Summer 2023

An Examination of Sustained Attention During Complex Multitasking Scenarios

Jonathan Cheshire Rann

An Examination of Sustained Attention During Complex Multitasking Scenarios

by

Jonathan Cheshire Rann

Bachelor of Science
Winston-Salem State University, 2007

Master of Science
North Carolina Agricultural and Technical State University, 2011

Master of Arts
University of South Carolina, 2020

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Psychology

College of Arts and Sciences

University of South Carolina

2023

Accepted by:

Amit Almor, Major Professor

Doug Wedell, Committee Member

Jennifer Vendemia, Committee Member

Anne Bezuidenhout, Committee Member

Ann Vail, Dean of the Graduate School

DEDICATION

Dedicated to Dr. Amit Almor, my advisor. Thank you so very much for this experience this meant everything to me. Also dedicated to Dr. Flossie Rann, my mother. Thank you for your ceaseless love and support during my journey through life.

ACKNOWLEDGEMENTS

First, I would like to acknowledge each of the members of my committee who helped and watched me grow throughout this process: Dr. Doug Wedell, Dr. Jennifer Vendemia, and Dr. Anne Bezuidenhout. I learned so much from each of you and am grateful that you took the time and interest in my personal, professional, and academic development. Thank you so very much, you guys are the absolute best.

Second, I would like to acknowledge each of the academic and professional mentors that have taken me under their wing over the years: Dr. Anna Yu., Dr. Gerry Dozier, Dr. Elva Jones, Dr. Rebecca Caldwell, Dr. Christov Dichev, Dr. Matt Riedy, Sarah Leitner, Josette Francis, David Allston, Mark Stillman, Kathy Adams, and Debbie McCoy. I would not have been able to make this journey without you all. Thank you for every door you helped me open (or opened for me), and every mile you walked with me. I have endless appreciation for each of you.

Last, I would like to acknowledge all of the people who are closest to my heart who have loved me into the man that I am today: Mary Alice Kelly Foxx, Ravonda Dalton-Rann, Laura Foxx, Mary Johnson, Eartha Brown, Marie Henderson, Faith Fox, Cheryl Oliver, James Foxx Jr., Donovan Alexander, Justin Brintle, Victoria Tinarwo, Chris Taylor, Jerry Akouala, Jon Garnett, Shawn Yale, Jeffrey Purcell, Isaac Regenbogen, Sori Kim, Tiphane Pate, Jennifer Olejarczyk, Greg Colleton, and Aurelia Joy Smith. I am forever grateful for your love (and patience) as you guided me through life. Thank you all so very much. I will do my best to make you all proud.

ABSTRACT

This dissertation project examined the time-course of vigilance decrements that occur when operators perform demanding multitasking activities. For this aim, we conducted three experiments which implemented a novel paradigm we developed that measured performance during two sessions: a Single-Task session in which participants performed a go-no-go target detection task in the absence of any other task for approximately 12 minutes; and a Multi-Task session in which participants performed the detection task simultaneously with a driving-based tracking task for the same duration. A total of 183 participants from the University of South Carolina Department of Psychology took part in this study. Growth curve analyses revealed quadratic trajectories across accuracy and error measures, and linear trajectories across the tracking and RT measures, and that the inclusion of late verbal tasks (Experiment 3) negatively affected all detection task measures during more difficult conditions, but did not affect the tracking task measure. Further, vigilance decrements during both sessions and across all measures had higher intensity and variability during the more difficult conditions, and practice reduced these effects, but only to an extent. Together, these findings support the cognitive overload and opportunity cost models of vigilance performance. Insight from this work can significantly inform the design and development of complex operator-system interfaces, and further develop theoretical understanding of attention and multitasking performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Sustained attention, generally defined as the ability to continuously engage in relevant activities (Esterman & Rothlein, 2019; Fortenbaugh et al., 2017; Sarter et al., 2001), is a fundamental aspect of many tasks that require operators to detect and respond to critical signals over long periods (e.g., x-ray inspection, quality assurance testing, military surveillance) (Cardoso et al., 2019; Davies & Parasuraman, 1982; Dorrian et at., 2007; Ghylin et al., 2007; Hancock & Hart, 2002; Heikoop et al., 2017; Masoudian & Razavi, 2019; Reinerman-Jones et al., 2016; Smith et al., 2021). During these scenarios, operators must exert varying levels of control for achieving task requirements while also maintaining adequate levels of alertness throughout task execution (Ballard, 1996; Fernandez-Duque & Posner, 2001; Hancock, 1989; Klösch et al., 2022; Langner & Eickhoff, 2013; van Schie et al., 2021). This can be difficult, however, since *vigilance decrements* typically occur in which certain measures of performance, such as target hit rate (hits) and reaction time (RT), deteriorate as a function of time-on-task (Dillard et al., 2019; Mackworth, 1948; Parasuraman et al., 1987). The two leading theories to explain this phenomenon are the *cognitive overload account*, which attributes decrements to a progressive depletion of processing resources that can occur during demanding scenarios (e.g., Caggiano & Parasuraman, 2004), and the *cognitive underload account*, which attributes decrements to a progressive increase in task disengagement that can occur during monotonous scenarios (e.g., Robertson et al., 1997).

Researchers utilize continuous performance tests (CPTs) to examine the source of vigilance decrements under highly controlled experimental settings (*for review*: Riccio et al., 2002). These powerful tools are configured to examine how certain task characteristics can influence operator performance during demanding scenarios (Fisk & Scneider, 1981; Parasuraman & Mouloua, 1987; Roca et al., 2011), and how increasing cognitive load can modulate these effects (Buckley et al., 2016; He & McCarley, 2011; Luna et al., 2021; Luna et al., 2022; McBride et al., 2007). Examining performance under these conditions is particularly relevant since operators often perform multiple simultaneous tasks as a function of their job requirements (Chen & Joyner, 2009; Chérif et al., 2018; Karpinsky et al., 2018; St. John et al., 1997; Strayer et al., 2011). However, generalizing results from these studies is not always easy since researchers sometimes fail to properly account for the specific attentional mechanisms tapped by the tasks used in their paradigms (Nichols & Waschbush, 2004; Roebuck et al., 2016; Riccio & Reynolds, 2001). Improving the methodological and ecological validity of these studies is critical for developing methods for mitigating vigilance decrements to help to prevent attentional failures in modern work environments (Hancock & Matthews, 2019; Helton & Russell, 2017; Neigel et al., 2020).

CHAPTER 2

BACKGROUND

Attention is a multi-faceted concept that broadly describes the ability to focus awareness on an intended target (Chun et al., 2011; Hommel et al., 2019; James, 1890; Krauzlis et al., 2023). This ability is considered to be a central component of human goal-directed behavior, and has been a topic of philosophical debate over the last few centuries (Murray & Ross, 1982; Nicholas, 1979), and perhaps since the times of Aristotle (Fiecconi, 2021). However, it was not until the late nineteenth century that experimental paradigms were created to advance our knowledge of attention and its different components (Pavlov, 1902; Wundt, 1880). For example, the Dutch medical scientist Franciscus Donders performed a pioneering study in which he examined the processing times involved with specific decision-making scenarios (Donders, 1868). He did this by comparing RTs for simple tasks which required participants to press a button as quickly as possible whenever a nearby light was illuminated, with those from more complex tasks which required participants to press individual buttons whenever a left or right light was illuminated. Results showed that the duration for stimulus processing was longer for the more complex task, thus suggesting that the mental processes involved with stimulus detection and response may occur in serial stages (Sternberg, 1969), and that processing times may be directly correlated with the number of mental processes involved with a task.

Later research examined how this stimulus-response processing can progressively deteriorate during occupational scenarios in which operators are required to sustain attention while performing their work duties (Ditchburn, 1943; Scerbo, 1998; Wyatt & Langdon, 1932). This was a particularly concerning issue for radar system operators in the British Royal Air Force (RAF) during World War II, since they often monitored for enemy vessels over long watch periods (Baca, 2017; Lichstein et al., 2000). To help improve operator performance, the British RAF tasked researcher Norman Mackworth with designing an experimental paradigm to elicit performance deteriorations, coined as "vigilance decrements", within a laboratory setting. For this, he developed 'Clock Test' experiments which required participants to sustain attention to a black pointer that moved around the face of a clock over the course of approximately 2 hours, and to verbally report whenever they detected any irregular pointer movements (Mackworth, 1948). Results showed that the number of correct signal detection events (i.e., hits) progressively declined over time, typically within the first 30 minutes of task engagement. Importantly, this work formed the basis of many subsequent studies which sought to examine the source of vigilance decrements across different measures of performance, and in relation to the different components associated with attentional processing (e.g., Dember et al., 1992; Teichner, 1974; Warm et al., 2018).

2.1 VIGILANCE DECREMENTS

There are several theories to explain the occurrence of vigilance decrements during sustained attention tasks. One set of theories attribute vigilance decrements to certain task-specific factors. This is highlighted in the work of Parasuraman & Davies (1977), who developed a taxonomy which described the source of decrements as a function of two interacting factors. The first factor, discrimination type, distinguished

between successive tasks in which non-critical targets are successively presented and compared with a critical target standard retained in working memory, and simultaneous tasks in which critical targets are presented among non-critical targets. The second factor, target presentation rate, distinguished between high rates (24 targets per minute or higher) or low rates (below 24 targets per minute). According to their taxonomy, successive discrimination tasks involving high target rates would result in the largest vigilance decrements, while simultaneous discrimination tasks involving low target rates would result in smallest decrements (Lanzetta et al., 1987; Parasuraman, 1979).

Another set of theories specifically attribute vigilance decrements to the internal state of the operator. These are often organized into two broad theories: the *resource depletion* (or cognitive overload) account, which suggests that vigilance tasks are stressful and demanding, and that decrements are related to progressive declines in available attentional resources as time-on-task increases (Caggiano & Parasuraman, 2004; Fisk & Scerbo, 1987; Gartenberg et al., 2018; Helton & Warm, 2008; Matthews et al., 1993; Wiener et al., 1984); and the cognitive underload account, which suggests that vigilance tasks are monotonous and boring (Cummings et al., 2016; Greenlee et al., 2018; McBain, 1970; Scerbo et al., 1992), and that decrements are related to a progressive disengagement from the task towards 'mindless behavior', potentially due to several factors, such as task routinization (Manly et al., 1999; Robertson et al., 1997), lowered arousal (Eysenck, 1976; Hebb, 1955), perceptual desensitization (Mackworth, 1968), and/or goal habituation (Ariega & Lleras, 2011).

Both accounts are supported by physiological indicators of stress and alertness, as well as self-reports of task engagement and demand. For example, the cognitive overload

predicted by the resource depletion account is indicated by increases and higher variability in eye activity (e.g., pupil size, blink rate), cortical activity (e.g., increasing theta, decreasing beta), and cardiovascular activity (e.g., heart rate) (Hancock, 1989; Innes, 1973; Mehrabi & Kim, 2022), as well as subjective reports of both increased workload (Hart & Staveland, 1988) and task-induced stress (Matthews et al., 1999). On the other hand, the lowered arousal and boredom predicted by the underload account are indicated by decreases in heart rate and increases and higher variability in blink duration and alpha frequency oscillations (Basacik et al., 2015; Jarosch et al., 2019; Lohani et al., 2019; McWilliams & Ward, 2021), as well as subjective reports of higher levels of sleepiness and lower levels of task engagement (Mathis & Hess, 2009; Schmidt et al., 2009).

Interestingly, however, operators may also report higher levels of demand while performing monotonous and boring tasks (Körber et al., 2015; Smallwood et al., 2004; Stevenson et al., 2011; Zhang & Kumada, 2017). This inconsistency may be explained by the *mind-wandering account* (Giambra, 1995; Smallwood & Schooler, 2006), which suggests that during monotonous scenarios attentional resources progressively shift towards demanding task-unrelated thoughts (TUTs), resulting in declining performance over time. The *resource-control account* (Thomson et al., 2015) shares a similar view, but specifically attributes vigilance decrements to progressive failures of executive control processes to sustain task goals and prevent TUTs from consuming available resources. Finally, the *opportunity-cost account* (Kurzban et al., 2013) goes one step further and suggests that progressive declines in control are the result of strategic shifts of cognitive effort based on factors such as reward or motivation.

Whether vigilance decrements occur due to task-specific factors, the depletion of attention resources, cognitive underload, etc., evidence supporting each of these theories ultimately suggests that there is no single vigilance decrement pattern that applies to all sustained attention tasks (Luna et al., 2022; McCarley & Yamani, 2021; Rubinstein, 2020). Importantly, however, there is also evidence that vigilance decrements can be reduced or even eliminated given certain conditions, such as taking brief periods of rest to reset task goal activation (Ariga & Lleras, 2011; Ross et al., 2014), increasing arousal during monotonous scenarios (Al-Shargie et al., 2019; Atchley & Chan, 2011; Atchley et al., 2014), automating task performance through practice (Fisk & Schneider, 1981; Fisk & Scerbo, 1987; Parasuraman & Giambra, 1991; Schneider and Shiffrin, 1977), and manipulating potential costs and rewards associated with task performance (Arrabito et al., 2007; Esterman et al., 2014; Gutzwiller et al., 2015; Hancock et al., 2016), to name a few.

2.2 CONTINUOUS PERFORMANCE TESTS

Continuous performance tests (CPTs) are specialized target detection tasks often used by researchers to investigate the source of vigilance decrements in laboratory settings (Ord et al., 2021; Riccio et al., 2002; Smid et al., 2006). CPTs have traditionally been used to assess attentional disorders in clinical populations, and are increasingly utilized in sustained attention research due to their versatility and sensitivity for measuring different aspects of attention (Ballard, 2001; Bearden et al., 2004; Conners et al., 2003).

Common versions follow a basic paradigm in which participants are required to either respond, or withhold response, whenever they detect a critical target stimulus presented among a series of distractor stimuli. For example, the standard (or CPT-X)

paradigm (Rosvold et al., 1956), developed to assess attention-deficit/hyperactivity disorder (ADHD), requires participants to visually monitor the presentation of a continuous stream of [non-target] letters, and then to manually respond whenever they detect the infrequent target letter 'X'. The sustained attention to response task (SART; Robertson et al., 1997), developed to assess patients with acquired brain injuries, requires participants to manually respond to the presentation of a continuous stream of [non-target] digits, and then to inhibit response whenever they detect the infrequent target digit '3'. Finally, the psychomotor vigilance test (PVT; Lim & Dinges, 2008; Wilkinson & Houghton, 1982), developed to study the effects of sleep deprivation, requires participants to manually respond as fast as possible to the presentation of a target presented at different intervals.

CPTs are particularly useful for experimental research since they produce several different measures of performance (Borgaro et al., 2003; Epstein et al., 2003; Roebuck et al., 2016). These include a measure of the time (or latency) between stimulus presentation and response (*RT*); measures of lapses (i.e., errors) in attention, such as failures to respond when a target is presented (*omissions)*, and incorrect responses when no target is presented (*commissions*); measures of detection accuracy, such as correct responses to target detection events (*hits*), and correct non-responses when no target is presented (*correct rejections*, or *CRs*); as well as measures derived from signal detection theory (SDT; Green & Swets, 1966). These include the ability to discriminate between targets and non-targets (*sensitivity*, or *d'*), considered by some to be a more powerful measure of detection accuracy than hits and false alarms alone (Lord, 1985); and the criterion (or strategy) used to characterize the willingness of an observer to respond to a presented

target. In vigilance studies, this criterion is often expressed as speed vs. accuracy tradeoffs in CPT responses (*response bias*, or β), where emphasis on speed-over-accuracy describes *liberal* bias, and emphasis on accuracy-over-speed describes *conservative* bias (Azizi et al., 2018).

Each of these measures are averaged across the duration of the CPT to assess overall performance on the task, and across blocks of trials to assess the ability to sustain attention (Bubnik et al., 2015; Cornblatt et al., 1989; Esterman & Rothlein, 2019). Resulting measures are then compared with the scores from the normative group to indicate whether individuals present behavioral markers related to common attention disorders, such as ADHD (Conners, 2000; Epstein & Loren, 2013; Gilden & Hancock, 2007; Kirlin, 2002; Reynolds et al., 2008). For example, *impulsivity*, generally associated with a degradation in response inhibition (Bari & Robbins, 2013; Dalley et al., 2011), is characterized by higher commissions, faster RTs, and more liberal response bias; and *inattentiveness*, generally associated with a degradation in attentional capacity (Diamond, 2005; Edwards et al., 2007), is characterized by higher omissions, lower sensitivity, and slower (and more variable) RTs. Moreover, while the ability to sustain attention in both normative and clinical populations is generally characterized by consistent measures of performance (i.e., detection accuracy and errors) across trial blocks, failure to sustain attention is typically characterized by high variability and/or progressive declines in performance across blocks, with clinical populations thought to be especially susceptible to these decrements (Fuermaier et al., 2022; Tucha et al., 2009).

Interpreting measures is not always straightforward, however, since common CPTs vary along several categories (stimulus modality, motor control requirements,

detection-response criteria, etc.) (Denney et al., 2005; Hall et al., 2016; Riccio et al., 2001). As a result, the relevance of certain measures as indicators of performance may vary depending on the configuration of the CPT. For example, since the standard (Rosvold et al., 1956) and the PVT (Lim & Dinges, 2008) paradigms both involve maintaining alertness for quickly detecting and responding to critical target stimuli, then the most relevant indictors of performance for these CPTs include hits and RTs, as well as errors of omission and commission (Basner et al., 2018; Levy et al., 2018). However, RTs may be a better indicator of performance in the PVT since it involves more frequent responding and thus better captures moment-to-moment fluctuations in RTs (Rosenberg et al., 2013). Alternatively, since the SART (Robertson et al., 1997) involves maintaining alertness for inhibiting prepotent (automatic) responses to critical target stimuli, then its most relevant indicators of performance include CRs and commissions, while hits, RTs, and omissions may be less relevant since they relate more to the secondary task of motor responding (Berger & Cassuto, 2014; Ogundele et al., 2011). Interestingly, this difference underscores the controversy in the literature regarding whether the SART measures the ability to sustain attention, or whether it measures the ability to sustain inhibition to response (i.e., Carter et al., 2013; Dang et al., 2018; Helton, 2009; Jun, 2021; Wilson et al., 2016).

Variations between CPT paradigms can also affect the interpretation of both sensitivity and response bias since they are derived from calculations involving hits and commissions (Broadbent & Gregory, 1965; Koelega et al., 1989; MacMillan & Creelman, 1990; Nuechterlein et al., 1983; Parasuraman & Mouloua, 1987; See et al., 1997). Therefore, these measures may be overall weak indicators of performance in

paradigms involving high signal-to-noise ratios, such as the SART, since these often produce disproportionate response distributions in which hits are prone to ceiling effects and commissions are prone to floor effects (Huang-Pollock et al., 2012). Relatedly, there is a debate in the literature regarding whether vigilance decrements can be attributed to a progressive loss in sensitivity, or rather due to a shifting strategy in which responses become more conservative with time-on-task (Thompson et al., 2016). While there is evidence for both, researchers should be mindful when interpreting SDT measures since certain indicators of performance, such as hits, commissions, and RTs, are highly confounded with response bias and may thus progressively increase as responses become more liberal, and decrease as responses become more conservative (Bedi et al., 2023; Craig, 1978; Helton et al., 2009).

Lastly, certain paradigms are specifically configured to determine the role of operator arousal in modulating CPT performance (Conners et al., 2003; Egeland & Kovalik-Gran, 2010). A common method of achieving this is by reducing the predictability of incoming stimuli by alternating interstimulus intervals (ISIs) within-task (MacLean et al., 2009). For instance, the PVT variably changes ISIs between 2 and 10 seconds throughout the duration of the task (Basner and Dinges, 2011), while the Conners (2014)'s CPT (CCPT) rotates ISIs between 1s, 2s, and 4s every set number of trials. The reported effects of these ISI manipulations, as well as the measures used to assess these effects, tend to vary between studies (Silverstein et al., 2004). Indeed, several studies report that shorter ISIs are more associated with higher omissions and slower and more variable RTs, possibly due to increased processing demands (Matthews et al., 2017), while longer ISIs are more associated with higher commissions, possibly due to

stimulus expectancy and motor preparation (Los et al., 2001; Nickerson, 1968); other studies suggest that target discrimination ability is better overall during the longer ISIs since respondents have more time to respond in these conditions (Huang-Pollock et al., 2012); and others suggest that variability in performance between ISIs is associated with difficulty adjusting to changes in task demands (Conners et al., 2003; Ord et al., 2021). While some of this inconsistency can be attributed to the nature of the CPT paradigms and/or the measures of interest for a particular study, another factor could be due to the length of time associated with each ISI, since trial blocks with longer ISIs also have disproportionately longer durations (and, as a result, more datapoints) than those with shorter ISIs (Ballard, 2001).

