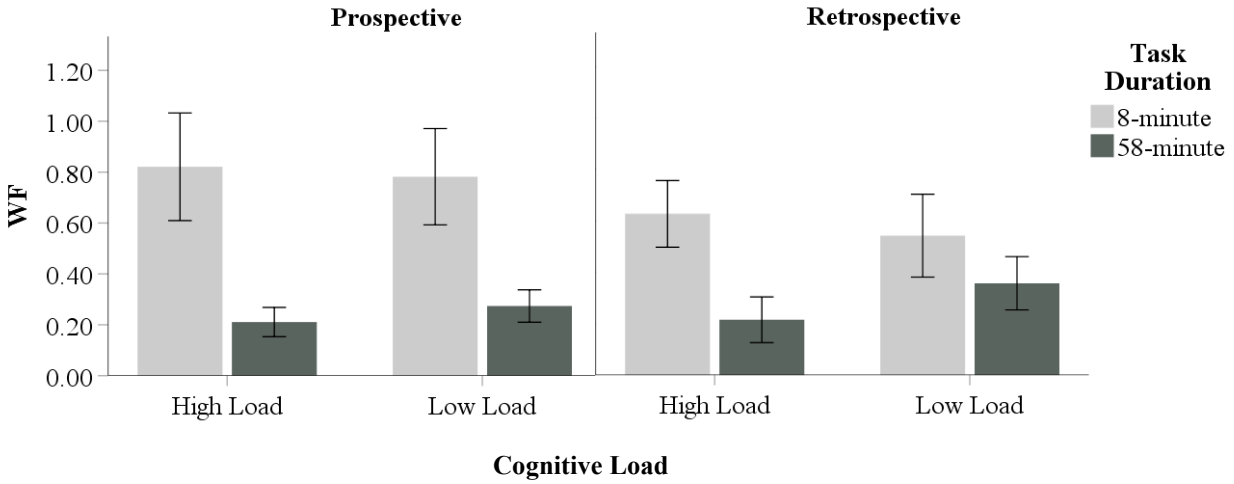


*Figure 8*. Weber's fraction variant (the difference between the participants' maximum and minimum time estimate over the actual task duration). Left panel = participants in the prospective time estimation paradigm; right panel = participants in the retrospective time estimation paradigm. Error bars show the 95% confidence intervals of the mean.

Like Tobin et al. (2010), the 58-minute task in the present study was perceived as proportionally shorter (as seen by the RATIO results), was less inaccurate (smaller ASE), and was less variable (smaller WF) compared to the shorter task. Finally, the WF results in this study and that of Tobin et al.'s study is opposite what would be expected of the scalar property of variance typically observed in very short tasks (Wearden, 2016). Considering longer durations should contain higher estimate variability, a second measure of variability was used to measure differences between groups and confirm this finding.

**The coefficient of variation (CV).** The CV for each condition was found by dividing the group's standard deviation by its mean. Overall, the average CV was 0.69 for the prospective

paradigm and 0.77 for the retrospective paradigm. Thus, the retrospective to prospective CV ratio was 1.11. In other words, retrospective estimates contained 10% more variability than prospective ones. This is slightly smaller than the CV ratio found in very short tasks (a ratio of 1.15; Block & Zakay, 1997). Yet, the difference in variability is in the same direction as shorter task (larger for retrospective estimates). Moreover, a variability difference was found between the high (CV = .66) and low (CV = .72) cognitive load conditions (CV ratio = 1.09). The low cognitive load condition contained 9% more estimation variability than the high load condition. Finally, the eight-minute to 58-minute CV ratio was 1.25. The eight-minute condition (CV = .40) thus contained 25% more estimation variability than the 58-minute one (CV = .32). Therefore, the scalar property of variance appears to break down when long intervals are used. This result confirmed the findings the from WF analysis.

## Exploratory Analyses

**Additional predictors of RATIO.** To explore whether boredom, engagement, and the feeling of time judgements (PoTJ) were related to participants' time estimates (RATIOs) and the independent variables under investigation (estimation paradigm, task duration, and cognitive load), correlations among all these variables were measured. The correlation matrix for the variables is given in Table 2. As expected, boredom and engagement were negatively correlated with each other. Additionally, the duration and the cognitive load of the task were correlated with how engaged or bored participants reported being. Next, engagement, boredom, and passage of time judgements were all significantly correlated with RATIO. They indicate that higher levels of boredom and the lower levels of engagement were associated with smaller time estimates. Finally, passage of time judgements were positively correlated with RATIO. Thus, when participants felt like time had gone by more quickly, they provided longer time estimates.

Table 2

*Correlations among the six variables of interest*

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1. RATIO | ---- | | | | | |
| 2. Estimation paradigm | .06 | ---- | | | | |
| 3. Cognitive Load | -.03 | .01 | ---- | | | |
| 4. Task Duration | -.59** | -.01 | .01 | ---- | | |
| 5. Engagement | .16* | -.04 | .25** | -.44** | ---- | |
| 6. Boredom | -.20* | .02 | -.23** | .44** | -.56** | ---- |
| 7. PoTJ | .27** | -.06 | .17* | -.47** | .44** | -.51** |

*Note.* * $p < .01$. ** $p < .001$.

To further examine the relationships among the variables, a hierarchical regression was performed. The predictors estimation paradigm, task duration, and cognitive load were added in a first step because they were the main variables of interest. A second step added engagement, boredom, and passage of time judgements (PoTJ) to the regression. One person did not provide a boredom rating and was thus excluded from the exploratory analysis. The regression results pertaining to the effects of these variables on RATIO are in Table 3.

Table 3

*Linear model of predictors of RATIO*

| | ***b* (95% CI)** | ***SE*** | **β** | ***t*** | ***p*** |
|---|---|---|---|---|---|
| **Step 1** | | | | | |
| Constant | 1.29 (1.20, 1.39) | .05 | | 26.05 | < .001 |
| Est. Paradigm | .06 (-.04, .16) | .05 | .06 | 1.12 | = .27 |
| Cognitive Load | -.02 (-.12, .08) | .05 | -.02 | -.45 | = .67 |
| Task Duration | -.60 (-.70, -.50) | .05 | -.59 | -11.83 | = .001 |
| **Step 2** | | | | | |
| Constant | 1.42 (1.06, 1.77) | .18 | | 7.91 | < .001 |
| Est. Paradigm | .05 (-.05, .15) | .05 | .05 | 1.03 | = .30 |
| Cognitive Load | .01 (-.09, .12) | .05 | .01 | .21 | = .84 |
| Task Duration | -.66 (-.78, -.53) | .06 | -.64 | -10.61 | < .001 |
| Engagement | -.04 (-.07, .002) | .02 | -.12 | -1.85 | = .07 |

| | | | | | |
|---|---|---|---|---|---|
| Boredom | .01 (-.03, .05) | .02 | .04 | .60 | = .55 |
| PoTJ | .01 (-.03, .05) | .02 | .04 | .58 | = .56 |

*Note.* $R^2$ = .35 for Step 1; $\Delta R^2$ = .01 for Step 2 ($p$ =.16).

Despite having significant correlations with RATIO, adding engagement, boredom and PoTJs to the model did not predict RATIO results over and above estimate paradigm, task duration, and cognitive load. In fact, the duration of the task was the most significant predictor, with little else explaining variation in RATIO.

**Exploring passage of time judgements (PoTJs)**. In the previous regression analysis, people's PoTJs were analysed as a predictor. Although these judgements were positively correlated with the outcome (RATIO), they did not explain a significant additional amount of variance in RATIO. PoTJs may also be viewed as an outcome variable rather than an independent one. Perhaps the speed at which time seems to pass does not influence the time estimates provided, but instead, time estimate paradigm, task duration, cognitive load, boredom, or engagement influence how fast or slow time seems to pass. A second hierarchical multiple regression was therefore conducted. The first three predictors entered were paradigm, duration, and cognitive load. The second step included boredom and engagement.

As seen in Table 4, both the duration of the task and the cognitive load of the task significantly explained variation in PoTJs in the first step of the regression. In model two, however, only the task's duration significantly explained PoTJ variance in addition to engagement and boredom. Cognitive load's b-coefficient was reduced by more than half when engagement and boredom were added to the model. Thus, time felt like it passed faster or slower depending on the load of the task, but this was perhaps because cognitive load relates to how boring or engaging a task was.

Table 4

*Linear model of predictors of PoTJs*

| | ***b* (95% CI)** | ***SE*** | **β** | ***t*** | ***p*** |
|---|---|---|---|---|---|
| **Step 1** | | | | | |
| Constant | 4.47 (4.13, 4.80) | .17 | | 26.9 | < .001 |
| Est. Paradigm | -.21 (-.55, .13) | .17 | -.07 | -1.23 | = .22 |
| Cognitive Load | .56 (.22, .90) | .17 | .17 | 3.23 | = .001 |
| Task Duration | -1.56 (-1.90, -1.22) | .17 | -.48 | -9.06 | < .001 |
| **Step 2** | | | | | |
| Constant | 4.58 (3.60, 5.57) | .50 | | 9.13 | < .001 |
| Est. Paradigm | -.16 (-.48, .16) | .16 | -.05 | -1.01 | = .31 |
| Cognitive Load | .25 (.09, .58) | .17 | .08 | 1.47 | = .14 |
| Task Duration | -.95 (-1.32, -.58) | .19 | -.29 | -5.01 | < .001 |
| Engagement | .12 (.01, .24) | .06 | .13 | 2.08 | = .04 |
| Boredom | -.25 (-.36, -.14) | .06 | -.29 | -4.56 | < .001 |

*Note.* $R^2 = .25$ for Step 1; $\Delta R^2 = .098$ for Step 2 ($p < .001$).

With the observation of cognitive load's b-coefficient change between steps, the possibility emerged that the effect of cognitive load on PoTJs might be mediated by task engagement and boredom. This hypothesis was tested with Lambert's mediation model using Hayes PROCESS tool in SPSS (Hayes, 2012). Bootstrapped 95% confidence intervals were obtained and are denoted by square brackets. Figure 9 shows the mediation model. When engagement and boredom were considered in the model, the direct effect of cognitive load on PoTJs became non-significant. That is, the b-coefficient of the direct effect became one-fifth the size of that seen in the total effected. Cognitive load indirectly effected PoTJs through engagement, $b = .16$ [.06, .33] and through boredom, $b = .25$ [.10, -.46]. Sobel tests indicated that engagement caused a significant reduction in the effect of cognitive load on PoTJs, $z = .253$ ($SE = .06$), $p = .01$, as did boredom, $z = 2.88$ ($SE = .09$), $p = .004$.