2.3 DISSOCIATING VIGILANCE

While sustained attention is a common measure among CPT paradigms, it is not always clearly defined in the literature (Esterman & Rothlein, 2019; Fortenbaugh et al., 2017; Langner & Eickhoff, 2013; Oken et al., 2006; Sarter et al., 2001). One reason for this is that it is highly related to other components of attention (Fisher, 2019; Mirsky et al., 1991; Posner & Boies, 1971; Shallice et al., 2008). For example, sustaining attention involves continuously allocating attentional resources towards performing a task over durations lasting several minutes to hours (Esterman et al., 2013; Manly et al., 1999; Rosenberg et al., 2016). These limited capacity resources can be focused on a particular stimulus or task goal, as in the case of selective attention (Treisman, 1969; Johnston & Dark, 1986), they can also be shared between multiple stimuli or task goals, as in the case of multitasking (Allport et al., 1972; Black et al., 2022; Koch et al., 2018; Poljac et al., 2018). Underlying these components are executive control processes that are essential for maintaining higher level functions involved in different aspects of task performance

(Kieras et al., 2000; Kieras & Meyer, 1997; Norman & Shallice, 1986; Rubinstein et al., 2001).

However, another reason for the lack of clarity is that sustained attention is often confounded with vigilance (Klösch et al., 2022; MacLean et al., 2009; Oken et al., 2006; van Schie et al., 2021). Mackworth (1968) describes vigilance as a state of readiness to react and respond to small irregular changes in the environment. This suggests an involvement of physiological components that are separate from (but still related to) those involved in sustained attention (Head, 1923; Klösch et al., 2022). These include alertness, defined as a quantifiable sensitivity to incoming stimuli (Posner, 2008), and arousal, defined as a sudden upward change in alertness associated with the sleep-wake spectrum (Stum & Williams, 2001; van Schie et al. 2021). Furthermore, van Schie et al. (2021) suggests that a major difference between the two concepts is that sustained attention has a horizontal directional aspect (i.e., goal directed behavior *over time*), while vigilance has a vertical directional aspect (i.e., *upward changes* in alertness), and that maintaining vigilance is necessary for sustaining attention, but the opposite is not true.

Posner & Petersen (1990)'s attention networks model discusses these concepts with relation to other attentional components. According to their model, there are three anatomically and functionally independent brain networks related to the main functions of attention (Posner & Petersen, 1990). The orienting network is involved with directing and prioritizing attention to important sensory stimuli. It consists of *overt* orienting, for physically directing the eyes toward an intended target or location, and *covert* orienting, for mentally shifting attention toward an intended target without physical eye movement. This network is most associated with the posterior areas of the parietal lobe, and includes

the superior colliculus and the pulvinar nucleus of the thalamus. The alerting (arousal/vigilance) network is involved with modulating arousal for achieving and maintaining an alert state. It consists of *phasic* alerting, for quickly moving from a non-alert to an alert state by influence of exogenous factors (salient stimuli, cues, etc.); and *tonic* arousal, involved with sustaining attention for detecting rare targets or for continuously responding to stimuli by influence of endogenous factors (task requirements, goals, etc.). This network is most associated with the reticular system, as well as the right inferior frontal, inferior parietal, and anterior cingulate regions of the brain. Finally, the executive network is involved with effortful allocation of attentional resources for focal conscious processing, conflict resolution, monitoring, error detection, planning, coordination, and inhibitory control. This network is most associated with the anterior areas of the frontal cortex, as well as the lateral frontal lobes, prefrontal cortex, anterior cingulate cortex, and supplementary motor area.

Fan et al. (2002) developed the Attention Network Test (ANT) paradigm to independently measure and evaluate these networks within a single experiment. It consists of a flanker task (Eriksen & Eriksen, 1974), in which participants must quickly respond to the direction pointed by a central target while ignoring the direction of flanking arrows, and a spatial cueing task (Posner, 1980), in which flanker trials (i.e., neutral, congruent, incongruent) are sometimes preceded by informative visual cues (i.e., center cue, double cue, spatial cue). Efficiency of the three attentional networks is assessed by measuring how RTs are influenced by the flankers and visual cues. Specifically, the orienting effect is calculated by subtracting the mean RT of the spatial cue conditions from the mean RT of the center cue, the alerting effect is calculated by

subtracting the mean RT of the double-cued conditions from the mean RT of the no-cue conditions, and the executive control effect is calculated by subtracting the mean RT from all congruent flanking conditions from the mean RT of incongruent flanking conditions.

Callejas et al. (2004) extended this with the ANT for Interactions (ANTI), which examined features of the interactions between the different networks. This paradigm modified the original ANT by adding an auditory tone as a warning signal cue to separately assess the alerting network, and an uninformative cue to assess the exogenous component of the orienting network. Results from this work found evidence that the alerting network produces an inhibitory effect on the executive network to support fast responses to detect infrequent stimuli, the orienting network positively influences the executive network for selectively attending to relevant stimuli, and the alerting network enhances the orienting network for faster orienting.

While both the ANT and ANTI have been successfully applied in different research settings to assess attentional functioning (e.g., Gamboz et al., 2010; Ishigami et al., 2016; Johnson et al., 2008; López-Ramón et al., 2011; MacLeod et al., 2010), they are not well equipped to measure vigilance as they only infer it using a phasic alertness task (Ishigami & Klein, 2010). To overcome this, Roca et al. (2011) developed the ANT for Interactions and Vigilance (ANTI-V), which separately assessed tonic alertness by embedding a SART-like secondary vigilance task within the main ANTI task. This provided several direct measures of vigilance performance, such as hits and commissions, as well as SDT indexes of sensitivity and response bias. However, these measures did not take into account some of the complexity involved with assessing vigilance. For example,

the effect of time-on-task was only indirectly examined using global (i.e., averaged) measures across the different conditions (e.g., cue vs. no cue blocks, etc.), and comparative measures between the first and last trial blocks (Ishigami & Klein, 2009). Furthermore, vigilance was only measured as a unitary concept (i.e., signal detection performance), without consideration of other perspectives discussed in the literature (e.g., Lim & Dinges, 2008).

To address these limitations, Luna et al. (2018) developed the ANT for Interactions and Vigilance – executive and arousal components (ANTI-Vea). This paradigm obtained independent measures of the three attentional networks using the ANTI, and directly measured vigilance as two dissociated components using a SART-like task to measure executive vigilance (EV), defined as the ability to sustain attention for detecting infrequent stimuli, and a PVT-like task to measure the arousal component of vigilance (AV), defined as the ability to achieve and maintain fast reaction to stimuli. Additionally, unlike the previous ANT-based paradigms, the ANTI-Vea used repeated measures ANOVAs to measure changes in performance across the different blocks and conditions, thereby allowing decrements in performance to be captured as a function of time on task. Results from this study found the main effects and interactions of the attentional networks observed in the ANTI task (Luna et al., 2021), and importantly, found an EV decrement in which hits progressively decreased and response bias incremented towards a more conservative criterion (i.e., Thomson et al., 2016), and an AV decrement in which mean and variability of RTs increased over time.

Recently, Luna et al. (2022) utilized the ANTI-Vea paradigm to examine the role of cognitive load in modulating vigilance performance. In line with the resource

depletion theory, they predicted that increasing cognitive load by combing vigilance tasks during multitasking conditions would result in larger EV and AV decrements than during single task conditions. Interestingly, results from their study showed that: for the EV task, mean hit rates were higher and response bias was more liberal during the dual- and triple-task conditions compared to the single-task condition, while neither measure showed significant changes across time-on-task only in the dual-task condition; and for the AV task, mean RT was slower in the dual-task compared to the single-task condition, while the variability of RT and omissions progressively increased with time-on-task only in the single-task condition. Overall, these results showed partial support for the resource depletion theory, since both EV and AV performance measures were worst in the more demanding conditions, however, other theories are needed to explain the mitigative influence of cognitive load on the vigilance decrements during the dual-task condition. To this end, the authors explained results in line with the inverted-U hypothesis (Wiener et al., 1984), which suggests that performance is most optimal when operators are neither under-aroused or over-aroused, and therefore supports other findings in the literature which suggest that increasing task load can reduce or eliminate certain vigilance decrements given the right conditions (e.g., Moray & Haudegond, 1998; Stearman & Durso, 2016).

Indeed, over-arousal and increased load are commonly implicated as the cause of operational failures associated with complex vigilance tasks (Engström et al., 2005; Lam, 2002; Lamble et al., 1999; Muhrer & Vollrath, 2011; Regan et al., 2011; Strayer & Fisher, 2016). These are often studied using multitasking paradigms which combine PVT-like tasks with visual motor tasks such as object tracking (Buckley et al., 2016;

Poudel et al., 2008; Tombu & Seiffert, 2008; Vater et al., 2017) and simulated driving (Bowden et al., 2019; Bruyas & Dumont, 2013; Cooper et al., 2009; Strayer et al., 2003; Strayer et al., 2019; Tillman et al., 2017). Seemingly less frequent are studies which pair complex choice and even SART-like CPTs with tracking-like tasks (Castro et al., 2019; Hahn et al., 2011; Mehler et al., 2008). This may be a key difference to consider when interpreting results from vigilance studies, since they commonly generalize findings based on the results from broad categories of CPTs (e.g., “EV-related” CPTs such as the SART), while failing to appropriately dissociate the attentional mechanisms tapped by the specific task configurations.

For example, unlike many other studies in vigilance literature (e.g., Baldwin & Lewis, 2017; Dang et al., 2018; Dillard et al., 2014; Hoonakker et al., 2017; Manly et al., 1999; Robertson et al., 1997; Seli et al., 2013), ANT-related studies fail to include CRs and/or commissions in their analysis of EV performance (i.e., Luna et al., 2018; Luna et al., 2022; Roca et al. 2011). This is a critical omission since executive control supports the phasic alertness involved with the primary tasks of inhibiting and strategizing responses in the SART (Scullin & Bugg, 2013), and also supports the allocation of attentional resources towards establishing the feed-forward motor routines required for the secondary responding task (Schneider and Shiffrin, 1977). Therefore, measures associated with executive processes are most relevant when interpreting performance on SART-like tasks (Carter et al., 2013).

Furthermore, ANT-related studies fail to utilize other psychophysical tasks that may provide more ecological validity than CPTs alone, like object tracking (Böttcher et al., 2023; Busk & Galbraith, 1975; Drew, 1951; Miall et al., 1993; Pylyshyn, 1994;

Pylyshyn & Storm, 1988). Indeed, object tracking is a classic CPT often used to measure and assess the cognitive functions associated with aspects of real-world tasks, such as the steering and curve negotiation involved with driving (Rajan et al., 2016; Salvucci & Taatgen, 2011; Tatham et al., 2014). Central to these functions are mechanisms associated with tonic arousal, involved with automatizing perceptual and motor routines for continuously monitoring the slow movement of target stimuli within the vision field (Barry et al., 2005; Rusby et al., 2007); and mechanisms associated with phasic alertness involved with focusing perceptual and motor control on the fast and/or unpredictable movement of target stimuli (Castro et al., 2019; Kinnear et al., 2013; Vardaki et al., 2014).

Interestingly, several studies also demonstrated mitigative effects of cognitive load on aspects of tracking performance (Cooper et al., 2013; He & McCarley, 2011; He et al., 2014; Li et al., 2018). However, in contrast to Luna et al. (2022), these findings were explained using the hierarchical control theory (Logan & Crump, 2009; Medeiros-Ward et al., 2014), which suggests that allocating attention to highly automatized tasks may actually disrupt performance rather than maintain or improve it, and that increasing task load may direct cognitive resources away from disrupting the automatized task (Forster & Lavie, 2009; Lavie et al., 2004).

While both ANT and tracking paradigms successfully demonstrate that they can be utilized to test different theories of performance under highly controlled settings, neither provide adequate explanation of the shared mechanisms involved with their performance during complex multitasking scenarios. Furthermore, while tracking tasks provide a more ecological method for evaluating visual motor performance, they often do

not utilize a theoretically based task structure for evaluating attentional components in line with the CPT literature. Therefore, we explore this knowledge gap in the following study.

2.4 THE PRESENT STUDY

In the present study, we conducted three experiments which implemented a novel paradigm we developed that measured performance during two fixed-order sessions: a *Single-Task* session in which participants performed a go-no-go target detection task in the absence of any other task for approximately 12 minutes; and a *Multi-Task* session in which participants performed the detection task simultaneously with a driving-based tracking task for the same duration. The target detection task was modeled after the Conners (2014) CPT and implemented using E-Prime 3.0 (Psychology Software Tools, Pittsburgh, PA), while the driving-based tracking task was the smooth pursuit continuous tracking and reaction (ConTRe; Mahr et al., 2012) task implemented in the OpenDS (Math et al., 2013) driving simulator environment.

The CPT provides several discrete measures of performance (i.e., hits, omissions, commissions, CRs, RT, sensitivity, and response bias), a design structure in which measures are averaged across six blocks of trials and where each block serves as a time-point of task performance, and a within-task difficulty manipulation in which interstimulus intervals (ISIs) alternate between shorter (i.e., more demanding) and longer (i.e., less demanding) ISIs after every set number of trials. The driving-based tracking task provided a single continuous measure of performance (i.e., average tracking distance), and was analyzed according to the block and ISI structure used with the target detection task. While all the measures are of interest in this study, our main focus is on the most relevant measures associated with the primary inhibition CPT task (i.e.,

commissions and CRs), the secondary responding CPT task (i.e., target hits and RT), and the tracking task (i.e., average distance).

The design of this paradigm was influenced by previous studies that examined vigilance performance using combined multitasking paradigms (Buckley et al., 2016; Castro et al., 2019; Luna et al., 2022; Rann & Almor, 2022). However, unlike these and other studies that utilized counterbalanced designs with alternating condition orders (e.g., Chiew & Brazer, 2013; Moran et al., 2020), the paradigm presented in this dissertation utilized a fixed-order design in which the Single-Task session was always performed first, and the Multi-Task session was always performed second. This was intentionally done because our focus was not on the differences between the Single- and Multi-Task sessions, but on changes in performance during the Multi-Task session. Always starting with just the CPT as the single task provided participants ample time to learn the demanding CPT so as to increase power for detecting vigilance decrements in both tasks during the Multi-Task session (Carter et al., 2013; Cornblatt et al., 1989; Hope et al., 1998; Lemay et al., 2004).

2.4.1 OBJECTIVES

The first objective for this dissertation was to determine whether our paradigm could elicit and measure the performance changes that occur when the target detection and tracking tasks are performed simultaneously. In line with the resource depletion account of vigilance performance (Caggiano & Parasuraman, 2004; Grier et al., 2003; Helton & Russell, 2017), operators may become overloaded when they simultaneously perform multiple tasks, especially when target event rates are faster compared to slower (Parasuraman & Davies, 1977). These effects may be supported by subjective state measures that suggest a positive correlation of stress and task demand (Desmond et al.,

1998; Matthews et al., 2017; Saxby et al., 2013). Therefore, our first critical hypothesis (H1) is that performance will be worst across CPT-related measures in the Multi-Task session, and our second critical hypothesis (H2) is that performance will be worst across all measures in this session during the faster compared to the slower ISIs.

The second objective for this project was to determine whether these changes exhibit non-linear trajectories throughout their duration. Traditionally, researchers analyzed performance on sustained attention tasks using simple statistical methods (e.g., correlations, ANOVAs, and regressions) that describe the general relationship between variables across independent time-points, with results typically represented as progressive linear declines in performance over time (Becker et al., 1991; Berardi et al., 2001; Caggiano & Parasuraman, 2004; Colquhoun, 1967; Conners et al., 2003; Dember et al., 1992; Fisk & Schneider, 1981; Grier et al., 2003; Helton & Russell, 2011; Luna et al., 2022; Mackworth, 1948; Nuechterlein et al., 1983; Parasuraman, 1979; Pattyn et al., 2008; Ralph et al., 2017; Roach et al., 2006; See et al., 1995). However, some suggest that these methods are not well suited for the analysis of vigilance since they do not properly account for the non-linear changes in performance that can occur during the course of task execution (Winter & Wieling, 2016), nor do they account for the different factors that influence these changes (e.g., nested dependencies, between and within-group variations, etc.) (McArdie & Nesselroade, 2003).

Indeed, operators often require extended periods of practice to learn how to perform their tasks most optimally for a given occupational scenario (Bherer et al., 2005; Brown, 2008). For more difficult tasks that require higher levels of executive control, these learning (or task attenuation) effects can result in rapid changes in performance as

operators become accustomed to task requirements (Fisk & Scerbo, 1987). Interestingly, these changes can mask, delay, or even counteract the occurrence of vigilance decrements during the early stages of task performance (Fisk & Scneider, 1981; Gartenberg et al., 2018; Smith et al., 2023; Thomson et al., 2015). To explore this, we utilize growth curve analyses (GCAs) in combination with a mixed-effects modeling framework as our method of data analysis (Boiteau et al., 2014; Rann & Almor, 2022). Therefore, given the predicted difficulty involved with performing the CPT and tracking tasks simultaneously (i.e., H1 and H2), as well as the relatively short duration in which operators must perform these tasks (i.e., approx. 12 minutes), our third critical hypothesis (H3) is that GCAs will reveal non-linear changes in performance across all measures in the Multi-Task session.

CHAPTER 3

EXPERIMENT 1

3.1 METHODS

3.1.1 PARTICIPANTS

A total of 63 native English-speaking participants (age: $M = 20.19$, $SD = 1.48$) from the University of South Carolina Department of Psychology undergraduate participant pool took part in the experiment. This number was based on an apriori power analysis with a medium effect-size (η2) = .50, alpha (α) = .05, and power (β) = .80 (Faul et al., 2007).

Of the 63 participants, 45 were female (age: $M = 20.2$, $SD = 1.32$) and 18 were male (age: $M = 20.2$, $SD = 1.86$). Participants were compensated with extra credit for their time. Recruitment criteria for this study specified that participants had to be native speakers of English and must have also held a valid driver's license. There were no other inclusion or exclusion criteria for selecting participants.

3.1.2 HARDWARE AND SOFTWARE

The CPT was coded and implemented using E-prime 3.0 (Psychology Software Tools, Pittsburgh, PA) stimulus presentation software. It ran on a Dell Precision Tower 7810 computer, and was displayed on a Dell 17in monitor with 1280 x 1040px resolution and a refresh rate of 60hz. Participants responded to the task stimuli by pressing a Linemaster T-91-S Treadlite II Foot Switch attached to a Psychology Software Tools 200A Serial Response Box that interfaced with the computer that ran the CPT.

The driving-based tracking task was coded and implemented using the OpenDS (Mahr et al., 2012; Math et al., 2013) driving simulator framework. It ran on a Dell XP 435t/9000 computer, and was displayed on a Dell 27in full HD flat panel monitor with

1920 x 1080px resolution and a refresh rate of 60hz. After collecting data for approximately half of the participants in E1, the computer malfunctioned and was replaced with a Dell OptiPlex 790 computer that executed the task and collected simulator data for the remainder of the study. Participants performed the driving-based tracking task using the Microsoft SideWinder Precision Racing Wheel (USB) that interfaced with the computer that ran the simulator.

All computers ran Windows 10 Pro operating system. R 4.2.2 (R Core Team, 2012), and R Studio version 1.1.447 (RStudio Team, 2016) were used for data analysis.

3.1.3 TASKS

The primary task was a go-no-go target detection task based on the Conner's (2014) CPT (CCPT). It required participants to continuously monitor a computer screen as black letter stimuli (*font*: Arial; *size*: 200px) were presented against a white background, and to respond by pressing a foot pedal when they detected target letters ('A'-'W', 'Y','Z') and withhold response when they detected the critical non-target letter 'X' (Figure 1). Each stimulus presentation (whether target or non-target), as well as the subsequent time window up until the next stimulus presentation, was considered a trial. There were 270 trials presented over the course of approximately 12 minutes, and trials continued to iterate regardless of whether participant responses were recorded. Experiment trials were divided into six blocks, with each consisting of 45 trials in which the time between stimulus presentations, called the inter-stimulus interval (ISI), randomly alternated every 15 trial sub-block for either 1000ms, 2000ms, or 4000ms (Figure 3.1). Target and non-target stimuli were randomly selected with an 80% to 20% ratio of targets to non-targets each presented for 250ms. Performance data for this task was sampled per trial and stored in E-prime output data files.