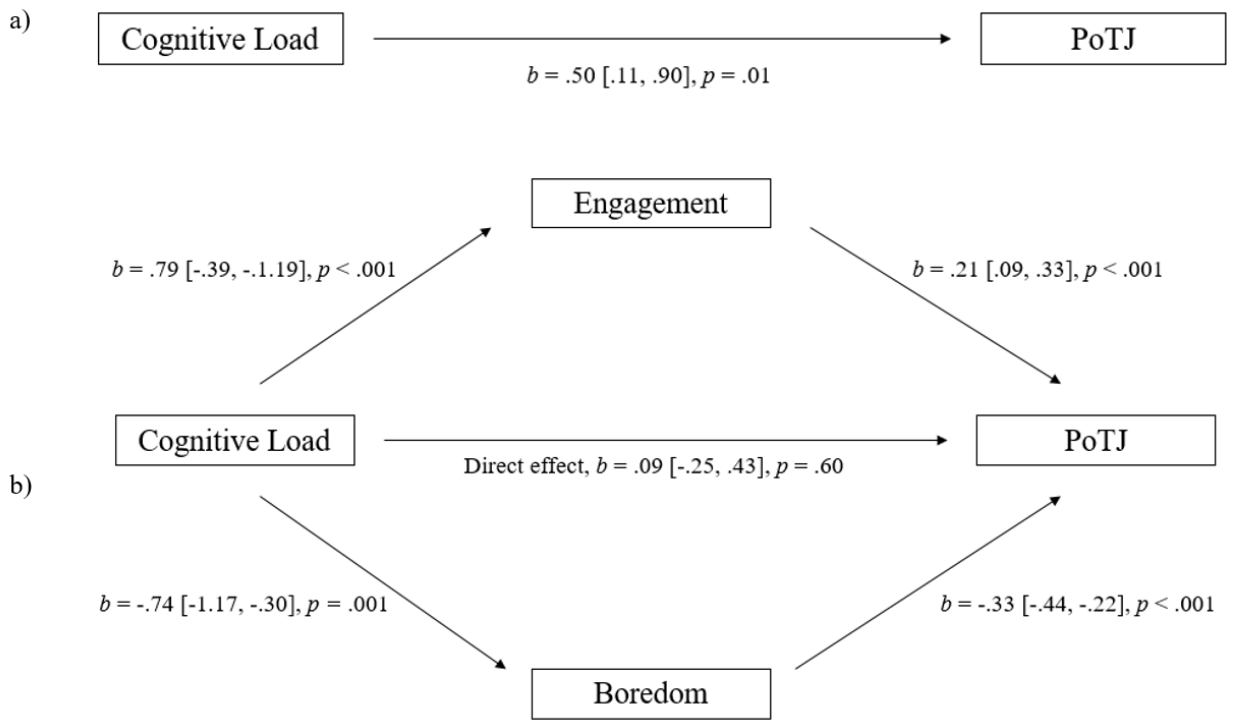


*Figure 9*. Complete mediation model: a) the total effect of cognitive load on passage of time judgements; b) the indirect effect of cognitive load on passage of time judgements through engagement and boredom.

## Discussion

The goal of this research was to consider the impact of the time estimate paradigm, task duration, and cognitive load on people's perception of time. Block et al. (2010) had reported a strong interaction between estimation paradigm and cognitive load, but, most of the literature on human time perception focuses on durations ranging from a few hundred milliseconds to a few minutes (Mattew & Meck, 2014). Thus, this experiment sought to replicate Block et al.'s (2010) findings for a long (i.e., eight-minute) and a very long (i.e., 58-minute) duration. Importantly, this experiment also directly assessed the amount of cognitive effort demanded of a task. Past research has either not considered the impact of cognitive load at long durations (e.g., Tobin et al. 2010) or has failed to quantify it (e.g., Tobin & Grondin, 2009). To achieve these goals, the present experiment used consistent and varied stimulus-response mappings in a visual and

memory search task (Schneider & Shiffrin, 1997). Under the consistent mapping condition, the same stimuli served as targets on all trials whereas under the varied mapping condition, each stimulus in turn served as either a target or distracter.

### The Effect of Estimation Paradigm

According to the Attentional Gate Model (AGM) of time perception, paying attention to time increases the amount of information used by the mind's internal clock to derive a time estimate (Zakay & Block, 1997). Because participants presumably payed more attention to time in the prospective estimation paradigm, these estimates were expected to be greater on average compared to the retrospective paradigm.

On average, participants in this experiment were very accurate at estimating their task's duration. The mean judgement ratio of all the estimates combined was 1.05. However, differences between estimation paradigms were found in the eight-minute task. Although participants overestimated the eight-minute task by over 33% in both estimation paradigms, overestimation was greater when estimates were made in the prospective paradigm. Thus, the effect of time estimation paradigm in the eight-minute task conformed to the AGM's predictions. Increased attention to time in this condition potentially allowed for more internal clock time units called *pulses* to be counted by an accumulator in the brain. Without directing attention toward time in the retrospective condition, fewer pulses were accumulated by the brain to derive a time estimate. Consequently, during the eight-minute task, participants in the prospective paradigm had accumulated more pulses than those in the retrospective paradigm, resulting in greater estimates. Participants in the eight-minute prospective condition thus overestimated of the task's duration much more than those in the eight-minute retrospective condition.

Alternatively, paradigm differences in the eight-minute task could also be explained using

a *contextual change hypothesis*. Rather than time estimates being derived from an accumulation of pulses, greater overestimation in the prospective paradigm might have been due to bias. This contextual change hypothesis claims that changes in context, such as changes in stimulus complexity or sensory modality, can act as events in memory. These events are encoded and act as cues for how much time has passed (Block & Zakay, 2008). Instructing participants to keep track of time could have changed the threshold for what was considered a cue. Perhaps smaller or less detailed environmental changes were stored in memory and thus considered when estimating the tasks' duration. With a more liberal threshold, participants in the prospective condition retrieved of a greater number of contextual cues when providing their time judgement. This led to prospective estimates being greater than retrospective ones in the eight-minute task.

This paradigm effect had been predicted to be larger in the 58-minute task compared to the eight-minute task when cognitive load was low. However, this difference was not found. Instead, participants underestimated the 58-minute task duration, and they did so to an equal degree in each estimation paradigm. The lack of paradigm effect at 58 minutes is incompatible with both the AGM and the contextual change hypothesis. Not only should there have been more pulses accumulated according to the AGM but instructing participants to keep track of time should have directed their attention toward smaller or less detailed contextual changes according to the contextual change framework. Attention toward contextual changes was expected to prompt the retrieval of a greater number of cues at the end of the task and thus produce longer prospective estimates on average. Yet, this did not occur.

This result thus does not replicate the significant overestimation effect found in Tobin et al.'s (2010) 58-minute, prospective condition. However, it does replicate Bakan's (1955) original finding that there is no estimation paradigm effect in a task almost one hour long. As such,

findings from the eight-minute and 58-minute conditions in this experiment suggest that the length of the task determines which underlying mechanisms produce a time estimate. It is possible that compared to incidental encoding, intentional encoding of temporal information resulted in pronounced overestimation at eight-minutes only. Given how few experiments have been conducted to date comparing prospective and retrospective estimates at very long durations, more research is needed to confirm whether this is true. However, the present experiment provides additional evidence against the existence of a paradigm effect in a task of 58 minutes.

As a cautionary note, a main effect of estimation paradigm was observed when absolute standard errors were analyzed. That is, prospective estimates in this experiment were further away from the actual task's duration on average compared to retrospective estimates. Thus, intentional encoding of time appears to have made estimates less accurate in this experiment. Although this main effect might suggest the possibility of an overall paradigm difference in estimation accuracy, even at 58-minutes, the effect size was small. Future research should attempt to replicate this effect.

Furthermore, in the eight-minute task, prospective estimates in the present experiment produced less variability than retrospective ones. That is, participants in the eight-minute task had higher Weber's fractions if they were in the prospective condition. This is opposite of what is typically found in the literature (Tobin et al. 2010). However, no main effect of estimation paradigm was found, suggesting that prospective estimates were no more variable overall than retrospective ones. Nonetheless, the coefficients of variations do suggest some difference in variability in the typical direction. Retrospective estimates contained 10% more overall variability. This finding replicates the result of Block and Zakay's (1997) meta-analysis. Taken

together, prospective estimates were more variable at eight compared to 58-minutes, but retrospective estimates displayed more variability overall.

## The Effect of Task Duration

In this experiment, underestimation occurred in the 58-minute condition and overestimation occurred in the eight-minute condition. What is surprising about this result is that participants in the 58-minute condition not only significantly underestimated the task's duration, but participants' responses displayed little variability. This finding replicates Tobin et al. (2010) who had also found little variability in their 58-minute condition compared to their two other durations (12 and 36 minutes). Thus, the scalar property of variance may break down when estimating a task of 58-minutes. Importantly, Matthew and Meck (2014) attributed underestimation to an increased response uncertainty at longer durations. The present variability data challenges this hypothesis. Furthermore, the pattern of overestimation and underestimation in this experiment suggests that even at very long durations, estimates conform to Vierordt's law. Yet, the question remains: what is causing this effect?

Roy and Christenfeld (2008) found that Vierordt's law predicted performance both when time estimates were made after completing a task and when estimates were made for future tasks. The latter is called a predictive time estimate, which is made prior to the experience of a target duration. Predictive estimates should not be confused with prospective estimates, which occur after a target duration is experienced. Predictions are typically made based on experience and knowledge of a task. Thus, drawing on memory is equally essential for predictive estimation and retrospective estimation.

Due to the presence of Vierordt's law in both predictive and retrospective estimation paradigms, Roy, Christenfeld, and McKenzie (2005) concluded that that a bias in prediction is

caused by a bias in memory. For example, a well-known memory bias of predictive estimation is the *planning fallacy*, first introduced by Kahneman and Tversky (1979). This fallacy refers to the tendency for people to underestimate how long a project will take to complete, even though they know that past similar projects have taken longer to complete than planned (Buehler, Griffin, & Peetz, 2010). Rather than memory for prior similar tasks being ignored while generating predictions, previous task information might be instead misremembered (Roy et al., 2005). Similarly, memory for a task that just transpired might be misremembered. For instance, if a person sets aside 45 minutes to complete a task that previously took them one hour to complete, they have engaged in the planning fallacy. Likewise, participants in the present experiment estimated the 58-minute task to take 42.5 minutes on average. Indeed, the duration of the task was misremembered.

A key similarity between planning and retrospectively estimating a task's duration is what is expected to happen when the interval is divided into smaller segments. Krueger and Evens (2004) found that unpacking a task into segments lead to longer and less biased predictions. This coincides nicely with past research finding that divided intervals are overestimated in retrospect (Lim, Kum, & Lee, 2015). Dividing the present task into subcomponents such as introducing short breaks, would likely have led to longer estimates. However, the present experiment did not manipulate the extent to which the task was segmented. Instead, underestimation of the 58-minute task occurred regardless of estimation paradigm and regardless of cognitive load. This is perhaps due to anchoring or compensation effects.

**Anchoring effects.** In the present experiment, participants performed practice trials for two minutes to become familiar with the task before they began. Thomas and Handley (2008) found that performing an identical but shorter task (i.e., an anchor) before the main task led to its

underestimation when participants were not told how long the practice trials would be. Conversely, when the anchor is longer than the main task, overestimation occurs. They reported that anchoring effects are also observed when the task instructions indicate the range of possible durations. Thus, the 90-minute cap provided by the experimental instructions could have lengthened time estimates.

Whether or not participants focussed on the shorter or longer anchor might have depended on which condition they were in. Participants in the eight-minute condition might have focussed on the 90-minute anchor. Perhaps they expected the task to be longer than it felt because they were provided with such a large cap. The expectation that they would be performing a longer task might have led them to overestimate the tasks' duration. In fact, anchoring has been outlined as an important contributor to overestimation a in prior research using long durations. Overestimation was observed when elementary school students performed an eight-minute video game or reading task (Tobin & Grondin, 2009). In that experiment, students were informed that the tasks would be performed within the confines of a regular class period. Likewise, overestimation occurred in a 12-minute video game condition when participants were told that the game could last a maximum of three hours (Tobin et al., 2010). It was concluded that this three-hour cap could have created an upward shift in time estimates.

Similarly, underestimation in the present experiment's 58-minute condition could be due to anchoring. Rather than focusing on the 90-minute cap, participants in this condition might have instead referenced their practice trials. These practice trials perhaps became a more salient reference in time as the tasks' duration increased. That is, in a task of this length, participants might have relied heavily on their memory for events occurring throughout the experiment to make a time judgement. Contextual changes are thought to be more salient toward the beginning

of an experiment (Hintzman et al., 1973), thus reliance on memory might have pulled estimates toward the practice trial anchor, leading to underestimation.

**Compensation effects.** Another explanation for the effect of task duration in the present experiment is that participants simply compensated for their level of boredom or engagement during the task. Indeed, higher boredom and lower engagement ratings led to shorter the time estimates. This finding is strange considering the well-known saying, "time flies when you're having fun." Instead, the present results imply that time flew when the participants were bored. Yet, PoTJs were negatively correlated with boredom and positively correlated with engagement. Therefore, participants who were bored in the present experiment felt like time passed slowly, but then provided a shorter time estimate rather than a longer one.

To explain this result, the participants might have overcompensated for the "slow" feeling they experienced. Possibly, they thought, "wow, time went by slowly. It felt like an hour, but it was probably only around 40 minutes." Then, they proceeded to provide the 40-minute estimate and thus underestimated the length of the 58-minute task. The opposite occurred for the situation in which time felt like it passed quickly (i.e., the eight-minute condition). In fact, this replicates prior research finding that the relative feeling of time passing does not always correspond with the accuracy of a time estimate. That is, although their time estimates may be accurate, participants might indicate that time passed slowly or quickly during a task (Droit-Volet & Wearden, 2016).

## The Effect of Cognitive Load

Not only could engagement and boredom ratings have influence whether the task was overestimated or underestimated in the present experiment, but these variables were related to cognitive load as well. Specifically, cognitive load impacted how fast or slow time felt like it

passed throughout the task, but this effect was completely mediated through the participants' level of engagement and boredom. In other words, when the cognitive load of the task was high, the task was more engaging (and less boring), and time felt like it passed quickly. When cognitive load was low, the task was less engaging (and more boring), and time felt like it passed slowly. This cognitive load effect was specific to PoTJs, however. Cognitive load had no effect on the actual time estimates provided. Once again, this indicates that time estimates and PoTJs are not the same construct (see Wearden, 2015 for a recent discussion).