The secondary task was a driving-based tracking task, known as the continuous tracking and reaction (ConTRe) task (Mahr et al., 2012), that was implemented in the OpenDS driving simulator (Math et al., 2013). It required participants to continuously track a yellow target moving across the simulator screen with a blue cursor that they controlled using the steering wheel (Figure 3.2). The yellow target was placed approximately 20 ft in front of the participants' view on the simulator screen, and moved horizontally (i.e., left-to-right, right-to-left) across the screen at constant lateral speed of 1 simulated meter per second. The yellow targets direction of movement (left-to-right, right-to-left) changed at random times. Participants only had control of the lateral movement of the blue cursor. Performance data for the tracking task was measured approximately every 19 milliseconds and stored in a MySQL database.

3.1.4 SETUP

Both tasks were displayed on monitors that were placed approximately two feet directly in front of the participants. The CPT was displayed on a smaller computer monitor that was placed behind and beneath a larger monitor that displayed the driving-based tracking task. For the CPT, the horizontal viewing angle was approximately .61° and the vertical visual angle was .64°; for the tracking task, the horizontal viewing angle was approximately .98° and the vertical visual angle was .55°. During the Single-Task session, the larger monitor was turned off so that participants only attended to the CPT displayed on the smaller monitor (Figure 3.3a). During the Multi-Task session, both tasks were simultaneously presented thus requiring participants to split their attention between both monitors as they performed the tasks (Figure 3.3b). The steering wheel used for the tracking task was placed directly in front of participants at chest level on the desk where the monitors were located. The foot pedal was placed directly in front of participants on

the floor close to where their right foot would naturally rest. Participant line of sight was located near the bottom of the monitor displaying the tracking task, at a point where the CPT and tracking task stimuli were approximately equal distance (red cross, Figure 3b). However, this point varied depending on the height of each participant.

3.1.5 QUESTIONNAIRES

Participants were administered questionnaires at different times throughout the study. Before starting the experiment, they were first asked to report basic demographic information, and to answer a few questions regarding their driving experience and sleep-related habits. They were also asked to report their pre-experiment subjective states of drowsiness and fatigue on two five-point scales in which ‘1’ indicated the lowest levels of drowsiness and fatigue, and ‘5’ indicated the highest levels. After completing the first experiment session, participants were asked to again report their subjective states of drowsiness and fatigue on two five-point scales. They were also asked to assess the difficulty of the previous experiment session on the same five-point scale. Then, after completing the second experiment session, they were asked one more time to report their drowsiness, fatigue, and perceived difficulty.

3.1.6 PROCEDURE

Prior to starting the experiment, participants were asked to show proof that they had a valid driver's license (as per the inclusion criteria), and that they had put their cell phones on silent so as not to distract their attention during either of the experimental sessions. Next, participants were given a consent form approved by the University of South Carolina IRB to review and sign, placed in a small room in front of the computers that ran the CPT and driving-based tracking task, and were then provided instructions on how to perform both tasks (e.g., give equal priority to both tasks in the Multi-Task

session). Before each session, variables and parameters (including participant number and session ID) were set by the researcher. Next, participants completed practice sessions to help prepare and acclimate them to the CPT and tracking tasks. They practiced the CPT for approximately 120 seconds before they performed the Single-Task session, and they practiced the combined CPT and tracking task simultaneously for approximately 120 seconds before they performed the Multi-Task session. The length of these practice trials was set to allow participants enough time to perform one block of trials in which they were exposed to each of the three ISIs in the CPT. During the experiment, the researcher sat outside of the room where participants performed the tasks in both sessions and listened for continuous pedal clicks and friction from steering wheel movements as auditory confirmation of compliance on both tasks.

3.1.7 VARIABLES AND MEASURES

The variables used for analysis included session (Single-Task, Multi-Task), block (1, 2, 3, 4, 5, 6), and ISI sub-block (1000ms, 2000ms, 4000ms). We collected seven measures in total: target hit rate (hits), omissions, commissions, correct rejections (CRs), and reaction time (RT) were collected from the CPT; sensitivity (*d’*) and response bias (*β*) were derived from hits and commissions CPT measures; and tracking distance was collected from the driving-based tracking task. Table 1 lists these measures, and provides a description of their source, definition (as relevant to the characteristics of the tasks used in this study), and the psychological processes they are most associated.

3.1.8 SIGNAL DETECTION THEORY

SDT provides a framework for describing the ability of an observer to detect (or discriminate) critical signals in the presence of background noise (Green & Swets, 1966; Marcum, 1947; Stanislaw & Todorov, 1999). According to SDT, there are four possible

outcomes to a signal detection situation: hits, omissions, commissions, and CRs. Statistical decision theory conceptualizes these outcomes as two probability distributions placed along an x-axis, with both being normally distributed with equal size and variance (Banks, 1970; Marcum, 1947). The noise distribution has a mean of 0, and the signal distribution has a mean greater than 0 and is thereby displaced to the right by the magnitude of the signal. Further, a criterion (or threshold) for determining whether a signal is detected is placed between both distributions. This criterion is controlled by the observer, and its placement (between the two distributions) may shift depending on their goals, expectations, and physiological state (Macmillan & Creelman, 1990; Neigel, Claypoole, & Szalma, 2019).

Performance in this framework is commonly assessed using two metrics (Lynn & Barrett, 2014; Swets, 1986). The first, *d*' (dee prime), describes the ability/sensitivity to detect a signal and is measured as the standardized difference between the means of the signal and noise distributions. Values for *d*' typically range between 0 (indicating chance performance) and 3 (indicating high signal detectability/sensitivity). It is calculated as the z-transformation of hits minus the z-transformation of commissions, where neither hits nor commissions are equal to 1 or 0 (note: this is often not the case when data is aggregated from multiple participants). The second, β (beta), describes the willingness of an observer to detect a signal (i.e. response bias), and is measured as the ratio of the heights of the signal and noise distributions at the criterion. In cases when the signal frequently occurs, and/or if the costs of making a false alarm are small, then the observer may shift the criterion to the left and thus increase the likelihood that they will respond that the signal was present. In cases when the signal rarely occurs, and or if the cost of

making a false alarm is high, then the observer may shift the criterion to the right and thus require more evidence of the signal before they will respond that it was present. Values for β are typically measured in base 10 log. When the base 10 log of β is less than 0, the observer is thought to have a more liberal response bias (i.e. they require less evidence to detect signal). When the base 10 log of β is greater than 0, the observer is thought to have a more conservative response bias (i.e., they require more evidence to detect signal). When the base 10 log of β equals 0, the observer is thought to be unbiased.

3.1.9 DATA PREPARATION

The CPT data from the practice blocks (blocks 1 and 3) was discarded and the CPT data from the Single- and Multi-Task blocks (blocks 2 and 4, respectively) was retained for analysis. The data from the driving-based tracking task in the Multi-Task session were stored in a MySQL table. After removing the data from the practice trial (block 3), the experimental data (block 4) were merged together (using customized SQL scripts), and then exported to .csv format for further data processing and analysis in R.

Prior to statistical analysis, the data files for both the CPT and driving-based tracking task were temporally aligned and merged so that tracking performance could be accurately matched with the CPT data. Following this alignment, tracking data was filtered to retain only the data from the 500ms before and after each CPT target stimuli presentation. This was done to create matched fixed time windows of tracking data across conditions since the number of tracking data points was four times higher in the 4000 ms ISI blocks than in the 1000ms ISI blocks (Ballard, 2001). Finally, target hit RTs less than 100ms (also referred to as perseverations) were filtered out of the data since those responses were likely due to errant and impulsive responding (Conners, 2014).

3.1.10 DATA ANALYSIS

GCAs are specialized statistical techniques for summarizing longitudinal data with best-fit lines (or smooth curves) that characterize performance trends within observed time windows (Bollen, 2007; Byrne & Crombie, 2003; Kristjansson, Kircher, & Webb, 2007). These lines are defined by estimates that take into account performance differences between longitudinal measurement 'clusters' (i.e., within-individual groups), as well as the individual differences in performance within clusters (i.e., between-individual variation) (Barr et al., 2013; Curran et al., 2010). Estimates are calculated within a mixed-effects framework in which regression models specify a small number of pre-determined parameters (i.e., fixed-effects), the random probability distributions around those effects (i.e., random-effects), as well as the orthogonal (i.e., statistically independent) time factors that represent the trajectory of the estimates over time (Baayen et al., 2008). The resulting *growth curves* plotted from these estimates can come in different forms, the most basic of which is the *random intercept* model in which best-fit lines characterize the overall differences between the repeated measurement groups as stable or flat, thus signifying no change in time (Barr, 2013). This model can then be expanded by considering the family of polynomial functions that can characterize performance trends as systematically increasing or decreasing over time (i.e., the *linear* model), or curving at some point and to varying degrees (i.e., the *quadratic* model, the *cubic* model, etc.) (Diallo et al., 2014).

As implied in the above description, developing mixed-effects models for GCA is a multi-step process. It involves building and comparing different models to find the one that explains the trajectory of the growth curve that best fits the data, given the theoretical considerations of the study at hand (Peugh, 2010). The most basic model often includes

baseline terms for the intercept, time order factors, and random factors used in the analysis, while subsequent models add other fixed factors to each of the time order terms. Furthermore, evaluating models involves assessing how well each of the compared models fit the data using chi-square likelihood ratio tests with degrees of freedom equal to the number of parameters used in the analysis (Baayen et al., 2008; Long, 2012). Once a model is selected, its parameter estimates and associated p-values are used to interpret the fixed-effects of the chosen growth curve (Winter, 2013).

In this study, we used R 4.2.2 (R Core Team, 2012) and the lme4 (Bates et al., 2014) statistical package to perform GCAs to analyze the longitudinal performance data of this study. The GCAs consisted of terms for linear (i.e., $time^1$) and quadratic (i.e., $time^2$) time orders, terms for session (Single-Task, Multi-Task) and ISI (1000ms, 2000ms, 4000ms), and a random effect of participants on the intercept. We attempted to fit more complex random factor terms to the data, but these models did not converge. For our analysis, we compared four models that increased in complexity: the *Base* model only included baseline time terms and the random factor (with no fixed terms representing the interaction of session and ISI terms) (Model 1, Table A.1); the *Intercept* model built upon this by adding an interaction of the fixed effects on the intercept (Model 2, Table A.1); the *Linear* model built upon this by adding the interaction on the linear (i.e., $time^1$) time order (Model 3, Table A.1); and the *Quadratic* model built upon this by adding the interaction on the quadratic (i.e., $time^2$) time order (Model 4, Table A.1). We then compared each of these models, using maximum likelihood estimates, to determine the best time order model to use for the analysis (Long, 2012). Following Long's recommendations, we considered a model to provide a better fit than a simpler model

using a $p < .1$ criterion. Note that the Base model included Intercept, Linear, and Quadratic terms/predictors but not any of the fixed factors we manipulated. This means that the Base model could capture the temporal characteristics of the data that were constant across session and ISI conditions. Finally, we interpreted the chosen model by visually inspecting the plot and coefficients that fit the selected model.

3.2 RESULTS

Data from seven participants were removed due to hardware issues that caused errors in data collection. Data from the remaining 56 participants (age: $M = 20.14$, $SD = 1.51$) were used for analysis. Of the 56 participants, 39 were female (age: $M = 20.1$, $SD = 1.33$) and 17 were male (age: $M = 20.2$, $SD = 1.89$).

3.2.1 CPT MEASURES

Target Hit Rate: The Intercept model provided a marginally significantly better fit of the data compared to the other models (Table B.1), $\chi^2(5) = 10.38$, $p = < .065$. Inspection of the model coefficients (Table B.2) and visual inspection of the graph (Figure 3.4a) show that: 1. target hit rate was not significantly different between sessions, but target hits during the faster ISIs (i.e., 1000ms and 2000ms) were numerically (but not significantly) lower (i.e., worse performance) in the Multi-Task session; and 2. target hits during all ISIs in both sessions may fit significant positive *linear* trajectories in which they steadily increased (i.e., improved) during the first several blocks, and then slightly leveled off during the end blocks.

Omissions: None of the models provided a significantly better fit of the data than the Base model (Table C.1), therefore we chose the Base model. Inspection of the model coefficients (Table C.2) and visual inspection of the graph (Figure 3.4b) show that: 1. omissions were not significantly different during all ISIs between sessions; and 2.

omissions during all ISIs in both sessions may fit significant negative *linear* trajectories in which they steadily decreased (i.e., improved performance) during the first several blocks, and then slightly leveled off during the end blocks.

Commissions: The Quadratic model provided a marginally significant better fit of the data (Table D.1), $\chi^2(5) = 9.67$, $p = <.09$. Inspection of the model coefficients (Table D.2) and visual inspection of the graph (Figure 3.4c) show that: 1. commissions were significantly worst during ISIs in the Multi-Task session; and 2. commissions during the 1000ms and 4000ms ISIs in the Multi-Task session may fit significant positive *quadratic* (curved) trajectories in which they initially decreased (i.e., improved performance) during the first several blocks, and then increased (i.e.., performance worsened) towards the end blocks.

CRs: The Quadratic model provided a marginally significant fit of the data (Table E.1), $\chi^2(5) = 10.42$, $p = <.065$. Inspection of the model coefficients (Table E.2) and visual inspection of the graph (Figure 3.4d) show that: 1. CRs were significantly worse during all ISIs in the Multi-Task session; and 2. CRs during the 1000ms and 4000ms ISIs in the Multi-Task session may fit significant negative *quadratic* trajectories in which they initially increased (i.e., improved performance) during the first several blocks, and then decreased (i.e., performance worsened) towards the end blocks.

Reaction Time: The Linear model provided a significantly better fit of the data compared to the other models (Table F.1), $\chi^2(5) = 11.64$, $p = <.05$. Inspection of the model coefficients (Table F.2) and visual inspection of the graph (Figure 3.5) show that: 1. RTs in both session were significantly best during the 1000ms ISI and significantly worst during the 4000ms ISI, and were significantly worst overall across all ISIs in the

Multi-Task session; and 2. RTs across all ISIs in both sessions may fit marginally significant *positive* linear trajectories in which they steadily increased (i.e., worsened performance) throughout the session.

3.2.2 SDT MEASURES

Sensitivity: The Quadratic model provided a significant fit of the data (Table G.1), $\chi^2(5) = 11.68$, $p = < .04$. Inspection of the model coefficients (Table G.2) and visual inspection of the graph (Figure 3.6a) show that: 1. sensitivity was significantly worse during all ISIs in the Multi-Task session; and 2. sensitivity during the 1000ms and 4000ms ISIs in the Multi-Task session may fit significant negative *quadratic* trajectories in which they initially increased (i.e., improved performance) during the first several blocks, and then decreased (i.e., performance worsened) towards the end blocks.

Response Bias: The Intercept model provided a significantly better fit of the data compared to the other models (Table H.1), $\chi^2(5) = 11.12$, $p = < .05$. Inspection of the model coefficients (Table H.2) and visual inspection of the graph (Figure 3.6b) show that: 1. response bias across all ISIs in both sessions was overall liberal; and 2. response bias during the 4000ms ISI in the Multi-Task session was significantly more liberal (i.e., further from '0') than in the other ISIs.

3.2.3 TRACKING MEASURE

Average Tracking Distance: The Linear model provided a marginally significant fit of the data (Table I.1), $\chi^2(2) = 2.77$, $p = <.10$. Inspection of the model coefficients (Table I.2) and visual inspection of the graph (Figure 3.7) show that: 1. distance during all ISIs may fit significant *positive* quadratic trajectories in which average distance initially decreased (i.e., improved performance) during the first several time blocks, and then slightly increased (i.e., performance worsened) towards the end blocks; and 2.

distance during the 1000ms ISI is significantly worst during the first few blocks, and then sharply decreases (i.e., performance improves) and levels off towards the end blocks.

3.2.4 QUESTIONNAIRE RESPONSES

We analyzed select survey responses with repeated measures ANOVAs using the anova_test function from the rstatix package (version 0.7.0; Kassambara, 2021) in R (R Core Team, 2012). Groups in this analysis include *order*, consisting of Baseline (i.e., before Single-Task session), Single-Task session, and Multi-Task session levels. For Drowsiness Responses, we found a significant effect of *order*, $F(1.77, 97.62) = 7.36$, $p < 0.003$, with visual inspection of the boxplot (Figure 3.8a) showing highest levels of drowsiness reported in the Multi-Task session. For Fatigue Responses, we found a significant effect of *order*, $F(2, 110) = 29.22$, $p < 0.001$, with visual inspection of the boxplot (Figure 3.8b) showing highest levels of drowsiness reported in the Multi-Task session. Finally, for Difficulty Responses, we found a significant effect of *order*, $F(1, 55) = 79.81$, $p < 0.001$, with visual inspection of the boxplot (Figure 3.8c) showing highest levels of difficulty reported in the Multi-Task session.

3.3 DISCUSSION

There were three critical hypotheses for E1, and each applied to the measures in the Multi-Task session: the first (H1) was that performance would be worst across CPT-related measures than in the Single-Task session; the second (H2) was that performance would be worst across all measures during the faster compared to the slower ISIs; and the third (H3) was that GCAs would reveal non-linear changes in performance across all measures.

Overall, the results from the GCAs for target hits and omissions data revealed no significant differences in performance between each of the ISIs across sessions; however,

GCAs for commissions, CRs, sensitivity, and RT data revealed significantly worst performance during each of the ISIs in the Multi-Task session, especially during the 1000ms and 4000ms ISIs. This difference was also evident for the response bias data in which performance was significantly more ‘liberal’ during the 4000ms ISI in the Multi-Task session than during the other ISIs. Furthermore, GCAs for average distance data revealed that performance was significantly worse in the early blocks (i.e., blocks 1, 2, and 3) during the 1000ms ISI compared to the other ISIs (*note*: unlike the CPT measures that were present in both sessions, the driving-based track task measure was only present in the Multi-Task session). Moreover, self-report questionnaire data suggested that participants perceived higher levels of drowsiness, fatigue, and task difficulty after performing the Multi-Task session; however, these findings apply more generally to the overall conditions and not to the specific measures associated with the tasks of the conditions. This is a trivial point, since the limitations of self-report questionnaires are well documented in the literature (e.g., Razavi, 2001; Williams et al., 2017). Therefore, these results provide partial support for H1 and H2.

Interestingly, GCAs revealed that much of the data in E1 was characterized by different performance trajectories. For example, target hits and omissions data during each ISI across both sessions were characterized by significant negative linear trajectories in which performance steadily worsened with time-on-task. In contrast, commissions, CRs, and sensitivity data during the 1000ms and 4000ms ISIs in the Multi-Task session, and only during the 1000ms ISI in the Single-Task session, were characterized by marginally significant quadratic trajectories in which performance initially worsened and then improved towards the later blocks (i.e., blocks 4, 5, and 6); however, while

commissions data was characterized by positive trajectories, CRs and sensitivity data was characterized by negative trajectories (a difference due to the nature of the performance criteria for these measures, as trajectory valences may have different implications depending on the measure). Other measures were less uniform with their trends. For example, RT data during each ISI across both sessions was characterized by significant positive linear trajectories but appeared to have some marginal quadratic trends towards the later blocks. Response bias data during each ISI across both sessions was shown to only vary on the intercept (i.e., no decrement recorded). Lastly, average distance data during the 2000ms and 4000ms ISIs was characterized by significant positive quadratic trajectories, but during the 1000ms ISI it appeared to have a significant negative linear trajectory in the early blocks with a positive quadratic trend towards the later blocks. Therefore, these results provide partial support for H3.

These results provide evidence that our paradigm achieved both of its objectives. First, vigilance decrements were found in all but one measure (i.e., response bias). This suggests that the task structure used in this paradigm (i.e., fixed-order design, alternating ISI manipulation) was sufficient for eliciting decrements in performance, especially in the Multi-Task session in which the interactive effects between the two tasks during the different ISIs was the primary focus of interest. Second, performance changed as a function of ISI, although this change was not consistent for all measures. This suggests that while arousal and alertness levels can change as a function of stimulus presentation rates (Matthews et al., 2017; Silverstein et al., 2004), this change may be mediated by task-specific factors (Fisk & Schneider, 1981; Wickens, 2008). Third, non-linear changes in performance characterized several measures, and suggests that GCAs were effective

tools for characterizing vigilance performance within our paradigm (Boiteau et al., 2014; Devlin et al., 2023; Grech et al., 2009; Rann & Almor, 2022).

However, GCA trajectories were not consistent for all measures. Indeed, GCAs revealed 'traditional' linear vigilance decrement patterns across some measures, and quadratic trajectories across other measures. One likely explanation for this difference is that quadratic trajectories occur due to the presence of learning effects in which performance initially improves during the early blocks as participants learn how to perform the tasks (Brown, 2008; Lara et al., 2014; Shiffrin & Schneider, 1977), then worsens during the later blocks as they become fatigued from performing the tasks over several minutes (Pattyn et al., 2008). More so, some tasks may be more susceptible to these effects than others (Parasuraman & Giambra, 1991; Tiwari et al., 2009; Wickens, 2002). For example, tasks that require more executive control in their execution may involve early learning effects (Fisk & Scerbo, 1987), this would explain why the measures associated with the primary response inhibition and tracking tasks exhibited quadratic trajectories in the Multi-Task session, while the secondary responding task exhibited linear trajectories (i.e., since it is highly automated and requires little executive control).