Despite the null effect of cognitive load on time estimation accuracy, cognitive load was successfully manipulated using consistent and varied stimulus-response mapping. The task was also easier to perform after 58-minutes compared to eight minutes as predicted. As such, cognitive load should have had its strongest influence on time estimation accuracy at eight-minutes. However, contrary to expectation, this manipulation did not interact with task duration nor did it interact with estimation paradigm on participants' time estimates.

To account for results specifically in the retrospective paradigm, it is possible that heightened cognitive load in the present experiment did not result in a greater number of contextual cues. After all, it is not cognitive load per se that influences retrospective estimation, but the number of memorable events, or segments, that occur during the task (Wearden, 2016). Specifically, although the high load task contained changes in stimulus arrangement, it did not result in the perception of each trial as a segment. The high load condition in the present experiment was just as memorable (contained the same number of contextual cues or segments) as the low load condition. Thus, either task was judged to be the same length in retrospect. The lack of cognitive load effect for prospective paradigm is surprising, however. Higher task demands typically result in underestimation, and this has been reported in a wide range of

research (Brown, 1997). Another underlying mechanism was thus considered as potentially being involved in keeping track of time.

**Prospective memory.** Rather than relying on contextual cues in the prospective paradigm, estimates might have relied on another long-term memory process. Even though people were told to keep track of time throughout this condition, it may not have been possible for them to constantly pay attention to temporal information for close to one hour. Actively timing intervals of this length might instead have required *prospective memory*. Prospective memory is memory for future intentions (Henry, McLeod, Phillips, & Crawford, 2004) such as remembering to pick up an item on the way home. Einstein and McDaniel (2005) argued that there is limited capacity for how much conscious control a person has over their behaviour. Many thoughts and behaviors are triggered automatically by cues or reminders. They highlighted how prospective memory can be time-based such as remembering to press a key after 10 minutes has elapsed. Indeed, participants in the present experiment were told to "keep track of time" and to provide a time estimate at the end of the task. Given that there were no cues to keep track of time throughout the task, participants had to rely on spontaneously and continuously remembering to do so. In fact, prospective memory can rely heavily on spontaneous retrieval rather than concrete cues (Einstein & McDaniel, 2005). As such, participants in this experiment were probably not keeping track of time continuously throughout the task. Rather, they were periodically remembering to encode time.

Previous evidence suggests that prospective memory performance is not influenced by task demands. This was found in an experiment conducted by Kliegel, Martin, McDaniel, and Einstein (2004). Participants in this experiment performed an ongoing word rating task while pressing a button whenever the saw a specific word. The demands of the word-rating task were

manipulated by having half the participants simultaneously perform a third auditory digit detection task. Thus, all participants were instructed to rate words while pressing a button to cued words, but only half of the participants also pulled a lever whenever they heard the number "9" in the auditory task. Regardless of whether participants also performed this third task, they equally complied with the button pressing instructions for the cued words. Thus, prospective memory for button pressing was not influenced by task difficulty.

These results are quite different from what is typically observed in working memory tasks. Working memory is the cognitive capacity to store, process, and manipulate information in the short term (Baddeley, 2012). This is likely the memory system used when performing tasks under two minutes, which is typical of time perception research. During short intervals, participants are not engaged in the task for long enough to off-load time-keeping into long-term memory. As such, the non-temporal task can draw attention away from accurate time-keeping. This leads to the typical underestimation effect seen in prospective estimation paradigms using short durations. In other words, a secondary task can compete for available attentional resources required of keeping track of time when both tasks are being processed using the same cognitive mechanisms (i.e., working memory).

In fact, Fortin et al. (1993) argued that it is not the difficulty of the non-temporal task that interferes with timing processes. Rather, timing performance is interrupted specifically by non-temporal tasks that involve short-term memory processing. Tasks that demand more short-term memory storage interrupt timing performance compared to tasks that do not demand much storage. In the present experiment, the varied mapping task was clearly more demanding on short-term memory than the consistent mapping task, as demonstrated by visual and memory

search performance. Yet, timing performance did not differ between the two conditions. The tasks were thus performed using separate processes and no resource allocation was necessary.

Indeed, attention-allocation models predict a bi-directional relationship between the time-keeping task (the prospective time estimation paradigm) and the concurrent task (Brown, 1997). This is also known as *dual-task interference* (Brown, Collier, & Night, 2013). The present RT findings are in the opposite direction of what would be expected on dual-task interference. Retrospective high load responses were slower than prospective high load responses, but only in the eight-minute condition. Perhaps participants in the latter condition simply responded quickly at the beginning of the task to ensure that they had enough time to keep track of time. As the task progressed (they ended up being in the 58-minute task), they slowed down and become more comfortable doing both the time keeping and the visual and memory search task simultaneously. Evidently, time-keeping and the visual and memory search task did not share common resources in this experiment. Prospective time-keeping for long or very long durations instead relied on long-term, prospective memory mechanisms, rather than the short-term memory system demanded of the visual and memory search task.

## Limitations and Future Directions

The visual and memory search task chosen for the present experiment had two main advantages. First, it allowed for maximum control over the difficulty of the experiment so that the impact of cognitive load could be carefully examined. Second, performance indices for visual and memory search are well-known and could therefore be used to determine whether participants were attentive to the task throughout the experiment. However, using a controlled task like visual and memory search does have its limitations. As outlined by Tobin et al. (2010), few experimental conditions in the literature display high ecological validity. That is, the tasks

sometimes do not reflect the types of activities that people perform daily. This observation applies to the present experiment as well. As such, the present results might not generalize beyond the type of activities that require repetitive and automatic stimulus-response processing. Although repetitive tasks do occur in daily life (e.g., some data entry tasks, scanning and labelling items in a retail environment, and repetition in video games), many tasks feature a mix of repetition, complexity, and novelty. These task characteristics might influence people's perception of time.

Nonetheless, the present experiment suggests that cognitive load does not influence time estimation accuracy in a repetitive task. Thus, rather than continuing to focus on cognitive load, future research might examine the effects of task engagement on time estimation. Past research has examined *enjoyment* levels on time estimation and found no significant relationship (Tobin et al. 2010). Considering engagement was positively correlated with the RATIO results in the present experiment, it is possible that enjoyment and engagement are two different concepts and contribute to time estimation differently. Moreover, task engagement was found to affect PoTJs in the present experiment. However, only one question assessed the participants' subjective level of engagement during the task. Furthermore, no a priori predictions were made regarding engagement level in this experiment. The exploratory results presented in this thesis must therefore be followed up with more precise measures and hypotheses.

In addition to the above considerations, there is also a paucity of research investigating the effects of different anchors on time estimation at long and very long durations. Thus, there were few studies on which to base an appropriate maximum anchor. For example, Tobin et al. (2010) recruited participants who were willing to play a video game for potentially three hours. Due to the repetitive nature of the visual and memory search task in the present experiment, a

three-hour maximum was deemed to be too large. A 90-minutes cap was chosen for this experiment to allow for variation while attempting to minimize an anchoring effect. Despite choosing a smaller maximum, significant overestimation still occurred in the eight-minute task. Future research could systematically vary the maximum and minimum anchors provided to participants. This would help determine how influential these anchors are in varying time estimates for a task of this length.

Moreover, the length of the task remains an important topic of future investigation. This is only the third known experiment to examine time estimation for a task close to one hour. As such, predictions for the present experiment were limited in that they were mainly based on literature that focussed on short durations (Block et al., 2010) or on literature that used very long duration in uncontrolled environments (Tobin et al., 2010). The present experiment should thus be the first of a series of experiment replicating and extending on this research. For example, the present experiment highlighted the need to operationalize the term “long” and when referring to task durations. Because an effect of estimation paradigm was found in the eight-minute condition but not in the 58-minute condition, a promising avenue of future research is in determining the point at which there is no longer a paradigm effect. That is, at what point do prospective and retrospective estimates converge? Could this be the point at which a duration should be considered “long”? To answer these questions, future experiments could systematically vary the amount of time people perform the same task. Therefore, rather than having two duration conditions like the present experiment, several durations can be examined within one experiment.

## Conclusion

This research examined the impact of the time estimate paradigm (prospective vs. retrospective), task duration (eight vs. 58 minutes), and cognitive load (high vs. low) on people's perception of time while also exploring the impact of engagement, boredom, and the speed at which time felt to have passed throughout a task. Participants in this experiment were very accurate at estimating the task's duration. However, differences were found between estimation paradigms in the eight-minute task where participants in the prospective paradigm overestimated the task to a greater extent than those in the retrospective paradigm. Although this replicated past research showing that prospective estimates tend to be greater on average than retrospective estimates (Block & Zakay, 1997), this result did not extend to the 58-minute task. Thus, while the eight-minute condition is compatible with both the AGM and the contextual change hypothesis proposed to explain differences between paradigms, the 58-minute condition is compatible with neither.

Furthermore, the present experiment found that Vierordt's law extended to the 58-minute task, replicating past research finding that this law occurs at long durations (Tobin & Grondin, 2009). Finally, cognitive load had no effect on time estimation accuracy, although it did effect passage of time judgements. This latter result was found to be due to differing levels of engagement and boredom in the high and low load conditions. When focussing on time estimation accuracy, however, the well-known cross-over interaction between cognitive load and time judgment condition (Block et al. 2010) does not seem to extend to longer durations.

High cognitive load in the present experiment probably did not lead to more available contextual cues. This would explain the lack of load effect in the retrospective paradigm. However, the lack of cognitive load effect in the prospective paradigm could signify the

existence of underlying cognitive mechanisms for time perception at long durations, not yet considered in the literature. Prospective memory might be more important for deriving time estimates at eight and 58-minutes than the working memory mechanisms thought to be involved for timing shorter durations. Future research should continue investigating the effect of task duration while directly comparing paradigms. This would further contribute to the field's understanding of the cognitive mechanisms involved in human time perception in long and very long tasks.

References

Avni-Babad, D., & Ritov, I. (2003). Routine and the perception of time. *Journal of Experimental Psychology: General*, *132*(4), 543-550.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438-456.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual review of psychology*, *63*, 1-29.

Bakan, P. (1955). Effect of set and work speed on time estimation. *Perceptual and Motor Skills*, *5*, 147-148.

Baudouin, A., Vanneste, S., Isingrini, M., & Pouthas, V. (2006). Differential involvement of internal clock and working memory in the production and reproduction of duration: A study on older adults. *Acta Psychologica*, *121*, 285-296.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *The American journal of psychology*, 181-199.

Bisson, N., Tobin, S., & Grondin, S. (2012). Prospective and retrospective time estimates of children: a comparison based on ecological tasks. *PLoS One*, *7*(3), e33049.

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

Block, R. A. (2003). Psychological timing without a timer: The roles of attention and memory. *Time and Mind II: Information Processing Perspectives*, 41-59.

Block, R. A., George, E. J., & Reed, M. A. (1980). A watched pot sometimes boils: A study of duration experience. *Acta Psychologica*, *46*(2), 81-94.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta psychologica*, *134*(3), 330-343.

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.

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, *4*(2), 184-197.

Block, R. A., & Zakay, D. (2008). *Timing and remembering the past, the present, and the future*. In S. Grondin (Ed.), Psychology of time. Bingley, UK: Emerald Group Publishing, 367–394.

Block, R. A., Zakay, D., & Hancock, P. A. (1998). Human aging and duration judgments: A meta-analytic review. *Psychology and Aging*, *13*(4), 584-596.

Boltz, M. (2005). Duration judgments of naturalistic events in the auditory and visual modalities. *Perception & Psychophysics*, *67*, 1362-1375.

Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics*, *38*(2), 115-124.

Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Perception & psychophysics*, *59*(7), 1118-1140.

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

Brown, S., & Bennett, E. (2002). The role of practice and automaticity in temporal and nontemporal dual-task performance. *Psychological Research*, *66*, 80-89.

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.

Brown, S. W., & Merchant, S. M. (2007). Processing resources in timing and sequencing tasks. *Perception & psychophysics*, *69*(3), 439-449.

Brown, S. W., & Stubbs, D. A. (1992). Attention and interference in prospective and retrospective timing. *Perception*, *21*, 545-557.