We will defer a full consideration of the theoretical implications of our results for the general discussion. That said, given these results, we made two changes to our paradigm. First, we removed the less informative 2000ms ISI condition to potentially increase the power of our paradigm to capture fine grained changes in performance across tasks. Second, we included an additional practice session to better acclimate

participants to the tracking task, and thus potentially reduce learning effects in the more difficult Multi-Task session.

Table 3.1 Description of Measures.

Measure	**Source**	**Definition**	**Psychological processes**
Hits	CPT	Correctly pressing foot pedal when target is presented	Detection accuracy, Prepotent motor response
Omissions	CPT	Failing to press the foot pedal when target is presented (i.e., misses)	Inattentiveness, degradation in attention capacity
Commissions	CPT	Pressing the foot pedal when no target is present (i.e., false alarms)	Impulsivity, Failure to inhibition prepotent motor response
CRs	CPT	Correctly not pressing the foot pedal when no target is present	Detection accuracy, Response inhibition
RT	CPT	The time between correctly detecting a presented target and pressing the foot pedal	Prepotent motor response processing speed
d'	SDT	Sensitivity for accurately detecting target	Detection accuracy, most related to the prepotent motor response
β	SDT	Response strategy regarding speed-vs-accuracy tradeoffs	Detection accuracy, Prepotent motor response, and Failure to inhibit prepotent motor response
Distance	Tracking task	Average distance between target and cursor during the tracking task	Detection accuracy, Prepotent motor response processing speed

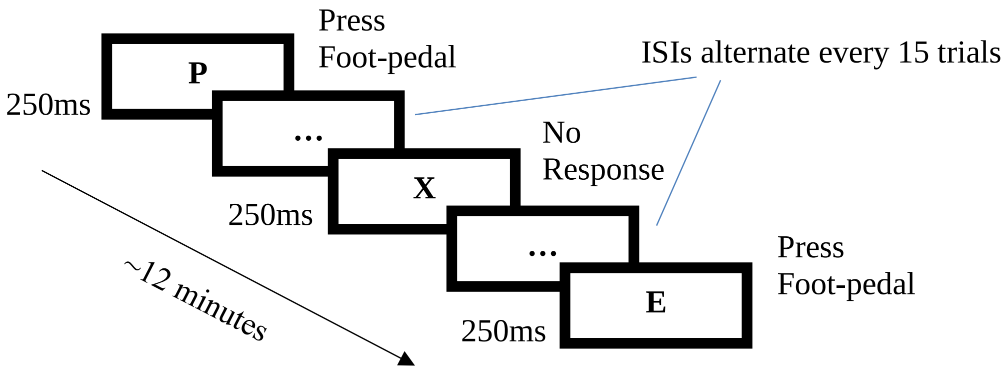

Figure 3.1: CCPT. Participants are required to press the foot-pedal when they detect a target letter, and withhold response when a non-target letter is detected.

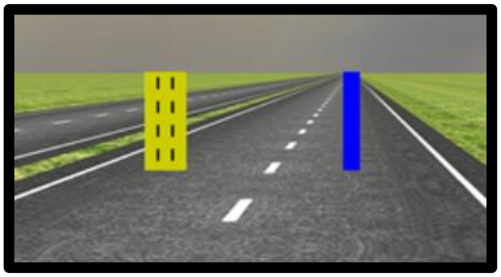

Figure 3.2: ConTRe Task. Participants are required to track the moving yellow target using the blue cursor they control using a steering wheel.

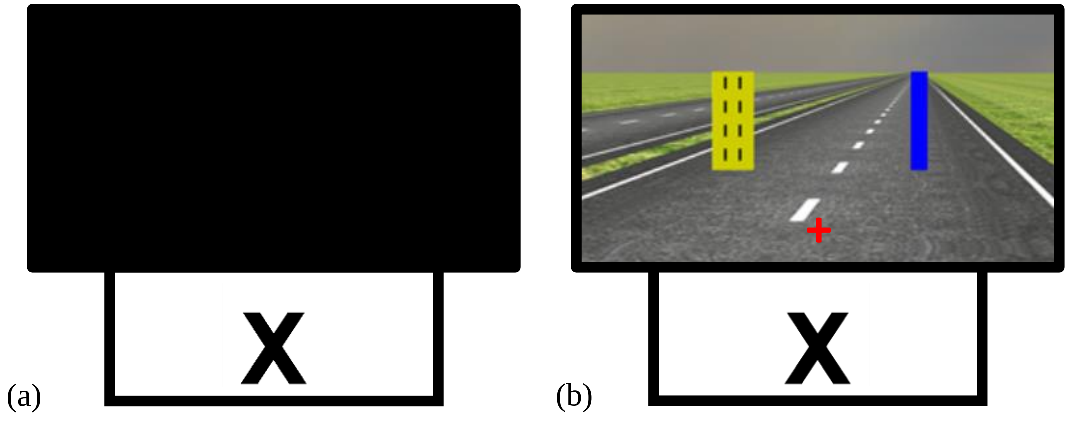

Figure 3.3: Setup – (a). Single-Task session in which the larger monitor was turned off and the CPT was displayed on the lower monitor; (b). Multi-Task session in which both monitors displayed their associated tasks. Approximate line of sign shown as red cross.

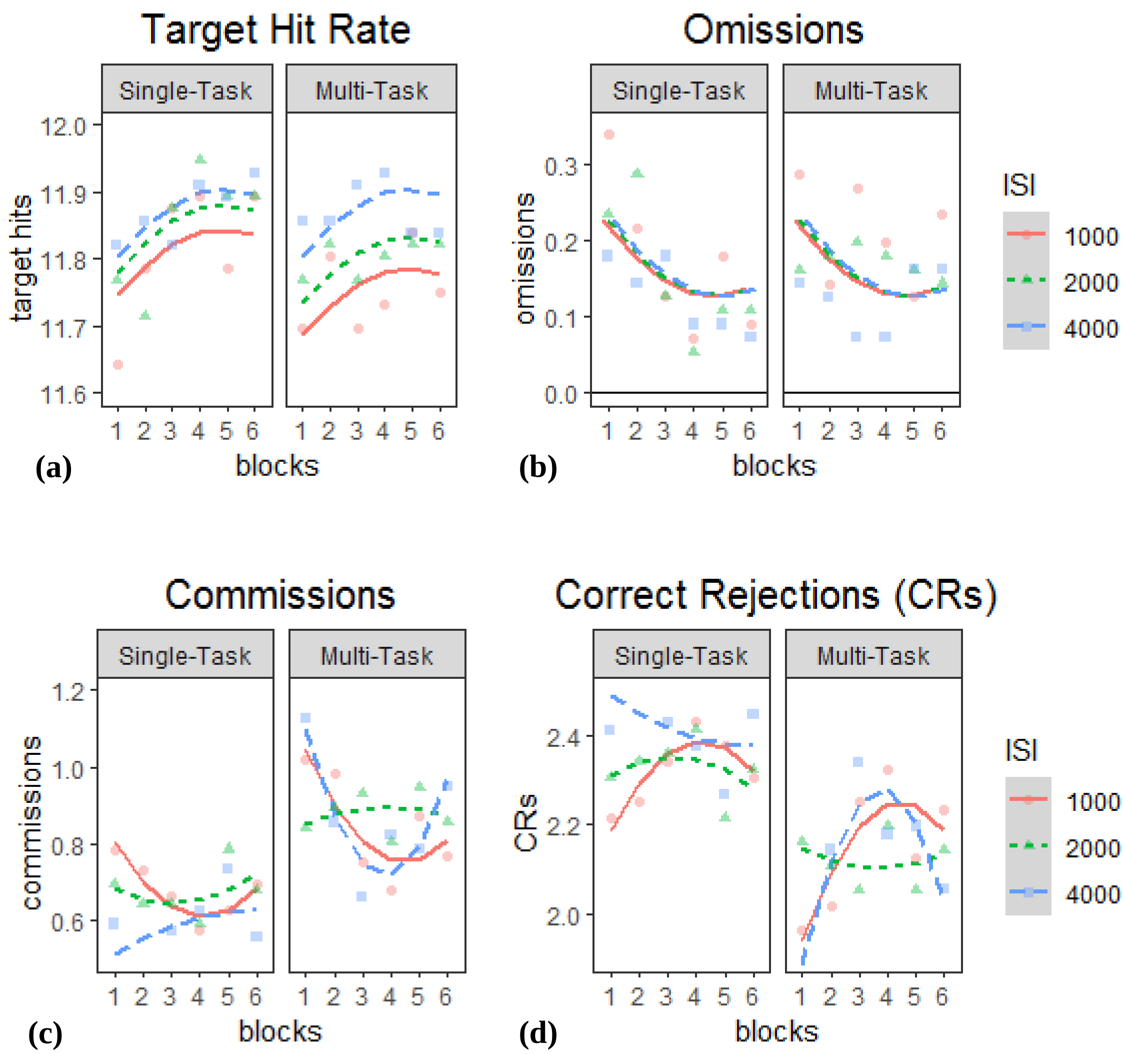

Figure 3.4: Growth curve analysis of CPT measures averaged per Block per ISI in E1 – (a). Target Hit Rate; (b). Omissions; (c). Commissions; and (d). CRs.

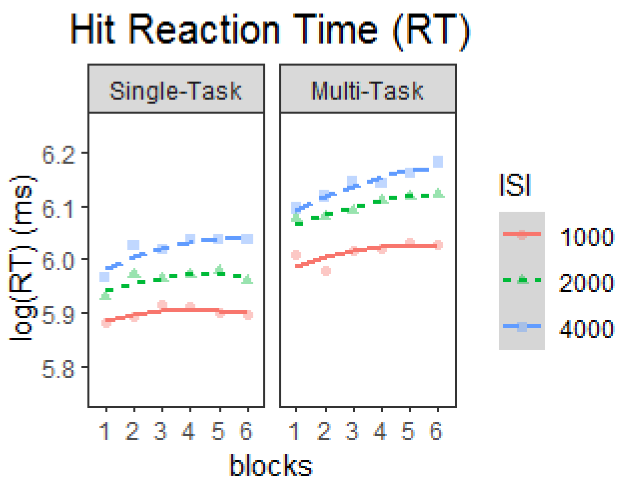

Figure 3.5: Growth curve analysis of Reaction Time (RT) averaged per Block per ISI in E1.

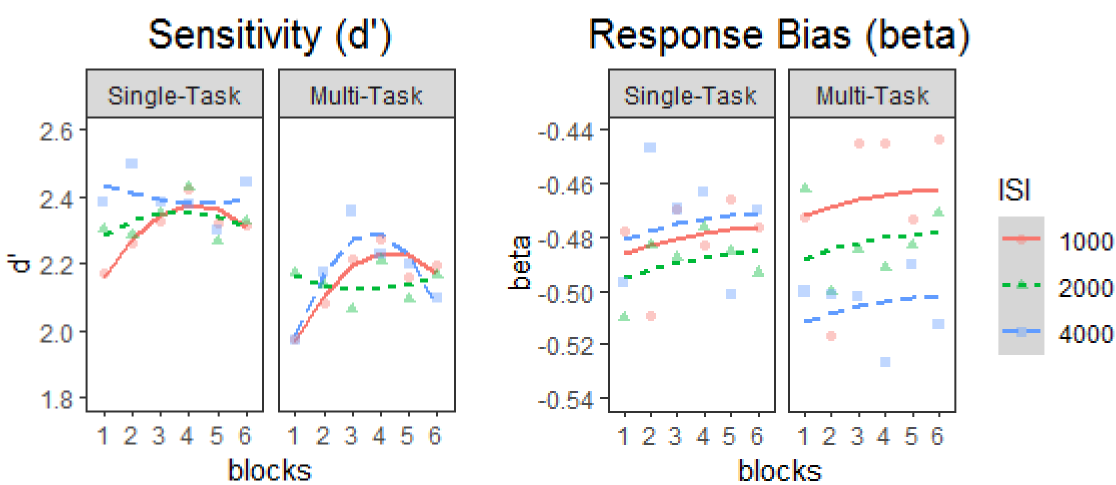

Figure 3.6: Growth Curve Analysis of SDT measures averaged per Block per ISI in E1 – (a). Sensitivity; and (b). Response Bias.

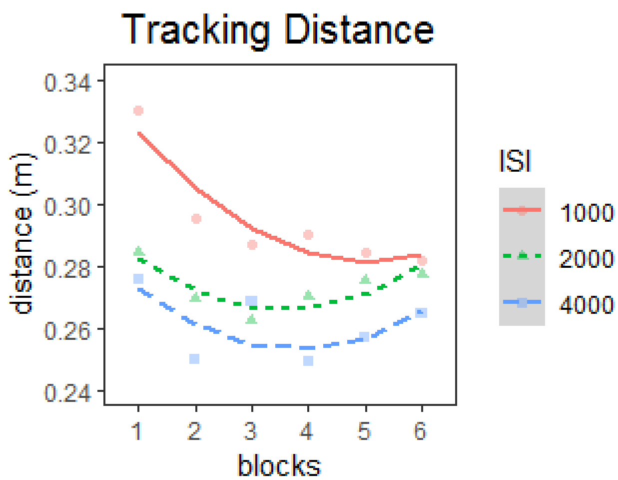

Figure 3.7: Growth Curve Analysis of Tracking Distance averaged per Block per ISI in E1.

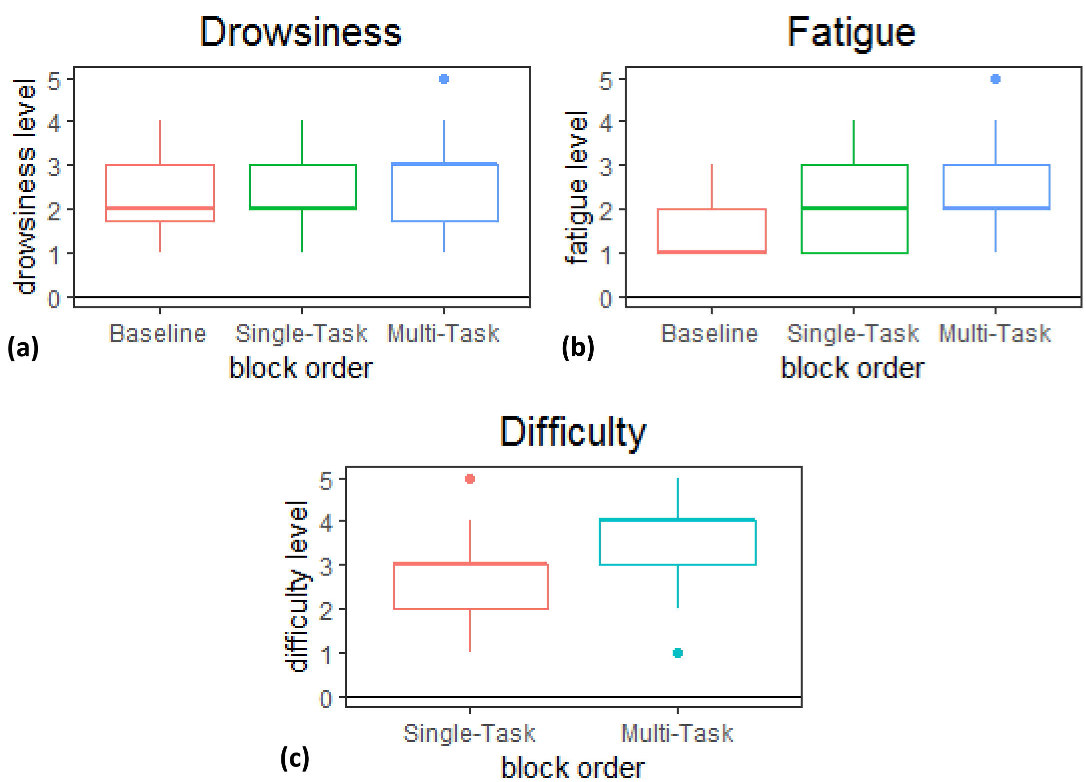

Figure 3.8: Boxplots for questionnaire responses in E1 – (a). Drowsiness; (b). Fatigue; and (c). Perceived Difficulty. Note: The top and bottom lines of the box represent the upper and lower quartiles of the data, respectively; the darker colored line represents the median value of the data; and the vertical lines attached to the boxes represent the range of the scores, while the colored dot represents the outlier for the maximum and minimum scores.

CHAPTER 4

EXPERIMENT 2

The purpose of Experiment 2 (E2) was to determine whether the E1 paradigm could be improved so as to further verify its effectiveness in eliciting and measuring changes in performance that occur when operators performed demanding multitasking activities.

Similar to E1, the first critical hypothesis (H1) is that performance will be worst across CPT-related measures in the Multi-Task session, and the second critical hypothesis (H2) is that performance will be worst across all measures in this session during the faster compared to the slower ISIs. However, our third critical hypothesis (H3) is that GCAs will reveal quadratic performance trajectories for the measures associated with the primary response inhibition task (i.e., commissions and CRs), and our fourth critical hypothesis (H4) is that GCAs will reveal linear performance trajectories for the measures associated with the secondary responding task (i.e., target hits and RT) and the tracking task (i.e., average distance). These last two hypotheses follow from the E1 findings which suggest that quadratic trajectories may characterize more demanding tasks in which early learning effects occur which eventually give way to later vigilance decrements, while linear trajectories may characterize less demanding automatized tasks.

4.1 METHODS

We made two key changes to the E1 paradigm. First, and most importantly, we removed the 2000ms ISI condition from the within-session difficulty manipulation. As a

result, there were more trials (20 instead of 15) in the remaining 1000ms and 4000ms ISIs, which increases power while at the same time results in fewer trials per block (40 instead of 45) and per session (240 instead 270), thereby reducing the overall time of each session in the paradigm (approximately 11instead of 12 minutes). We predict this duration reduction will have a negligible effect since observed changes in performance occurred throughout the experiment blocks in both sessions. Second, we implemented an additional practice trial in which participants performed the driving-based tracking task in the absence of the CPT for approximately 30 seconds. This new trial occurred after participants completed the Single-Task session, and before they begun the Multi-Task session practice trial.

We also utilized empirically validated and theoretically motivated questionnaires to assess subjective states of stress before and after each session (using the SSSQ; Helton, 2004), and perceived workload ratings after each session (using the NASA-TLX; Hart, 2006). All other methods in E2 (i.e., participant selection criteria, hardware and software, tasks, setup, data preparation, and data analysis) were the same as in E1.

4.2 RESULTS

4.2.1 CPT MEASURES

Target Hit Rate: The Intercept model provided a significantly better fit of the data compared to the other models (Table J.1), $\chi^2(3) = 25.4$, $p = <.001$. Inspection of the model coefficients (Table J.2) and visual inspection of the graph (Figure 4.1a) show that: 1. Target hits were significantly worse during the 1000ms ISI in the Multi-Task session; and 2. target hits during both ISIs across sessions may fit significant negative *quadratic* trajectory in which they steadily increased (i.e., improved performance) during the first

several blocks, and then steadily decreased (i.e., worsened performance) during the end blocks.

Omissions: The Intercept model provided a significantly better fit of the data compared to the other models (Table K.1), $\chi^2(3) = 18.47$, $p = < .001$. Inspection of the model coefficients (Table K.2) and visual inspection of the graph (Figure 4.1b) show that: 1. omissions were significantly worse during the 1000ms ISI in the Multi-Task session; and 2. omissions across all ISIs in both sessions may fit significant negative *quadratic* trajectories in which they steadily decreased (i.e., improved performance) during the first several blocks, and then steadily increased (i.e., worsened performance) during the end blocks.

Commissions: The Intercept model provided a significantly better fit of the data compared to the other models (Table L.1), $\chi^2(3) = 11.13$, $p = <.02$. Inspection of the model coefficients (Table L.2) and visual inspection of the graph (Figure 4.1c) show that: 1. commissions were slightly but not significantly worst during the 4000ms ISI in the Multi-Task session; and 2. commissions across conditions may initially fit significant negative *linear* trajectories in which they sharply decreased (i.e., improved performance) during the first several blocks, and then positive *quadratic* trajectory in which they increased (i.e.., worsened performance) in a curved fashion towards the end blocks.

CRs: The Quadratic model provided a significantly better fit of the data compared to the other models (Table M.1), $\chi^2(3) = 11.71$, $p = <.01$. Inspection of the model coefficients (Table M.2) and visual inspection of the graph (Figure 4.1d) show that: 1. CRs slightly but not significantly worse during the 4000ms ISI in the Multi-Task session; and 2. CRs during both ISIs across sessions may initially fit significant positive *linear*

trajectories in which they sharply increased (i.e., improved performance) during the first several blocks, and then negative *quadratic* trajectory in which they decreased (i.e., worsened performance) in a curved fashion towards the end blocks.

Reaction Time: The Liner model provided a significantly better fit of the data compared to the other models (Table N.1), $\chi^2(3) = 19.61$, $p = <.001$. Inspection of the model coefficients (Table N.2) and visual inspection of the graph (Figure 4.2) show that: 1. RTs in both session were significantly best during the 1000ms ISI and significantly worst during the 4000ms ISI, and were significantly worst overall across all ISIs in the Multi-Task session; and 2. RTs across all ISIs in both sessions may fit marginally significant *positive* linear trajectories in which they steadily increased (i.e., worsened performance) throughout the session.

4.2.2 SDT MEASURES

Sensitivity: The Intercept model provided a significantly better fit of the data compared to the other models (Table O.1), $\chi^2(3) =16.43$, $p = < .001$. Inspection of the model coefficients (Table O.2) and visual inspection of the graph (Figure 4.3a) show that: 1. sensitivity was significantly worst during both ISIs in the Multi-Task session; and 2. sensitivity may initially fit significant positive *linear* trajectory in which they sharply increased (i.e., improved performance) during the first several blocks, and then negative *quadratic* trajectory in which they decreased (i.e., worsened performance) in a curved fashion towards the end blocks.

Response Bias: The Intercept model provided a significantly better fit of the data compared to the other models (Table P.1), $\chi^2(3) = 9.68$, $p = < .03$. Inspection of the model coefficients (Table P.2) and visual inspection of the graph (Figure 4.3b) show that: 1. response bias during both ISIs across sessions was overall liberal; and 2. response bias

during the 4000ms ISI was significantly more liberal (i.e., further from ‘0’) than during the 1000ms ISI in the Multi-Task session.

4.2.3 TRACKING MEASURE

The Intercept model provided a significantly better fit of the data compared to the other models (Table Q.1), $\chi^2(1) = 44.5$, $p = <.001$. Inspection of the model coefficients (Table Q.2) and visual inspection of the graph (Figure 4.4) show that: 1. average distance was overall worst during the 1000ms ISI; and 2. average distance during both ISIs may fit significant positive *linear* trajectories in which average distance steadily increased (i.e., worsened performance) throughout the session.

4.2.4 QUESTIONNAIRE RESPONSES

We analyzed select survey responses with repeated measures ANOVAs using the anova_test function from the rstatix (version 0.7.0) library in R (R Core Team, 2012). For SSSQ Subjective State Responses, we found a significant effect of *order*, $F(1.71, 160.06) = 164.13$, $p < 0.001$, with visual inspection of the boxplot (Figure 4.5) showing highest levels of distress in the Single-Task subjective response session. For NASA-TLX workload Responses, we found a significant effect of *order*, $F(1, 59) = 99.86$, $p < 0.001$, with visual inspection of the boxplot (Figure 4.6) showing highest levels of workload reported in the Multi-Task session.

4.3 DISCUSSION

There were four critical hypotheses for E2, and each applied to the measures in the Multi-Task session: the first (H1) was that performance would be worst across CPT-related measures than in the Single-Task session; the second (H2) was that performance would be worst across all measures during the faster compared to the slower ISIs; the third (H3) was that GCAs would reveal quadratic performance trajectories for the

measures associated with the primary response inhibition task (i.e., commissions and CRs); and the fourth (H4) was that GCAs would reveal linear performance trajectories for the measures associated with the secondary responding task (i.e., target hits and RT) and the tracking task (i.e., average tracking distance).

Overall, the results from the GCAs revealed that performance was worst in the Multi-Task session for the target hits and omissions data during the 1000ms ISI, and for the sensitivity and RT data during the 4000ms ISI, while there were no significant differences in performance during both ISIs in the commissions and CRs data. GCAs for the response bias data revealed that participants were significantly more liberal in their responses during the 4000ms ISI in the Multi-Task session. GCAs for the average distance data revealed that performance was significantly worst during the 1000ms ISI. Finally, self-report questionnaire data from the SSSQ suggests that participants perceived similar subjective states of stress across the sessions (i.e., Baseline, Single-Task, Multi-Task), while data from the NASA-TLX suggest that participants perceived the Multi-Task session as more demanding than the Single-Task session. Therefore, these results provide partial support for H1 and H2.

As in E1, GCAs revealed a variation in performance trajectories during both ISIs across both sessions for most measures. For example: target hit and omissions data were characterized by symmetrical quadratic trajectories in which performance initially worsened and then gradually improved towards the later blocks; CR, sensitivity, and RT data were characterized by 'check mark'-shaped trajectories in which performance sharply improved in a linear fashion during the early blocks (i.e., blocks 1, 2, and 3), and then worsened in a curved (quadratic) fashion during the later blocks; response bias data

was characterized by flat trajectories (that were slightly, but not significantly, linear); and lastly, average distance data was characterized by linear trajectories in which performance steadily worsened with time-on-task. Therefore, these results provide full support for H3, and partial support for H4.

These results show that the changes we made to our paradigm did in fact improve its effectiveness in eliciting and measuring the performance changes that occur when operators perform demanding multitasking activities. First, removing the 2000ms ISI condition revealed significant differences in performance between the 1000ms and 4000ms ISIs across most measures. Like E1, these differences were inconsistent across measures, however there was no difference between ISIs for the response inhibition measures (i.e., commissions and CRs). These results further support the role of target presentation rate on modulating arousal and alertness on certain tasks (Matthews et al., 2017; Silverstein et al., 2004), and suggest that these effects may be mitigated during more demanding tasks, such as the primary response inhibition task, since it requires more executive control to perform and is thus more cognitively demanding. Second, including the additional tracking practice session resulted in tracking data that was characterized by linear trajectories in which performance progressively degraded over time. This is a marked difference from E1 in which tracking data was characterized by quadratic trajectories, especially during the 1000ms ISI. This suggests that participants benefited from the additional practice, and that our paradigm may be especially susceptible to task attenuation effects during the early session blocks, which can be reduced or eliminated through practice (Parasuraman & Giambra, 1991).

One interesting finding was that the inconsistencies in growth curve trajectories continued to persist across several of the measures in E2, albeit differently than those revealed in E1. For example, in E1, the primary inhibition task measures were characterized by quadratic trajectories, while the secondary responding task measures were characterized by linear trajectories; however, in E2, both sets of measures were characterized by quadratic trajectories, although the former was highly skewed during the early blocks while the latter was symmetrically curved throughout the session. These findings, in addition to the tracking data findings, may suggest that, in addition to task demand and attenuation, workload mitigation strategies could have also influenced performance across measures (Epling et al., 2016; Finomore et al., 2009; Kurzban et al., 2013). Specifically, the extra time given for the participants to practice the tracking task may have reduced the amount of control required in its performance. As a result, participants may have allocated more attentional resources towards performing the more demanding CPT than the tracking task, and thus increased the task attenuation effects observed during the early blocks in the session.

Again, we will defer a full consideration of the theoretical implications of our results for the general discussion. That said, given these results, we proposed two final changes to our paradigm. First, and most importantly, we included an additional verbal task towards the later blocks of the Multi-Task session to increase demand and further determine whether participants engage in workload mitigation strategies. Second, we allowed participants as much time as they needed to practice each of the tasks.

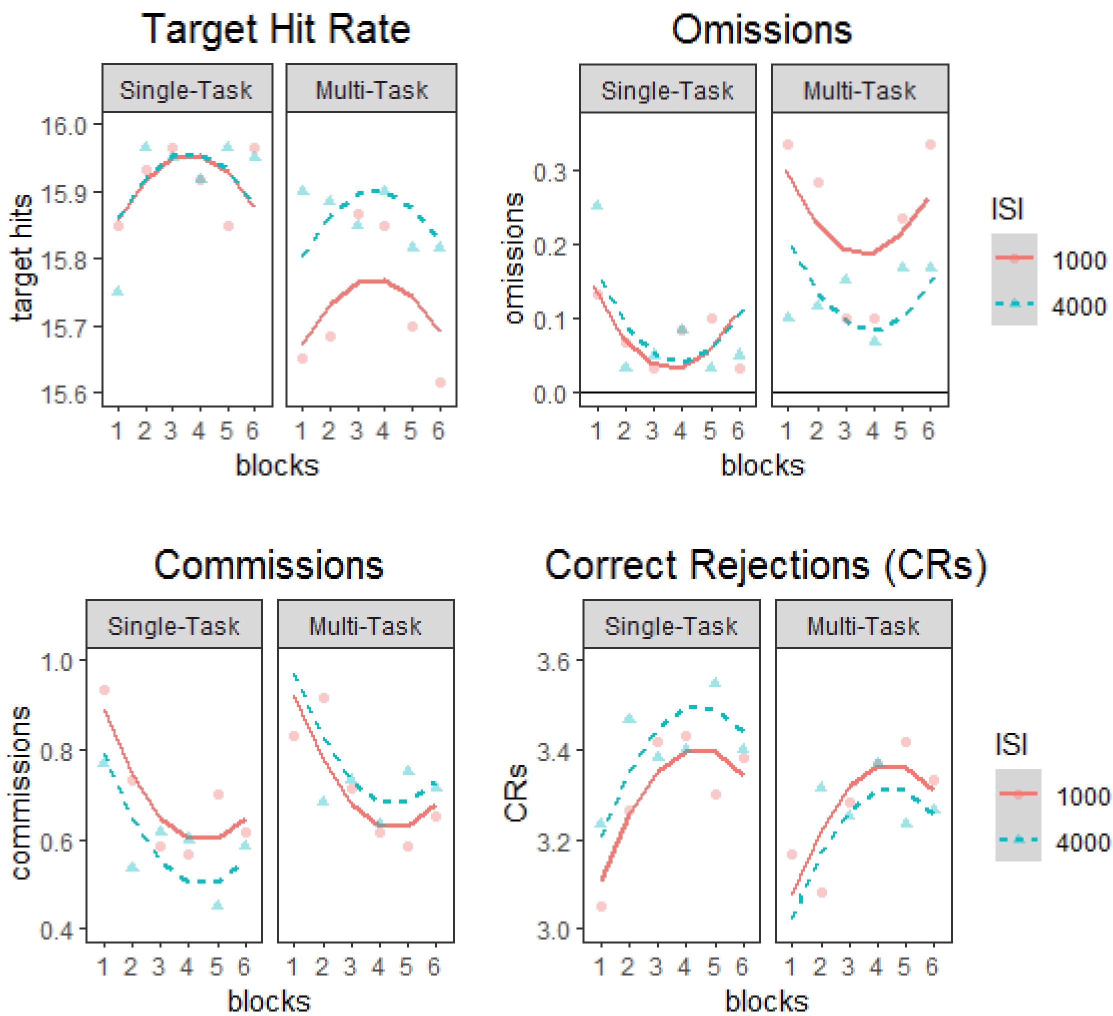

Figure 4.1: Growth curve analysis of CPT measures averaged per Block per ISI in E2 – (a) Target Hit Rate, (b) Omissions, (c) Commissions, (d) CRs, (e) Sensitivity, and (f) Response Bias.

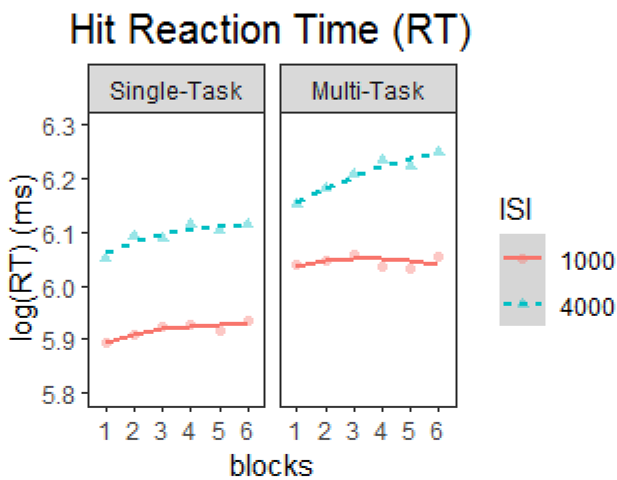

Figure 4.2: Growth curve analysis of Reaction Time (RT) averaged per Block per ISI in E2.

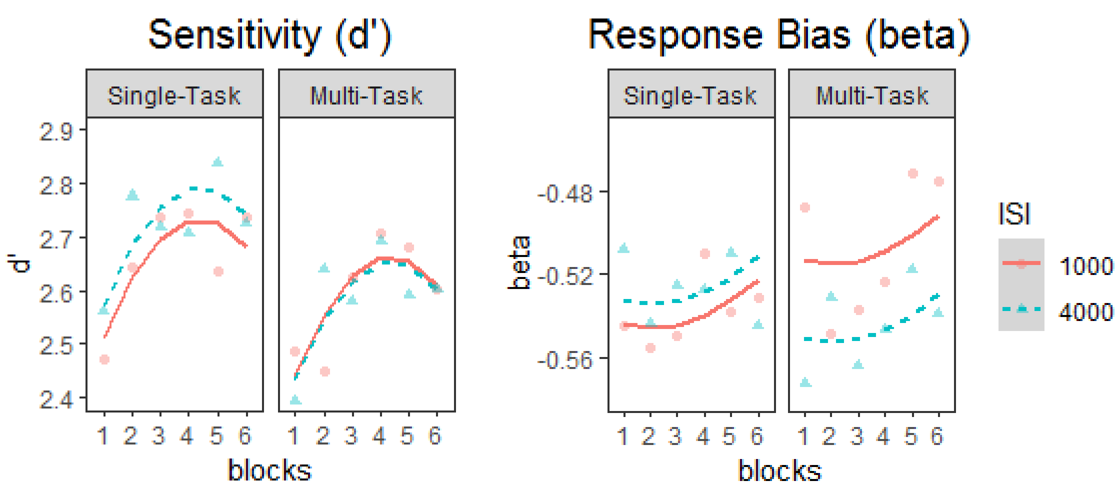

Figure 4.3: Growth curve analyses of SDT measures averaged per Block per ISI in E2 – (a) Sensitivity and (b) Response Bias.

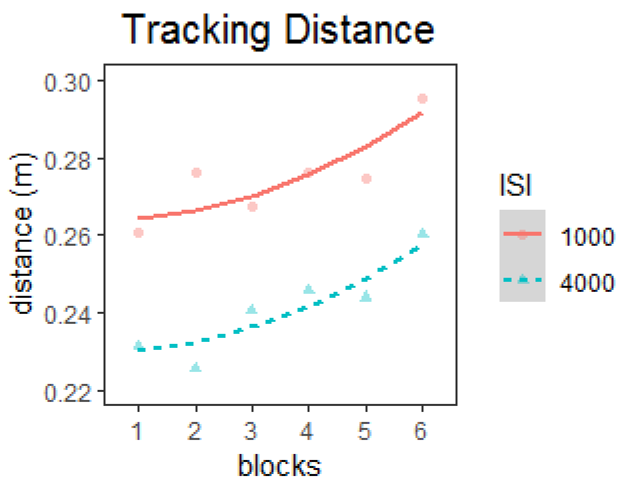

Figure 4.4: Growth curve analysis of Tracking Distance averaged per Block per ISI in E2.

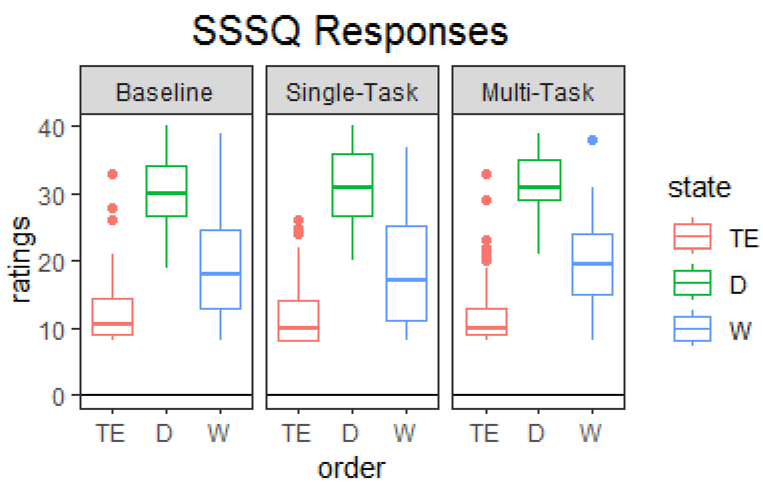

Figure 4.5: Boxplots for SSSQ questionnaire responses in E2 – Task Engagement (TE), Distress (D), and Worry (W). Note: The top and bottom lines of the box represent the upper and lower quartiles of the data, respectively; the darker colored line represents the median value of the data; and the vertical lines attached to the boxes represent the range of the scores, while the colored dot represents the outlier for the maximum and minimum scores.

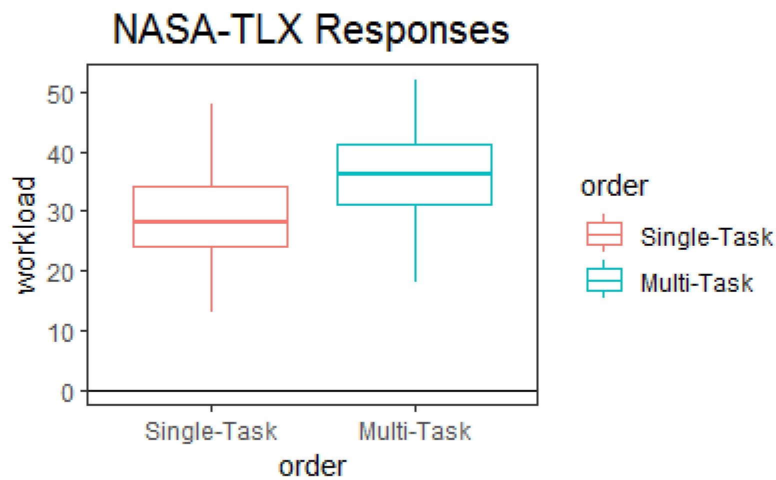

Figure 4.6: Boxplots for NASA-TLX questionnaire responses in E2. Note: The top and bottom lines of the box represent the upper and lower quartiles of the data, respectively; the darker colored line represents the median value of the data; and the vertical lines attached to the boxes represent the range of the scores.

CHAPTER 5

EXPERIMENT 3

The purpose of Experiment 3 (E3) was to determine whether and to what extent engaging in demanding verbal exchange can modulate performance across sustained attention tasks.

This is of particular significance since operators often engage in conversational activities throughout task performance (Atchley et al., 2011; Beede & Kass, 2006). For pilots, these activities may include continuously exchanging information over the radio with air traffic control at different times during flight (Dismukes & Nowinski, 2007; Falzon, 2009), while long haul truck drivers may use their CB radio to communicate upcoming road conditions to assist other drivers (Claveria et al., 2019). Regardless of the scenario, these operators must always remain vigilant since their primary goal is to properly operate and control their vehicles from the beginning of their journey to the end (Cabon et al., 1993; Wiggins, 2011).

Nevertheless, engaging in conversation may affect this ability in different ways (Almor, 2008; Cohen et al., 2017; Large et al., 2018; Regan et al., 2011; Wijayaratna et al., 2019; Zuercher, 1965). For example, Atchley et al., (2014) demonstrated that engaging in verbal tasks can improve a drivers’ ability to remain alert when driving for longer durations (over 60 minutes); and Rann & Almor (2022) demonstrated that engaging in verbal tasks can degrade a drivers’ performance when driving for shorter durations (under 5 minutes), especially when they produce compared to when they listen

to speech (Kubose et al., 2006). Given the results from our previous experiments, we chose to integrate aspects of Atchley et al. (2014) and Rann & Almor (2022) to determine how engaging in conversation can affect performance in our paradigm.

Similar to the previous experiments: the first critical hypothesis (H1) is that performance will be worst across CPT-related measures in the Multi-Task session, and the second critical hypothesis (H2) is that performance will be worst across all measures in this session during the faster compared to the slower ISIs. However, the third critical hypothesis (H3) is that GCAs will reveal quadratic performance trajectories for the measures associated with the primary inhibition (i.e., CRs and commissions) and secondary responding (i.e., target hits and RT) CPT tasks; the fourth critical hypothesis (H4) is that GCAs will reveal linear performance trajectories for the average distance tracking measure; and the fifth critical hypothesis (H5) is that average distance will be worst overall when participants respond compared to when they listen to verbal prompts. These last three hypotheses follow from the E2 findings which suggests that during demanding scenarios, operators strategically allocate attentional processing resources towards less cognitively demanding tasks (Kurzban et al., 2013).