Buehler, R., Griffin, D., & Peetz, J. (2010). The planning fallacy: Cognitive, motivational, and social origins. *Advances in experimental social psychology*, 43, 1-62.

Buhusi, C. V., & Meck, W. H. (2006). Interval timing with gaps and distracters: evaluation of the ambiguity, switch, and time-sharing hypotheses. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*(3), 329-338.

Burle, B., & Casini, L. (2001). Dissociation between activation and attention effects in time estimation: implications for internal clock models. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 195-205.

Cahoon, D., & Edmonds, E. M. (1980). The watched pot still won't boil: Expectancy as a variable in estimating the passage of time. *Bulletin of the Psychonomic Society*, *16*, 115-116.

Casini, L., & Macar, F. (1997). Effects of attention manipulation on judgments of duration and of intensity in the visual modality. *Memory & Cognition*, *25*(6), 812-818.

Clément, A., & Droit-Volet, S. (2006). Counting in a time discrimination task in children and adults. *Behavioural Processes*, *71*, 164-171.

Cousineau, D., & Larochelle S. (2004). Visual-memory search: An integrative perspective. *Psychological Research, 69,* 77-105.

Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17*(4), 327-352.

Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992). Automatization and training in visual search. *The American journal of psychology*, 271-315.

Droit-Volet, S., Trahanias, P., & Maniadakis, M. (2017). Passage of time judgments in everyday life are not related to duration judgments except for long durations of several minutes. *Acta Psychologica*, *173*, 116-121.

Droit-Volet, S., & Wearden, J. (2016). Passage of time judgments are not duration judgments: evidence from a study using experience sampling methodology. *Frontiers in psychology*, *7*.

Droit-Volet, S., Wearden, J. H., & Zélanti, P. S. (2015). Cognitive abilities required in time judgment depending on the temporal tasks used: a comparison of children and adults. *The Quarterly Journal of Experimental Psychology*, *68*(11), 2216-2242.

Duzcu, H., & Hohenberger, A. (2014). Prospective duration judgments: the role of temporality and executive demands of concurrent tasks. *Acta psychologica*, *147*, 34-41.

Einstein, G. O., & McDaniel, M. A. (2005). Prospective memory: Multiple retrieval processes. *Current Directions in Psychological Science*, *14*(6), 286-290.

Eisler H, Eilser AD, Hellstrom A. (2008). *Psychophysical issues in the study of time perception*. In: S. Grondin (Ed.), Psychology of time. Bingley, UK: Emerald Group Publishing, 75–109.

Fetterman, J. G., Killeen, P. R., & Hall, S. (1998). Watching the clock. *Behavioural processes*, *44*, 211-224.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Fortin, C. (1999). Short-term memory in time interval production. *International Journal of Psychology*, 34, 308–316.

Fortin, C., & Rousseau, R. (1987). Time estimation as an index of processing demand in memory search. *Perception & Psychophysics*, *42*(4), 377-382.

Fortin, C., Rousseau, R., Bourque, P., & Kirouac, E. (1993). Time estimation and concurrent nontemporal processing: Specific interference from short-term-memory demands. *Perception & Psychophysics*, *53*(5), 536-548.

Franssen, V., Vandierendonck, A., & Van Hiel, A. (2006). Duration estimation and the phonological loop: Articulatory suppression and irrelevant sounds. *Psychological research*, *70*, 304-316.

French, R. M., Addyman, C., Mareschal, D., & Thomas, E. (2014). Unifying prospective and retrospective interval-time estimation: A fading-Gaussian activation-based model of interval-timing. *Procedia-Social and Behavioral Sciences*, *126*, 141-150.

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

Grondin, S., & Killeen, P. R. (2009). Tracking time with song and count: Different Weber functions for musicians and nonmusicians. *Attention, Perception, & Psychophysics*, *71*(7), 1649-1654.

Grondin, S., & Laflamme, V. (2015). Stevens's law for time: A direct comparison of prospective and retrospective judgments. *Attention, Perception, & Psychophysics*, *77*, 1044-1051.

Grondin, S., & Macar, F. (1992). *Dividing attention between temporal and nontemporal tasks: A performance operating characteristic-POC-analysis*. In Time, action and cognition. Dordrecht, Netherlands: Kluwer Academic Publishers, 119-128.

Grondin, S., Ouellet, B., & Roussel, M. E. (2004). Benefits and limits of explicit counting for discriminating temporal intervals. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *58*, 1-12.

Grondin, S., & Plourde, M. (2007). Judging multi-minute intervals retrospectively. *Quarterly Journal of Experimental Psychology*, *60*(9), 1303-1312.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.

Hélie, S., & Cousineau, D. (2011). The cognitive neuroscience of automaticity: behavioral and brain signatures. *Cognitive sciences*, *6*, 35.

Hemmes, N. S., Brown, B. L., & Kladopoulos, C. N. (2004). Time perception with and without a concurrent nontemporal task. *Perception & Psychophysics*, *66*(2), 328-341.

Henry, J. D., MacLeod, M. S., Phillips, L. H., & Crawford, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and aging*, *19*, 27-39.

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*, 719-730.

Hintzman, D. L., Block, R. A., & Summers, J. J. (1973). Contextual associations and memory for serial position. *Journal of Experimental Psychology*, *97*(2), 220-229.

Hornstein, A. D., & Rotter, G. S. (1969). Research methodology in temporal perception. *Journal of Experimental Psychology*, *79*, 561-564.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*, 269-299.

Kahneman, D., & Tversky, A. (1979). Intuitive prediction: Biases and corrective procedures. *TIMS Studies in Management Science, 12,* 313–327.

Kikkawa, M. (1983). The effect of distribution of attention upon prospective and retrospective estimations of long temporal intervals. *Journal of Child Development, 19,* 14-22.

Klapproth, F. (2007). Time perception, estimation paradigm, and temporal relevance. *Perceptual and Motor Skills*, *104*, 749-757.

Kliegel, M., Martin, M., McDaniel, M., & Einstein, G. (2004). Importance effects on performance in event-based prospective memory tasks. *Memory*, *12*(5), 553-561.

Kurtz, R. M., & Strube, M. J. (2003). Hypnosis, attention, and time cognition. *International Journal of Clinical and Experimental Hypnosis*, *51*(4), 400-413.

Lefebvre, C., Cousineau, D., & Larochelle, S. (2008). Does training under consistent mapping conditions lead to automatic attention attraction to targets in search tasks? *Perception & psychophysics*, *70*(8), 1401-1415.

Lejeune, H. (2000). Prospective timing, attention and the switch: A response to 'Gating or switching? Gating is a better model of prospective timing' by Zakay. *Behavioural Processes*, *52*, 71-76.

Lim, E. A. C., Kum, D., & Lee, Y. H. (2015). Understanding how changes within service experiences impact prospective vs. retrospective time judgments. *Journal of the Academy of Marketing Science*, *43*(6), 730-745.

Logan, G. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492-527.

Matthews, W. J., & Meck, W. H. (2014). Time perception: the bad news and the good. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(4), 429-446.

McClain, L. (1983). Interval estimation: Effect of processing demands on prospective and retrospective reports. *Perception & Psychophysics*, *34*(2), 185-189.

Miller, G. W., Hicks, R. E., & Willette, M. (1978). Effects of concurrent verbal rehearsal and temporal set upon judgements of temporal duration. *Acta Psychologica*, *42*, 173-179.

Moors, A., & De Hower, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2)*,* 297-326.

Ogden, R. S., Wearden, J. H., & Jones, L. A. (2008). The remembrance of times past: Interference in temporal reference memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(6), 1524-1544.

Ogden, R. S., Salominaite, E., Jones, L. A., Fisk, J. E., & Montgomery, C. (2011). The role of executive functions in human prospective interval timing. *Acta psychologica*, *137*(3), 352-358.

Ogden, R. S., Wearden, J. H., & Montgomery, C. (2014). The differential contribution of executive functions to temporal generalisation, reproduction and verbal estimation. *Acta Psychologica*, *152*, 84-94.

Posner, M. I., & Snyder, C. R. R. (1975). *Attention and cognitive control*. In R.L. Solso (Ed.), Information processing and cognition: The Loyola symposium. Hillsdale, NJ: Erlbaum, 55-85.

Predebon, J. (1996). The effects of active and passive processing of interval events on prospective and retrospective time estimates. *Acta Psychologica*, *94*, 41-58.

Predebon, J. (1999). Time judgments as a function of clock duration: Effects of temporal paradigm and an attention-demanding nontemporal task. *Perceptual and Motor Skills*, *88*, 1251-1254.

Rattat, A. C., & Droit-Volet, S. (2012). What is the best and easiest method of preventing counting in different temporal tasks?. *Behavior Research Methods*, *44*, 67-80.

Roeckelein, J. E. (2000). *The concept of time in psychology*. Westport, London: Greenwood Press.

Roy, M. M., & Christenfeld, N. J. (2008). Effect of task length on remembered and predicted duration. *Psychonomic Bulletin & Review*, *15*, 202-207.

Roy, M. M., Christenfeld, N. J., & McKenzie, C. R. (2005). Underestimating the duration of future events: Memory incorrectly used or memory bias?. *Psychological Bulletin*, *131*(5), 738.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*, *84*, 1-66.

Schuster, D., Rivera, J., Sellers, B, C., Fiore, S. M., & Jentsch, F. (2013). Perceptual training for

visual search. *Ergonomics, 56*(7), 1101-1115.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127-190.

Staddon, J. E. R., & Higa, J. J. (1999). Time and memory: Toward a pacemaker-free theory of interval timing. *Journal of the experimental analysis of behavior*, *71*(2), 215-251.

Taatgen, N. A., Van Rijn, H., & Anderson, J. (2007). An integrated theory of prospective time interval estimation: The role of cognition, attention, and learning. *Psychological Review*, *114*, 577-598.

Thomas, K. E., & Handley, S. J. (2008). Anchoring in time estimation. *Acta Psychologica*, *127*, 24-29.

Thomas, K. E., Newstead, S. E., & Handley, S. J. (2003). Exploring the time prediction process: The effects of task experience and complexity on prediction accuracy. *Applied Cognitive Psychology*, *17*, 655-673.

Tobin, S., Bisson, N., & Grondin, S. (2010). An ecological approach to prospective and retrospective timing of long durations: a study involving gamers. *PloS One*, *5*(2), e9271.

Tobin, S., & Grondin, S. (2009). Video games and the perception of very long durations by adolescents. *Computers in Human Behavior*, *25*, 554-559.

Underwood, G., & Swain, R. A. (1973). Selectivity of attention and the perception of duration. *Perception*, *2*, 101-105.

Walker, J. A., & Cousineau, D. (Submitted). Into the mind's eye: Exploring the fast-same effect in the same-different task. *American Journal of Psychology*.

Wearden, J. H. (2015). Passage of time judgements. *Consciousness and cognition*, *38*, 165-171.

Wearden, J. (2016). *The psychology of time perception*. Keele, UK: Macmillan Ltd.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology*, *61*(4), 569-587.

Wolfe, J. M., Palmer, E. M., & Horowitz, T. S. (2010). Reaction time distributions constrain models of visual search. *Vision research*, *50*(14), 1304-1311.

Zakay, D. (1992). The role of attention in children's time perception. *Journal of Experimental Child Psychology*, *54*, 355-371.

Zakay, D. (1993). Relative and absolute duration judgments under prospective and retrospective paradigms. *Perception & Psychophysics*, *54*(5), 656-664.

Zakay, D. (1998). Attention allocation policy influences prospective timing. *Psychonomic Bulletin & Review*, *5*(1), 114-118.

Zakay, D., & Block, R. A. (1995). An attentional-gate model of prospective time estimation. *Time and the Dynamic Control of Behavior*, 167-178.

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

Zakay, D., & Block, R. A. (2004). Prospective and retrospective duration judgments: An executive-control process. Acta Neurobiologiae Experimentalis, 64, 319−328.

Zakay, D., & Fallach, E. (1984). Immediate and remote time estimation—a comparison. *Acta psychologica*, *57*, 69-81.

Zakay, D., Nitzan, D., & Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. *Perception & Psychophysics*, *34*(5), 451-456.

Zakay, D., & Shub, J. (1998). Concurrent duration production as a workload measure. *Ergonomics*, *41*(8), 1115-1128.

Zakay, D., Tsal, Y., Moses, M., & Shahar, I. (1994). The role of segmentation in prospective and retrospective time estimation processes. *Memory & Cognition*, *22*(3), 344-351.