5.1 METHODS

We made two key changes to the E2 paradigm: first, we included an interactive verbal task in which participants actively listened and then verbally responded to pre-recorded speech stimuli during the last two blocks (~4 minutes) of the Multi-Task session; and second, to potentially reduce some of the practice effects observed in the previous experiments, we allowed participants as much time as they needed to practice the tasks until they were ready to begin the experiment trials. Most other methods in E3

(e.g., participant selection criteria, hardware and software, setup, etc.) were the same as in E2.

The secondary verbal task in E3 was influenced by the Rann & Almor (2022) paradigm in which participants performed a driving-based tracking task while they actively listened and then immediately responded to pre-recorded verbal prompts presented via computer speakers. These prompts involved a speaker greeting the listener (i.e., participant), and then stating their name, occupation, and place of employment. For example, the speaker would say “Hello my name is Josh, and I am a teller at Wells Fargo.” Participants responded to the prompts by verbally greeting the speaker, and repeating back the details of what they heard as best as possible. For example, in response to the prompt described above, participants would say “Hello Josh, teller at Wells Fargo.” There were 12 verbal prompts presented throughout each conversation block at a rate of one prompt per every 20 seconds, and the mean duration of the verbal prompts were 4395 ms (SD = 771.58).

For E3, we incorporated this interactive verbal task structure and associated verbal stimuli. However, we made several changes to specifically examine the role of language processing in modulating performance within our vigilance paradigm. For example, we only presented verbal prompts during the last two blocks of the Multi-Task session (i.e., approximately four minutes) instead of throughout the entire session, as was done in Rann & Almor (2022). We did this to examine how engaging in a conversation-like verbal task could affect performance when session-related practice effects may start to subside (Rakauskas et al., 2004), and potentially at a time when participants may be more susceptible to task-related fatigue (Atchely et al., 2014; Gastaldi et al., 2014).

Furthermore, we increased the number of verbal prompts from 12 to 20 to increase the power for finding effects for the verbal task manipulation, and randomly jittered prompt onsets within- and between-participants to enhance their response preparation state and thus keep them "on their toes" (Smith et al. 2019; Wodka et al., 2009). We also presented prompts in fixed-order across participants (i.e., participants heard and responded to prompts in the same order), as we did not feel that randomizing the sound files played by the prompts would provide any additional benefit since we already randomized prompt onsets.

Importantly, we also ensured that the verbal prompts were consistently distributed between ISIs so that two prompts would always occur during the 1000ms ISI, and that eight prompts would always occur in the 4000ms ISI. This was done to mitigate concerns regarding the distribution of verbal prompts between the ISIs in the final two blocks of the Multi-Task session, since datapoints in the 4000ms ISI outnumbered the datapoints in the 1000ms ISI by a proportion of 4:1. Moreover, to ensure that participants had enough practice listening and responding to the verbal prompts, we included several prompts in the practice trials for both the tracking-only practice session (i.e., practice trial #2) and Multi-Task session (practice trial #3). Furthermore, we also allowed participants to practice each of the tasks as much as they felt that they needed before they moved on to the experiment trials.

To verify that participants complied with the verbal task, we video recorded participants during the Multi-Task session, and then reviewed the recordings to ensure that participants quickly and accurately responded to each of the verbal prompts. We also administered fill-in-the-blank verbal memory tests to participants after the Multi-Task

session. These required participants to fill in the missing pieces of information from a transcribed list of the verbal prompts they heard in the session. While we were not interested with how well they performed on this test, we did intend to motivate participants to fully engage with the verbal task by letting them know before the experiment that they would be tested on what they remembered from the prompts upon completing the Multi-Task session.

We examined the effects of verbal tasks on multitasking performance using two GCA-based approaches. First, we examined performance trajectories that occurred across each of the block time-points across all measures (using the same method used in the previous two experiments), making careful consideration of the trends that occurred in the last two blocks of the session. Second, we then examined performance trajectories that occurred in the driving-based tracking task data by treating each of the 20 verbal stimuli onsets as individual time-points of performance distributed across the last four minutes of the Multi-Task session. Each of these onsets corresponded to a spatial-temporal location within the driving simulator environment running the task. They were pre-coded into the OpenDS task files for each participant, so that they could easily be crossed-referenced and aligned with the average distance data. This allowed us to code verbal task data into segments in which no verbal tasks were present (none), segments in which participants actively listened to the verbal prompts (listen), and segments in which participants quickly responded to what they heard (respond). We used a liberal criterion of approximately five seconds to filter average distance data into the different segments instead of the approximately 4.5 seconds used in Rann & Almor (2022); this was done to accommodate for potential between-participant variation in response times, since review

of the video recordings showed that not all participants responded immediately to the verbal prompts. Given the power of the analyses in this experiment, we predicted that any variation lost as a result of such constraints would be minimal. Finally, since the average distance data was not normally distributed, and contained outliers that significantly further skewed the data, we removed outliers above three standard deviations from the median for the final analysis of performance during the verbal task blocks.

The overall analysis utilized the same GCA structure that was also used in E1 and E2 (i.e., Table A1). This consisted of terms for linear (i.e., $time^1$) and quadratic (i.e., $time^2$) time orders, terms for session (Single-Task, Multi-Task) and ISI (1000ms, 4000ms), and a random effect of participants on the intercept. However, to account for more dynamic changes in driving-based tracking task performance due to the effect of language processing, we filtered out the average distance data from blocks 1 through 4 so that only data from blocks 5 and 6 remained, replaced the block time-points used in the previous experiments with terms that corresponded to the 20 verbal onset time-points, and included additional terms for cubic (i.e., $time^3$) and quartic (i.e., $time^4$) time orders to account for dynamic changes in performance that may not be fully captured by linear and quadratic trajectories (i.e., Boiteau et al., 2014; Rann & Almor, 2022). We then gradually built models that increased in complexity: the *Base* model (Model 1, Table A2); the *Intercept* model (Model 2, Table A2); the *Linear* model (Model 3, Table A2); the *Quadratic* model (Model 4, Table A2); the *Cubic* model (Model 5, Table A2); and the *Quartic* model (Model 6, Table A2). All other methods were the same as previous experiments (e.g., methods to fit complex random factor terms, model comparisons, plot and coefficients inspections).

5.2 RESULTS

Data from 60 participants (age: $M = 20.13$, $SD = 1.38$), 44 females (age: $M = 20.1$, $SD = 1.24$), 16 males (age: $M = 20.2$, $SD = 1.77$), were used for analysis. This number was determined from the apriori power analysis performed in E2 (i.e., Faul et al., 2007).

5.2.1 CPT MEASURES

Target Hit Rate: The Quadratic model provided a significant fit of the data (Table R.1), $\chi^2(3) = 47.91$, $p = <.001$. Inspection of the model coefficients (Table R.2) and visual inspection of the graph (Figure 5.1a) show that: 1. target hits were significantly worst during both ISIs in the Multi-Task session, and worst overall in the 1000ms ISI; 2. target hits during the 1000ms ISI in the Multi-Task session may initially fit significant negative *quadratic* trajectory in which they decreased (i.e. worsened performance) in a curved fashion during the first several blocks, then significant negative *linear* trajectory in which they sharply decreased (i.e., worsened performance) towards the end blocks; and 3. target hits during the 4000ms ISI in the Multi-Task session may initially fit significant positive *linear* trajectory in which they slightly increased (i.e., improved performance) during the first several blocks, then significant negative *quadratic* trajectory in which they slightly decreased (i.e., worsened performance) towards the end blocks.

Omissions: The Quadratic model provided a significant fit of the data (Table S.1), $\chi^2(3) = 50.04$, $p = < .001$. Inspection of the model coefficients (Table S.2) and visual inspection of the graph (Figure 5.1b) show that: 1. omissions were significantly worst during both ISIs in the Multi-Task session, and worst overall in the 1000ms ISI; 2. omissions during the 1000ms ISI in the Multi-Task session may initially fit significant positive *quadratic* trajectory in which they decreased (i.e. improved performance) in a

curved fashion during the first several blocks, then a significant positive *linear* trajectory in which they sharply increased (i.e., worsened performance) towards the end blocks; and 3. omissions during the 4000ms ISI in the Multi-Task session may initially fit a significant negative *linear* trajectory in which they slightly decreased (i.e., improved performance) during the first several blocks, then a significant negative *quadratic* trajectory in which they slightly increased (i.e., worsened performance) towards the end blocks.

Commissions: The Quadratic model provided a significant fit of the data (Table T.1), $\chi^2(3) = 16.62$, $p = <.001$. Inspection of the model coefficients (Table T.2) and visual inspection of the graph (Figure 5.1c) show that: 1. commissions were significantly worse during both ISIs in the Multi-Task session, and worst overall during the 4000ms ISI; 2. commissions in the Multi-Task session during both ISIs may initially fit significant positive *quadratic* trajectories in which performance sharply decreased (i.e., improved performance) during the first several blocks and then increased (i.e., worsened performance) towards the middle blocks in a curve fashion, and then significant positive *linear* trajectories in which they sharply increased (i.e., worsened performance) towards the end blocks; and 3. commissions in the Single-Task session revealed interactions in which performance during the 1000ms ISI had a significant negative *linear* trajectory across blocks, while performance during the 4000ms ISI had a significant positive *linear* trajectory across blocks.

CRs: The Quadratic model provided a significant fit of the data (Table U.1), $\chi^2(3) = 17.79$, $p = <.001$. Inspection of the model coefficients (Table U.2) and visual inspection of the graph (Figure 5.1d) show that: 1. CRs were significantly worse across

all ISIs in the Multi-Task session, and worst overall during the 4000ms ISI; 2. CRs in the Multi-Task session during both ISIs may initially fit significant negative *quadratic* trajectories in which performance sharply increased (i.e., improved performance) during the first several blocks and then decreased (i.e., worsened performance) towards the middle blocks in a curve fashion, and then significant negative *linear* trajectories in which they sharply decreased (i.e., worsened performance) towards the end blocks; and 3. CRs in the Single-Task session revealed interactions in which performance during the 1000ms ISI had a significant positive *linear* trajectory across blocks, while performance during the 4000ms ISI had a significant negative *linear* trajectory across blocks.

Reaction Time: The Quadratic model provided a significant fit of the data (Table V.1), $\chi^2(3) = 31.6$, $p = <.001$. Inspection of the model coefficients (Table V.2) and visual inspection of the graph (Figure 5.2) show that: 1. RTs in both sessions were significantly best during the 1000ms ISI, and were significantly worst overall during both ISIs in the Multi-Task session; 2. RTs during the 4000ms ISI in both sessions may fit significant positive *linear* trajectories in which they steadily increased (i.e., worsened performance) throughout the session; and 3. RTs during the 1000ms ISI in both sessions may fit quadratic trajectories in which they exhibit opposite curving behavior throughout the sessions.

5.2.2 SDT MEASURES

Sensitivity: The Quadratic model provided a significant fit of the data (Table W.1), χ^2 (3) = 49.56, $p < .001$. Inspection of the model coefficients (Table W.2) and visual inspection of the graph (Figure 5.3a) show that: 1. sensitivity was significantly worse in the Multi-Task session during both ISIs, and worst overall during the 1000ms ISI; 2. sensitivity in the Multi-Task session during both ISIs may fit significant negative

quadratic trajectories in which performance slightly improved in the early blocks and then sharply worsened towards the end blocks; and 3. sensitivity in the Single-Task session revealed interactions in which performance during the 1000ms ISI had a significant positive linear trajectory across blocks, while performance during the 4000ms ISI had a significant negative linear trajectory across blocks.

Response Bias: The Linear model was the simplest model that provided a significantly better fit of the data compared to the other models (Table X.1), χ^2 (3) = 107.52, $p < .001$. Inspection of the model coefficients (Table X.2) and visual inspection of the graph (Figure 5.3b) show that: 1. response bias in both sessions across all ISIs was overall liberal; 2. response bias in the Multi-Task session during the 1000ms ISI was significantly more conservative, especially during the end blocks.

5.2.3 TRACKING MEASURE

The Intercept model provided a significantly better fit of the data compared to the other models (Table Y.1), $\chi^2(1) = 56.06$, $p = <.001$. Inspection of the model coefficients (Table Y.2) and visual inspection of the graph (Figure 5.4) show that: 1. average distance was significantly worst in the 1000ms ISI; 2. average distance during both ISIs may fit significant negative *linear* trajectories in which performance progressively worsened with time on task, and then slightly (but not significantly) leveled off towards the last two verbal task blocks (i.e., blocks 5 and 6).

5.2.3 TRACKING MEASURE DURING VERBAL TASK

The Quadratic model provided a significantly better fit of the data compared to the other models (Table Z.1), $\chi^2(3) = 8.40$, $p = <.04$. Inspection of the model coefficients (Table Z.2) and visual inspection of the graph (Figure 5.5) show that: 1. average distance may be slightly (but not significantly) worst overall during the 1000ms ISI in both verbal

task segments, and marginally significantly worst overall in the Respond segment, especially at the transition between blocks 5 and 6 (i.e., verbal onset prompts #10 and 11); 2. average distance during both ISIs may fit significant negative *quartic* trajectories in which distance slightly decreased (i.e., worsened) during the first few verbal task onsets and then slightly increased (i.e., improved) towards the middle onsets, and then slightly decreased and increased again towards the end of the verbal task onsets.

5.2.4 QUESTIONNAIRE RESPONSES

We analyzed select survey responses with repeated measures ANOVAs using the anova_test function from the rstatix (version 0.7.0) library in R (R Core Team, 2012). For SSSQ Subjective State Responses, we found a significant effect of *order*, $F(3.09, 182.57) = 195.81$, $p < 0.001$, with visual inspection of the boxplot (Figure 5.6) showing highest levels of distress in the Baseline session. For NASTA-TLX workload Responses, we found a significant effect of *order*, $F(1, 59) = 184.08$, $p < 0.001$, with visual inspection of the boxplot (Figure 5.7) showing highest levels of workload reported in the Multi-Task session.

5.2.5 VERBAL MEMORY TEST

We analyzed responses for the fill-in-the-blank verbal memory test by simply calculating the mean score across all participants. Of the 24 verbal prompts (i.e., 12 prompts for block 5, 12 prompts for block 6), participants correctly answered 7.125 questions on average.

5.3 DISCUSSION

There were five critical hypotheses for E3, and each applied to the measures in the Multi-Task session: the first critical hypothesis (H1) was that performance would be worst across CPT-related measures than in the Single-Task session; the second critical

hypothesis (H2) was that performance would be worst across all measures during the faster compared to the slower ISIs; the third critical hypothesis (H3) was that GCAs would reveal quadratic performance trajectories for the measures associated with the primary inhibition (i.e., CRs and commissions) and secondary responding (i.e., target hits and RT) CPT tasks; the fourth critical hypothesis (H4) was that GCAs would reveal linear performance trajectories for the average tracking distance measure; and the fifth critical hypothesis (H5) was that average distance would be worst overall when participants responded compared to when they listened to verbal prompts.

Overall, the results from the GCAs revealed that performance was worst in the Multi-Task session for most CPT measures, with target hits, omissions, and sensitivity data having worst performance during the 1000ms ISI, and commissions, RT, and CRs data having worst performance during the 4000ms ISI. GCAs for the response bias data revealed that participants were significantly more conservative in their responses during the 1000ms ISI. GCAs for the average distance data revealed that performance was significantly worst overall during the 1000ms ISI, and marginally significantly worst during the Respond segments in the late verbal task blocks (i.e., blocks 5 and 6). Finally, self-report questionnaire data from the SSSQ suggests that participants perceived similar (but not significantly different) subjective states of stress across the sessions (i.e., Baseline, Single-Task, Multi-Task), while data from the NASA-TLX suggests that participants perceived the Multi-Task session as more demanding than the Single-Task session. Therefore, these results provide full support for H1 and H5, and partial support for H2.

As in the previous experiments, GCAs revealed a variation in performance trajectories across most measures in the Multi-Task session. For example, target hits, omissions, and sensitivity were characterized by large 'check mark'-shaped quadratic trajectories with long linear tails during the 1000ms ISI in which performance initially improved during the early blocks, then sharply degraded towards the end blocks, and by small curved trajectories during the 4000ms ISI in which performance slightly improved during the early blocks, then degraded towards the end blocks; CRs, commissions, and sensitivity data were also characterized by ‘check mark’-shaped trajectories, however they were more parabolic in that performance during both ISIs sharply improved during the early blocks, then sharply degraded towards the later blocks; RT data was characterized by a large ‘check mark’-shaped quadratic trajectory with a long linear tail during the 1000ms ISI in which performance slowly improved in the blocks and then early sharply improved towards the later blocks, and a small curved trajectory during the 4000ms ISI in which performance slightly degraded during the early blocks and then improved towards the end blocks (this pattern was similar to the target hits, omissions, and sensitivity trajectories, however performance in these trended towards performance degradations, while RT data trended towards performance improvements); response bias data during both ISIs became more conservative towards the later blocks, however, the data was characterized by a large sharp positive linear trajectory during the 1000ms ISI, with a small curved quadratic trajectory during the 4000ms ISI; average distance data during both ISIs were characterized by positive linear trajectories in which performance decreased with time on task; Finally, during the verbal task blocks, average distance data was characterized by quartic trajectories in which performance fluctuated across verbal

task onsets in that they initially degraded during the early onsets of block 5, and then improved towards the end of block 5, then degraded during the early onsets of block 6 and improved towards the end of block 6. Therefore, these results provide full support for H3 and H4.

There were several interesting findings in E3. First, and perhaps most importantly, there was a clear and significant effect of the verbal task across most measures, especially during the 1000ms ISI. Specifically, target hits and inhibition task measures (i.e., commissions and CRs) degraded towards the last two verbal task blocks, while the RT improved, and response bias became more conservative. This falls in line with the resource depletion account of vigilance performance (Helton & Warm, 2008; Matthews et al., 1993), and suggests that response strategy can at least be partly attributed to shifts in resource availability that occur when cognitive load increases (i.e., Fox et al., 2007; Helton et al., 2009). However, the overall average distance measure appeared to show no significant effect of the verbal task, as performance progressively worsened with time on task in a similar linear fashion as it did in E2. This suggests that response strategy also shifted across tasks, such that the most highly automatized task (i.e., tracking) got the highest priority (and was thus less affected by the verbal task), while the least relevant task (i.e., secondary responding) got the lowest priority (and was thus most affected by the verbal task). This is particularly interesting since average distance during the last two blocks was characterized by quartic trajectories across the verbal task onsets. While this fluctuation in performance falls in line with the results from Rann & Almor (2022), which demonstrated dynamic changes in performance as drivers engaged in different aspects of conversation (e.g., listening, planning, speaking), these results differ in that

there were no significant differences in performance between conversation segments. This was perhaps due to the lack of power from our design, since there were only 12 prompts presented per block, and these were disproportionately divided between the ISIs (i.e., 4 prompts total for the 1000ms ISI, 8 prompts total for the 4000ms ISI). However, this could also be due to general difficulty of participants to *both* listen and respond while also simultaneously performing the CPT and tracking tasks.

The additional practice time allotted to participants for each of the E3 tasks may have helped to make these effects more discernable. This is especially the case for most of the CPT data in the Single-Task session which were characterized by interacting linear trajectories in which performance steadily improved during the 1000ms ISI and steadily worsened during the 4000ms ISI. In line with E2, this finding suggests that measures from highly practiced tasks are characterized by linear trajectory vigilance decrements, while tasks that require high levels of control are more prone to quadratic trajectories in which task attenuation and learning occur in the early blocks which give way to vigilance decrements towards the later blocks.

Together with the results from E1 and E2, these findings provide key insights for how resources are allocated during the course of sustained multitasking, and show that task automaticity and relevancy are significant factors involved with workload mitigation strategies (Fisk & Schneider, 1981; Koch et al., 2018; Monsell, 2003; Schumacher et al., 2001; Tsang & Chan, 2018; Wickens, 2008).

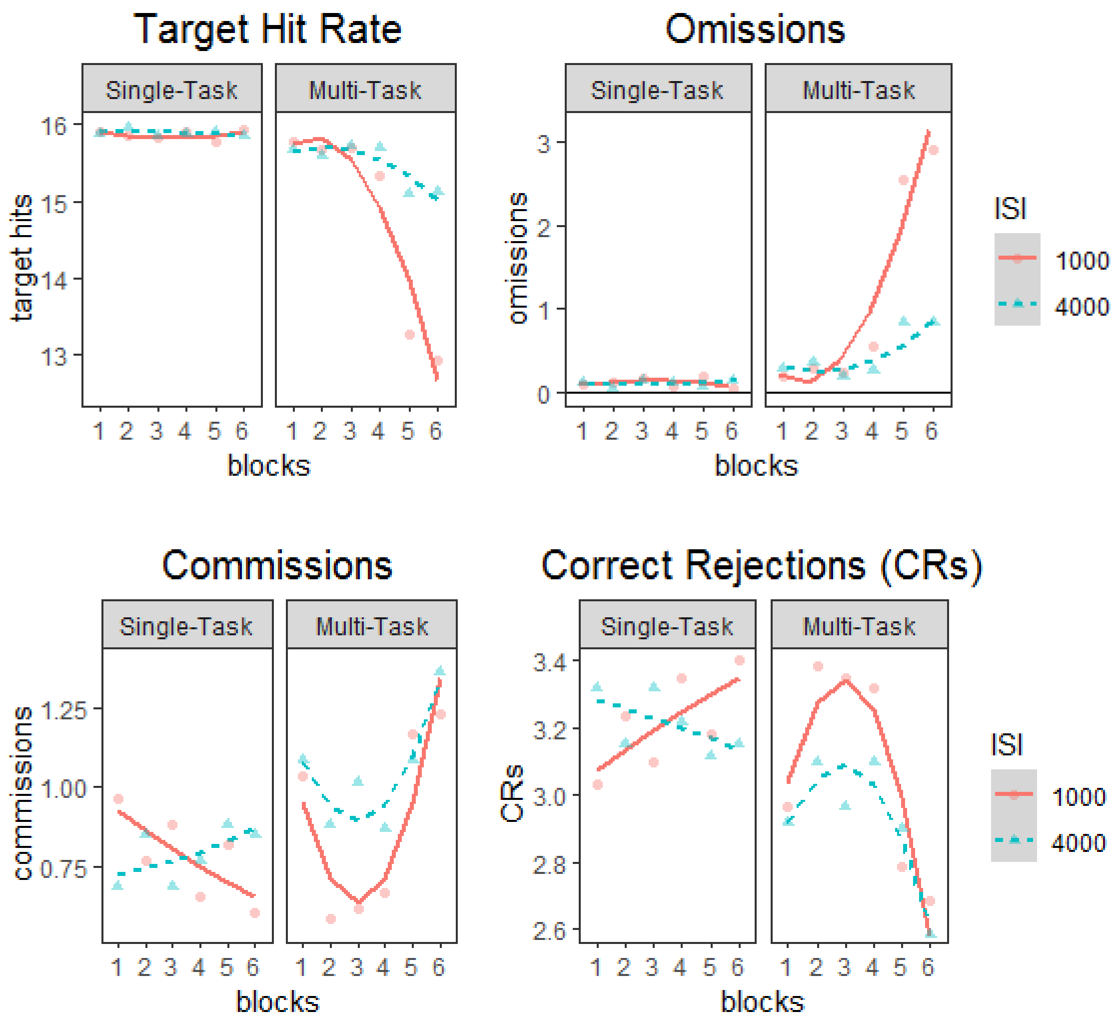

Figure 5.1: Growth curve analysis of objective CPT measures averaged per Block per ISI in E3 – (a) Target Hit Rate, (b) Omissions, (c) Commissions, (d) CRs, (e) Sensitivity, and (f) Response Bias.

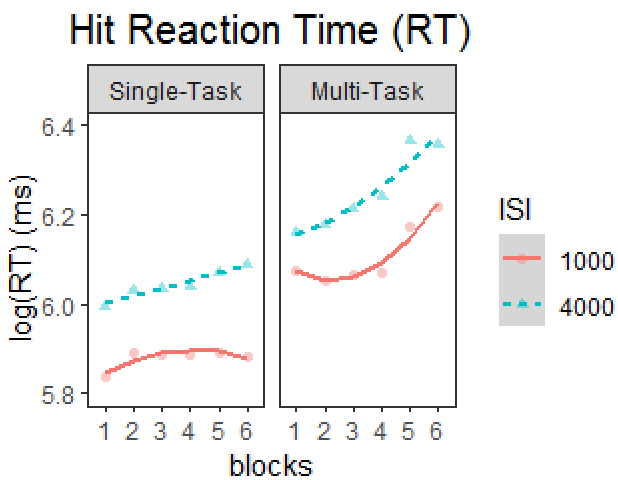

Figure 5.2: Growth curve analysis of Reaction Time (RT) averaged per Block per ISI in E3.

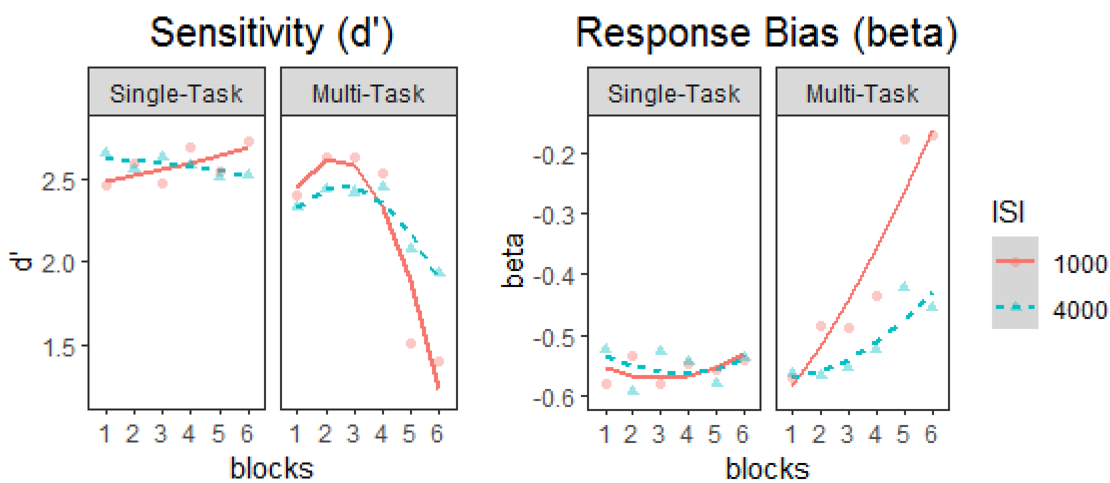

Figure 5.3: Growth curve analyses of SDT measures averaged per Block per ISI in E3 – (a) Sensitivity and (b) Response Bias.

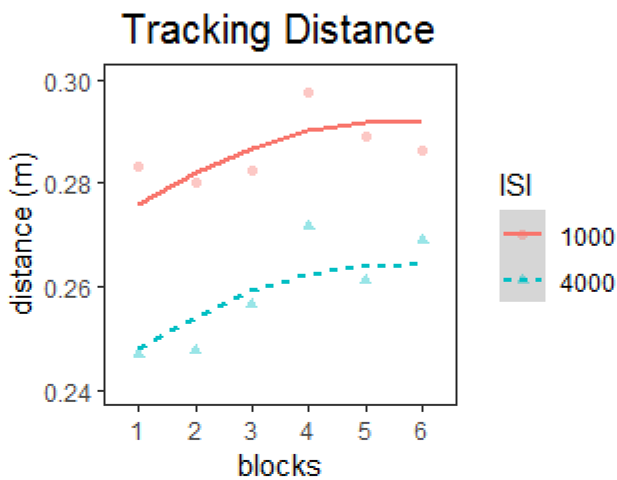

Figure 5.4: Growth curve analysis of Tracking Distance averaged per Block per ISI in E3.

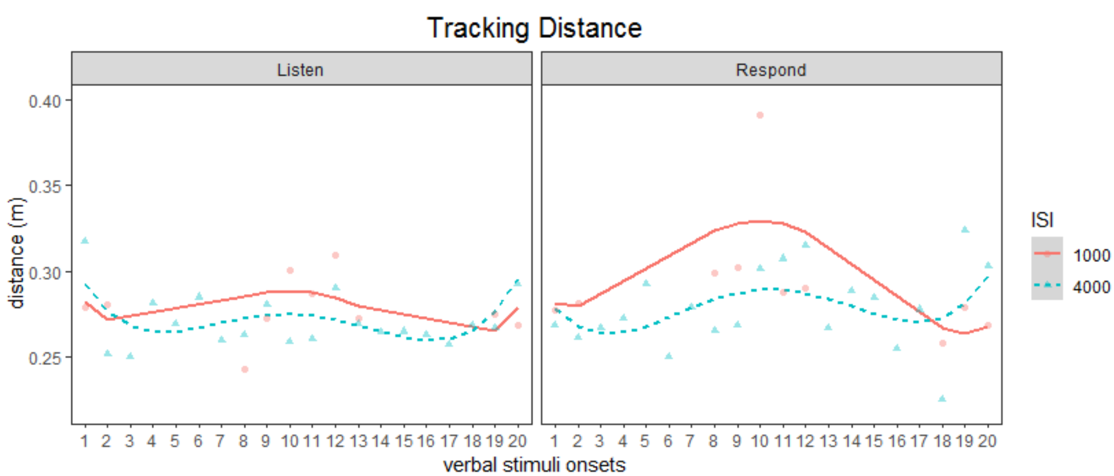

Figure 5.5: Growth curve analysis of Tracking Distance averaged during each verbal stimuli onset per ISI in E3.

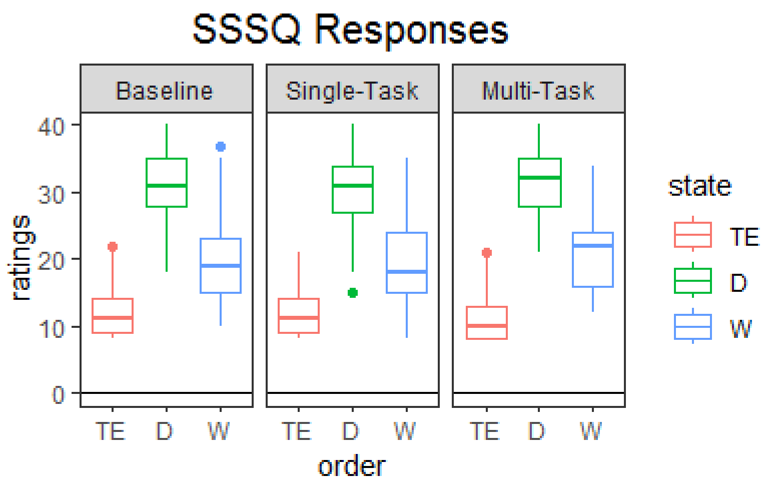

Figure 5.6: Boxplots for SSSQ questionnaire responses in E3 – Task Engagement (TE), Distress (D), and Worry (W). Note: The top and bottom lines of the box represent the upper and lower quartiles of the data, respectively; the darker colored line represents the median value of the data; and the vertical lines attached to the boxes represent the range of the scores, while the colored dot represents the outlier for the maximum and minimum scores.

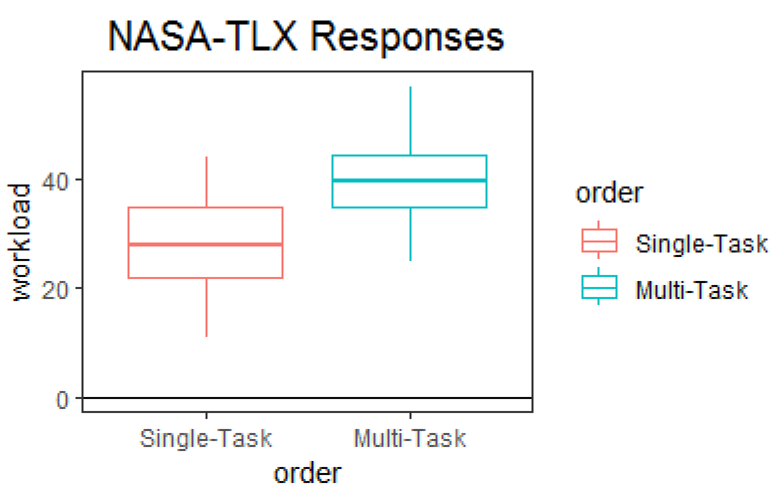

Figure 5.7: Boxplots for NASA-TLX questionnaire responses in E3. Note: The top and bottom lines of the box represent the upper and lower quartiles of the data, respectively; the darker colored line represents the median value of the data; and the vertical lines attached to the boxes represent the range of the scores.

CHAPTER 6

GENERAL DISCUSSION

In this dissertation, we conducted three experiments which implemented a novel paradigm we developed that measured performance during two sessions: a Single-Task session in which participants performed a go-no-go target detection task in the absence of any other task for approximately 12 minutes; and a Multi-Task session in which participants performed the detection task simultaneously with a driving-based tracking task for the same duration.

Our general predictions for this study were that performance would be worst in the more demanding Multi-Task session, and during faster rather than slower ISIs. While these are more so related to multitasking performance, and not to vigilance per se, we made these predictions to validate whether our paradigm effectively modulated task demand between sessions and ISI conditions, which therefore allowed us to examine the role of task demand in modulating vigilance performance (i.e., Luna et al., 2022). That said, we based these predictions on the resource depletion account of vigilance performance, which attributes vigilance decrements to the depletion of available attentional resources during demanding multitasking scenarios (Caggiano & Parasuraman, 2004; Helton & Russell, 2011; Helton & Warm, 2008; Warm et al., 2008; Wickens, 2008).

Across three experiments, we found that performance was indeed worst overall for most measures, and varied depending on ISI within- and between-sessions.

Specifically, in E1, target hits and tracking performance was worst during the 1000ms ISI, RT was worst in the 4000ms ISI, and there was no difference between the 1000ms and 4000ms ISIs for omissions, commissions, CRs, and sensitivity. In E2, similar general patterns in performance occurred, except that omissions were also worst during the 1000ms ISI; and in E3, performance was notably worst in the 1000ms ISI for most measures, and worst in the 4000ms ISI for commissions, CRs, and RTs. These results, as well as the questionnaire data taken before and after each session across experiments, partially supported both predictions. Importantly, these results also demonstrate that operator arousal and alertness levels (manipulated via alternating ISIs) may significantly affect their ability to sustain attention on relevant tasks (Silverstein et al., 2004), and that these changes may be mediated by certain task-specific factors (Wickens, 2002), such as the levels of control and automaticity required to perform both tasks (Fisk & Schneider, 1981).

We examined this more directly with our specific predictions for each experiment. Our prediction for E1 was that GCAs would reveal non-linear performance trajectories across all measures in the Multi-Task session. We based this on the findings from studies that suggest that performance can rapidly change as operators learn how to perform their tasks most optimally (Basner et al., 2018; Brown, 2008; Fisk & Schneider, 1981; Parasuraman & Giambra, 1991), especially for more difficult tasks that require higher levels of executive control, and that these changes can also affect the occurrence of vigilance decrements during the early stages of task performance (Fisk & Scneider, 1981; Gartenberg et al., 2018; Smith et al., 2023; Thomson et al., 2015). Results partially supported this prediction in that: linear trajectories in which performance progressively improved across the session characterized target hits and omissions, and trajectories in which performance progressively worsened

across the session characterized RTs and average tracking distance (during the 1000ms ISI); while quadratic trajectories in which performance initially improved and then worsened towards the end of the session characterized commissions, CRs, sensitivity, and average distance (during the 2000ms and 4000ms ISIs). These results demonstrate that operators are indeed susceptible to learning and practice effects during the early stages of task performance, and that these changes may be influenced by the level of control and automaticity involved with performing their tasks (Findley et al., 1999; Fisk & Schneider, 1981; Lara et al., 2014). Most notably, the quadratic performance trajectories revealed by the GCAs may be associated with tasks involving higher levels of executive control, while the linear trajectories may be associated with tasks involving higher levels of automaticity.

Along these same lines, our predictions for E2 were that GCAs would reveal quadratic performance trajectories for commissions and CR measures, since these are associated with the primary inhibition CPT task and may thus require higher level of executive control for its execution; and linear performance trajectories for target hits, RTs, and average distance measures, since these are associated with the secondary responding CPT and tracking tasks, and may thus involve higher levels of automization for their execution. Results fully supported these predictions with the exception of target hits, which was characterized by quadratic trajectories. These results confirm the E1 predictions that both task demand and learning processes may influence performance across measures, and demonstrate that operators may invoke workload mitigation strategies in which they allocate attentional resources towards performing the more demanding tasks during sustained multitasking scenarios (Epling et al., 2016; Finomore et al., 2009; Kurzban et al., 2013; Wickens, 2008).

Finally, based on the results of E2, our predictions for E3 were that GCAs would reveal quadratic performance trajectories for the primary inhibition (i.e., CRs and commissions) and secondary responding (i.e., target hits and RT) CPT measures, and linear performance trajectories for the average distance measure; we also predicted that average distance would be worst overall when participants responded compared to when they listened to verbal prompts. We based this on the findings from Rann & Almor (2022), Lee et al. (2017), and others which suggest that producing speech is more demanding than listening to speech. Results fully supported these predictions. These results confirm the E2 prediction that operators may selectively attend to higher priority tasks during demanding sustained multitasking scenarios (Kurzban et al., 2013), and that the level of control involved with performing the tasks can influence which tasks operators prioritize.

6.1 THEORETICAL IMPLICATIONS

These results provide useful insights regarding the vigilance theories discussed in this dissertation. For example, as mentioned earlier in the introduction, the two most prominent theories of vigilance decrements are the resource depletion account (Helton & Warm, 2008), which suggests that vigilance tasks are stressful and demanding, and that operators' attentional resources are progressively depleted over the course of task performance; and the cognitive underload account (Scerbo et al., 1992), which suggests that vigilance tasks are monotonous and boring, and that operators mindlessly disengage from performing tasks over the course of task performance. Of the two theories, we propose that the resource depletion account is best supported by the results from our study. Specifically, the performance trajectories that characterized the CPT data in the Multi-Task session, as well as the subjective measures reported in the questionnaires, suggests that simultaneously performing

the sustained attention tasks is highly demanding in comparison to performing the CPT by itself in the Single-Task session.

That said, cognitive underload could be a factor if either task is performed by itself, as the CPT is in the Single-Task session. However, the CPT requires higher levels of control for the primary inhibition task, so perhaps cognitive underload would best apply to the highly automatized (and potentially monotonous) tracking task if it was performed by itself in the Single-Task session instead of the CPT (we discuss these points in more detail in the next section). Relatedly, the higher demands reported in the Multi-Task session can also provide support for the mind-wandering account (Smallwood & Schooler, 2006), which suggests that during monotonous scenarios operators progressively shift their attention away from the primary task and towards demanding task-unrelated thoughts; as well as the resource-control account (Thomson et al., 2015), which suggests that the operators' ability to sustain task goals and prevent mind-wandering from consuming available resources progressively diminishes over the course of task performance.

We can only speculate on the relevancy of these theories since we did not assess for mind-wandering in this study (discussed in more detail in the next section). However, the results across studies, and especially in E3, provide the most convincing support for the opportunity-cost account (Kurzban et al., 2013), which suggests that operators strategically allocate their cognitive effort based on the reward and/or motivation involved with performing sustained attention tasks. This is supported by the differing performance trajectories that characterized the measures of the CPT and tracking tasks in each experiment, and especially during the verbal task blocks in E3. Specifically, steady performance declines occurred in E3 during both ISIs for the tracking task, while sharp performance declines

occurred during the 1000ms ISI in the verbal task blocks for the CPT. This clearly demonstrates that the verbal task interfered only with the CPT (which is also predicted by the resource depletion account), and that participants strategically allocated resources towards performing the tracking and verbal tasks instead of the demanding CPT.

These results also provide useful insights regarding the validity of the measures assessed in our paradigm. First, differentiating measures according to what aspect of the CPT they are most associated helps to better assess the attentional mechanisms underlying CPT performance (Ord et al., 2021; Riccio et al., 2002; Roebuck et al., 2016). We suggest that the CPT used in this study consisted of three sub-tasks: the primary inhibition task, in which CRs were associated with target detection accuracy and inhibition ability, and commissions were associated with impulsivity and/or the inability to inhibit pre-potent motor responding; the secondary responding task, in which target hits were associated with target detection accuracy and RT was associated with the processing speed involved with pre-potent motor response; and an arousal manipulation, in which performance across all measures were potentially affected when target presentation rates changed throughout the sessions (Conners et al., 2003; Huang-Pollock et al., 2012; MacLean et al., 2009). Of note, the arousal manipulation only applies to CPTs which alternate difficulty within-task, like the CCPT (Conners, 2014), and not traditional SART-like paradigms. Furthermore, given the sub-tasks that we outlined for SART-like paradigms, we suggest that the vigilance components dissociated by Luna et al. (2018) should be re-examined to best account for the many different aspects involved with performing vigilance tasks.

Second, while analysis of the SDT sensitivity measure (i.e., *d*’) may not be the most reliable means of assessing sustained attention performance on SART-like tasks (Huang-

Polluck et al., 2020; Sarter et al., 2001; Thomson et al., 2016), the response bias measure (i.e., β) appears to be a useful indicator of performance, particularly since it is highly associated with other relevant measures of performance (e.g., RTs, hits, and commissions), and can indicate how and when operators employ response strategies during demanding sustained multitasking scenarios (Bedi et al., 2023; Helton et al., 2009; Rubinstein, 2020). For example, response bias was overall liberal and showed little change across sessions for E1 and E2. However, response bias for E3 sharply shifted towards a more conservative criterion during the 1000ms ISI when participants performed the verbal task simultaneously with the other tasks. Interestingly, this change did not affect the average distance tracking measure. Therefore, these results further support our findings suggesting that operators may engage in mitigation strategies in which they weigh the opportunity and costs associated with performing certain tasks (Gutzwiller & Sitzman, 2017; Kurzban et al., 2013), and that that this choice is highly influenced by task relevance as well as the executive control required to perform the task (Finomore et al., 2009; Henderson, 2017; Henderson et al., 2009; Thomson et al., 2016).

Lastly, combining tracking tasks with CPTs may indeed provide a more ecological method for examining multitasking performance since it mimics several different real-world tasks. For example, the combination of SART-like CPTs and tracking tasks may be comparable to military occupations which require gunners to rapidly fire at enemy targets while avoiding friendly fire incidents (Munnik et al., 2020; Wilson, 2015). Therefore, the results from these combined paradigms may help with the development of new strategies to lower the number of accidents that occur during critical military missions (Abich IV et al.,

2017; Chérif et al., 2018; Nicolae et al., 2016; Shappell & Wiegmann, 2004; Thomas & Russo, 2007).

6.2 LIMITATIONS AND FUTURE RESEARCH

This study has several limitations that should be addressed before considering the future directions of this work. First, unlike other studies that implemented similar paradigms (e.g., Buckley et al., 2013), our paradigm utilized fixed-order sessions, and did not include a tracking-only session. We did this because we wanted to give participants more time to practice the CPT (compared to the tracking task) so as to elucidate the changes in performance that occur during the more demanding Multi-Task session. While the results from our study support both assumptions, not including a tracking-only session eliminated our ability to compare sessions within our paradigm, and to generalize results to other tracking-related studies in the literature. Although using these fixed order sessions was necessary for the purposes of this study, it may be useful for future studies to compare vigilance decrements in single and multiple task contexts by manipulating the order of single and multiple task blocks.

Second, although we administered subjective state questionnaires before and after sessions throughout our study (Hart & Staveland, 1988; Helton, 2004), these questionnaires were not as informative as originally anticipated. In line with other related vigilance studies (Körber et al. 2015; Mooneyham & Schooler, 2013; Seli et al., 2016; Smallwood et al., 2004; Thomson et al., 2015; Weinstein, 2018; Yanko & Spalek, 2014), it may be worthwhile to also assess whether and to what extent operators engage in mind-wandering activities during demanding multitasking scenarios. This could provide useful insight for assessing more recent theories for vigilance decrements not directly examined in this study, such as the resource-control account (Thomson et al., 2015). That said, it may also be interesting to

collect continuous subjective measures of workload and stress, such as heart rates, eye movement activity, and EEG (Hancock, 1989; Mehrabi & Kim, 2022), to further examine the role of cognitive load on sustained multitasking performance.

Finally, while the duration of sustained attention tasks can range from several minutes to hours (Arrabito et al., 2007), our paradigm specifically measured performance within a small time window (approximately 12-minutes). This was not a limitation *per se*, since our interest was on the performance changes that occur during relatively short periods (Conners et al., 2003; Loh et al., 2004; Roach et al., 2006); that said, extending the duration of our paradigm would be a useful contribution to the literature since this relates more to the scenarios operators typically encounter in the real-world (Canisius & Penzel, 2007; Mackie, 1987; Mackworth, 1968; Popp et al., 2015).

6.3 CONCLUSION

In this dissertation, we examined the time-course of vigilance decrements that occur when operators perform demanding multitasking activities. Our findings suggest that cognitive overload and opportunity costs play significant roles in vigilance performance, and that dissociating the underlying mechanisms involved in task performance can help better interpret measures from common CPTs used in sustained attention research. In addition to contributing to the theoretical understanding of attention and multitasking performance, the insights gained from our study can inform the development of strategies and interventions to mitigate vigilance decrements and improve performance in demanding work environments.

REFERENCES

Abich IV, J., Reinerman-Jones, L., & Matthews, G. (2017). Impact of three task demand factors on simulated unmanned system intelligence, surveillance, and reconnaissance operations. *Ergonomics*, *60*(6), 791-809.

Al-Shargie, F., Tariq, U., Mir, H., Alawar, H., Babiloni, F., & Al-Nashash, H. (2019). Vigilance decrement and enhancement techniques: a review. *Brain sciences*, *9*(8), 178.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. Quarterly journal of experimental psychology, 24(2), 225-235.

Almor, A. (2008). Why does language interfere with vision-based tasks?. *Experimental Psychology*, *55*(4), 260.

Ariga, A., & Lleras, A. (2011). Brief and rare mental "breaks" keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition*, *118*(3), 439-443.

Arrabito, G. R., Abel, S. M., & Lam, K. (2007). Methods for mitigating the vigilance decrement in an auditory sonar monitoring task: A research synthesis. *Canadian acoustics*, *35*(4), 15-23.

Atchley, P., & Chan, M. (2011). Potential benefits and costs of concurrent task engagement to maintain vigilance: A driving simulator investigation. *Human factors*, *53*(1), 3-12.

Atchley, P., Chan, M., & Gregersen, S. (2014). A strategically timed verbal task improves performance and neurophysiological alertness during fatiguing drives. Human factors, 56(3), 453-462.

Atchley, P., Dressel, J., Jones, T. C., Burson, R. A., & Marshall, D. (2011). Talking and driving: applications of crossmodal action reveal a special role for spatial language. *Psychological research*, *75*, 525-534.

Azizi, E., Stainer, M. J., & Abel, L. A. (2018). Is experience in multi-genre video game playing accompanied by impulsivity?. *Acta psychologica*, *190*, 78-84.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of memory and language, 59(4), 390-412.

Baca, R. H. (2017). *Operational Research in the Royal Air Force During World War II and How It Can Be Applied to Big Data in Future War*. MARINE CORPS UNIV QUANTICO VA.

Baldwin, C. L., & Lewis, B. A. (2017). Positive valence music restores executive control over sustained attention. *PLoS One*, *12*(11), e0186231.

Ballard, J. C. (1996). Computerized assessment of sustained attention: a review of factors affecting vigilance performance. *Journal of clinical and experimental neuropsychology*, *18*(6), 843-863.

Ballard, J. C. (2001). Assessing attention: Comparison of response-inhibition and traditional continuous performance tests. *Journal of Clinical and Experimental Neuropsychology*, *23*(3), 331-350.

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: behavioral and neural basis of response control. *Progress in neurobiology*, *108*, 44-79.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. Frontiers in psychology, 4, 328.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of memory and language, 68(3), 255-278.

Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M., & Rushby, J. A. (2005). Arousal and activation in a continuous performance task. *Journal of psychophysiology*, *19*(2), 91-99.

Basacik, D., Waters, S., & Reed, N. (2015). Detecting cognitive underload in train driving: a physiological approach. In *Proceedings of the 5th International Rail Human Factors Conference* (pp. 14-17).

Basner, M., & Dinges, D. F. (2011). Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss. *Sleep, 34*(5), 581-591.

Basner, M., Hermosillo, E., Nasrini, J., McGuire, S., Saxena, S., Moore, T. M., ... & Dinges, D. F. (2018). Repeated administration effects on psychomotor vigilance test performance. *Sleep*, *41*(1), zsx187.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

Bearden, T. S., Cassisi, J. E., & White, J. N. (2004). Electrophysiological correlates of vigilance during a continuous performance test in healthy adults. *Applied Psychophysiology & Biofeedback*, *29*(3).

Becker, A. B., Warm, J. S., Dember, W. N., & Hancock, P. A. (1991, September). Effects of feedback on perceived workload in vigilance performance. In Proceedings of the Human Factors Society Annual Meeting (Vol. 35, No. 20, pp. 1491-1494). Sage CA: Los Angeles, CA: SAGE Publications.

Bedi, A., Russell, P. N., & Helton, W. S. (2023). Go-stimuli probability influences response bias in the sustained attention to response task: A signal detection theory perspective. *Psychological Research*, *87*(2), 509-518.

Beede, K. E., & Kass, S. J. (2006). Engrossed in conversation: The impact of cell phones on simulated driving performance. *Accident Analysis & Prevention*, *38*(2), 415-421.

Berardi, Raja Parasuraman, James V. Haxby, A. (2001). Overall vigilance and sustained attention decrements in healthy aging. *Experimental aging research*, *27*(1), 19-39.

Berger, I., & Cassuto, H. (2014). The effect of environmental distractors incorporation into a CPT on sustained attention and ADHD diagnosis among adolescents. *Journal of neuroscience methods*, *222*, 62-68.

Bherer, L., Kramer, A. F., Peterson, M. S., Colcombe, S., Erickson, K., & Becic, E. (2005). Training effects on dual-task performance: are there age-related differences in plasticity of attentional control?. *Psychology and aging*, *20*(4), 695.

Black, S. C., Bender, A. D., Whitney, S. J., Loft, S., & Visser, T. A. (2022). The effect of multi-tasking training on performance, situation awareness, and workload in simulated air traffic control. *Applied cognitive psychology*, *36*(4), 874-890.

Boiteau, T. W., Malone, P. S., Peters, S. A., & Almor, A. (2014). Interference between conversation and a concurrent visuomotor task. *Journal of experimental psychology: General*, *143*(1), 295.

Borgaro, S., Pogge, D. L., DeLuca, V. A., Bilginer, L., Stokes, J., & Harvey, P. D. (2003). Convergence of different versions of the continuous performance test: Clinical and scientific implications. *Journal of clinical and experimental neuropsychology*, *25*(2), 283-292.

Bollen, K. A. (2007). On the origins of latent curve models. In Factor analysis at 100 (pp. 93-112). Routledge.

Böttcher, A., Adelhöfer, N., Wilken, S., Raab, M., Hoffmann, S., & Beste, C. (2023). TRACK—a new algorithm and open-source tool for the analysis of pursuit-tracking sensorimotor integration processes. *Behavior research methods*, 1-14.

Bowden, V. K., Loft, S., Wilson, M. K., Howard, J., & Visser, T. A. (2019). The long road home from distraction: Investigating the time-course of distraction recovery in driving. *Accident Analysis & Prevention*, *124*, 23-32.

Broadbent, D. E., & Gregory, M. (1965). Effects of noise and of signal rate upon vigilance analysed by means of decision theory. *Human Factors*, *7*(2), 155-162.

Brown, S. W. (2008). The attenuation effect in timing: Counteracting dual-task interference with time-judgment skill training. *Perception*, *37*(5), 712-724.

Bruyas, M. P., & Dumont, L. (2013, June). Sensitivity of Detection Response Task (DRT) to the driving demand and task difficulty. In *Driving Assesment Conference* (Vol. 7, No. 2013). University of Iowa.

Bubnik, M. G., Hawk, L. W., Pelham, W. E., Waxmonsky, J. G., & Rosch, K. S. (2015). Reinforcement enhances vigilance among children with ADHD: comparisons to typically developing children and to the effects of methylphenidate. *Journal of Abnormal Child Psychology*, *43*, 149-161.

Buckley, R. J. (2013). Sustained attention lapses and behavioural microsleeps during tracking, psychomotor vigilance, and dual tasks.

Buckley, R. J., Helton, W. S., Innes, C. R., Dalrymple-Alford, J. C., & Jones, R. D. (2016). Attention lapses and behavioural microsleeps during tracking, psychomotor vigilance, and dual tasks. Consciousness and cognition, 45, 174-183.

Busk, J., & Galbraith, G. C. (1975). EEG correlates of visual-motor practice in man. *Electroencephalography and clinical neurophysiology*, *38*(4), 415-422.

Byrne, B. M., & Crombie, G. (2003). Modeling and testing change: An introduction to the latent growth curve model. Understanding Statistics, 2(3), 177-203.

Cabon, P. H., Coblentz, A., Mollard, R., & Fouillot, J. P. (1993). Human vigilance in railway and long-haul flight operation. *Ergonomics*, *36*(9), 1019-1033.

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic bulletin & review*, *11*(5), 932-937.

Callejas, A., Lupiánez, J., & Tudela, P. (2004). The three attentional networks: On their independence and interactions. *Brain and cognition*, *54*(3), 225-227.

Canisius, S., & Penzel, T. (2007). Vigilance monitoring–review and practical aspects.

Cardoso, M., Fulton, F., Callaghan, J. P., Johnson, M., & Albert, W. J. (2019). A pre/post evaluation of fatigue, stress and vigilance amongst commercially licensed truck

drivers performing a prolonged driving task. *International journal of occupational safety and ergonomics*, *25*(3), 344-354.

Carter, L., Russell, P. N., & Helton, W. S. (2013). Target predictability, sustained attention, and response inhibition. Brain and Cognition, 82(1), 35-42.

Castro, C., Padilla, J. L., Doncel, P., Garcia-Fernandez, P., Ventsislavova, P., Eisman, E., & Crundall, D. (2019). How are distractibility and hazard prediction in driving related? Role of driving experience as moderating factor. *Applied ergonomics*, *81*, 102886.

Chen, J. Y., & Joyner, C. T. (2009). Concurrent performance of gunner's and robotics operator's tasks in a multitasking environment. *Military Psychology*, *21*(1), 98-113.

Chérif, L., Wood, V., Marois, A., Labonté, K., & Vachon, F. (2018). Multitasking in the military: Cognitive consequences and potential solutions. *Applied cognitive psychology*, *32*(4), 429-439.

Chiew, K. S., & Braver, T. S. (2013). Temporal dynamics of motivation-cognitive control interactions revealed by high-resolution pupillometry. Frontiers in psychology, 4, 15.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual review of psychology*, *62*, 73-101.

Claveria, J. B., Hernandez, S., Anderson, J. C., & Jessup, E. L. (2019). Understanding truck driver behavior with respect to cell phone use and vehicle operation. *Transportation research part F: traffic psychology and behaviour*, *65*, 389-401.

Cohen, J., LaRue, C., & Cohen, H. H. (2017). Attention interrupted: Cognitive distraction & workplace safety. *Professional Safety*, *62*(11), 28-34.

Colquhoun, W. P. (1967). Sonar target detection as a decision process. *Journal of Applied Psychology*, *51*(2), 187.

Conners, C. K. (2000). *Conners' continuous performance test*. Multi-Health Systems.

Conners, C. K. (2014). Conners continuous performance test 3rd edition (Conners CPT 3) & connors continuous auditory test of attention (Conners CATA): Technical manual. MHS.

Conners, C. K., Epstein, J. N., Angold, A., & Klaric, J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of abnormal child psychology*, *31*, 555-562.

Cooper, J. M., Medeiros-Ward, N., & Strayer, D. L. (2013). The impact of eye movements and cognitive workload on lateral position variability in driving. *Human factors*, *55*(5), 1001-1014.

Cooper, J. M., Medeiros-Ward, N., Seegmiller, J., & Strayer, D. L. (2009, October). Shifting eyes and thinking hard keep us in our lanes. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 53, No. 23, pp. 1753-1756). Sage CA: Los Angeles, CA: SAGE Publications.

Cornblatt, B. A., Lenzenweger, M. F., & Erlenmeyer-Kimling, L. (1989). The continuous performance test, identical pairs version: II. Contrasting attentional profiles in schizophrenic and depressed patients. Psychiatry research, 29(1), 65-85.

Craig, A. (1978). Is the vigilance decrement simply a response adjustment towards probability matching?. *Human Factors*, *20*(4), 441-446.

Cummings, M. L., Gao, F., & Thornburg, K. M. (2016). Boredom in the workplace: A new look at an old problem. *Human factors*, *58*(2), 279-300.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. Journal of cognition and development, 11(2), 121-136.

Dalley, J. W., Everitt, B. J., & Robbins, T. W. (2011). Impulsivity, compulsivity, and top-down cognitive control. *Neuron*, *69*(4), 680-694.

Dang, J. S., Figueroa, I. J., & Helton, W. S. (2018). You are measuring the decision to be fast, not inattention: the Sustained Attention to Response Task does not measure sustained attention. *Experimental brain research*, *236*, 2255-2262.

Dember, W. N., Galinsky, T. L., & Warm, J. S. (1992). The role of choice in vigilance performance. *Bulletin of the Psychonomic Society*, *30*, 201-204.

Denney, C. B., Rapport, M. D., & Chung, K. M. (2005). Interactions of task and subject variables among continuous performance tests. *Journal of Child Psychology and Psychiatry*, *46*(4), 420-435.

Desmond, P. A., Hancock, P. A., & Monette, J. L. (1998). Fatigue and automation-induced impairments in simulated driving performance. *Transportation Research Record*, *1628*(1), 8-14.

Devlin, S. P., Brown, N. L., Drollinger, S., Alami, J., & Riggs, S. L. (2023). Workload transition rate matters: Evidence from growth curve modeling. *Applied Ergonomics*, *106*, 103885.

Diallo, T. M., Morin, A. J., & Parker, P. D. (2014). Statistical power of latent growth curve models to detect quadratic growth. Behavior research methods, 46(2), 357-371.

Diamond, A. (2005). Attention-deficit disorder (attention-deficit/hyperactivity disorder without hyperactivity): A neurobiologically and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Development and psychopathology*, *17*(3), 807-825.

Dillard, M. B., Warm, J. S., Funke, G. J., Nelson, W. T., Finomore, V. S., McClernon, C. K., ... & Funke, M. E. (2019). Vigilance tasks: Unpleasant, mentally demanding, and stressful even when time flies. *Human factors*, *61*(2), 225-242.

Dismukes, R. K., & Nowinski, J. (2007). Prospective memory, concurrent task management, and pilot error. Attention: From theory to practice, 225, 236.

Ditchburn, R. W. (1943). Some factors affecting efficiency of work of lookouts. *Admiralty Res. Lab. Rep., Great Britain*.

Donders, F. C. (1868). Die schnelligkeit psychischer processe: Erster artikel. Archiv für Anatomie, Physiologie und wissenschaftliche Medicin, 657-681.

Dorrian, J., Roach, G. D., Fletcher, A., & Dawson, D. (2007). Simulated train driving: fatigue, self-awareness and cognitive disengagement. Applied ergonomics, 38(2), 155-166.

Drew, G. C. (1951). Variations in reflex blink-rate during visual-motor tasks. *Quarterly Journal of Experimental Psychology*, *3*(2), 73-88.

Edwards, M. C., Gardner, E. S., Chelonis, J. J., Schulz, E. G., Flake, R. A., & Diaz, P. F. (2007). Estimates of the validity and utility of the Conners' Continuous

Performance Test in the assessment of inattentive and/or hyperactive-impulsive behaviors in children. *Journal of abnormal child psychology*, *35*, 393-404.

Egeland, J., & Kovalik-Gran, I. (2010). Validity of the factor structure of Conners' CPT. *Journal of attention disorders*, *13*(4), 347-357.

Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation research part F: traffic psychology and behaviour*, *8*(2), 97-120.

Epling, S. L., Russell, P. N., & Helton, W. S. (2016). A new semantic vigilance task: vigilance decrement, workload, and sensitivity to dual-task costs. *Experimental brain research*, *234*, 133-139.

Epstein, J. N., & Loren, R. E. (2013). Changes in the definition of ADHD in DSM-5: subtle but important. *Neuropsychiatry*, *3*(5), 455.

Epstein, J. N., Erkanli, A., Conners, C. K., Klaric, J., Costello, J. E., & Angold, A. (2003). Relations between continuous performance test performance measures and ADHD behaviors. *Journal of abnormal child psychology*, *31*, 543-554.

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current opinion in psychology*, *29*, 174-180.

Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral cortex*, *23*(11), 2712-2723.

Esterman, M., Reagan, A., Liu, G., Turner, C., & DeGutis, J. (2014). Reward reveals dissociable aspects of sustained attention. *Journal of Experimental Psychology: General*, *143*(6), 2287.

Eysenck, M. W. (1976). Arousal, learning, and memory. *Psychological bulletin*, *83*(3), 389.

Fairclough, S. H., & Mulder, L. J. M. (2012). Psychophysiological processes of mental effort investment.

Falzon, P. A. (2009). Discourse segmentation and the management of multiple tasks in single episodes of air traffic controller-pilot spoken radio communication. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, (4).

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of cognitive neuroscience*, *14*(3), 340-347.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods, 39(2), 175-191.

Fernandez-Duque, D., & Posner, M. I. (2001). Brain imaging of attentional networks in normal and pathological states. *Journal of Clinical and Experimental Neuropsychology*, *23*(1), 74-93.

Fiecconi, E. C. (2021). Aristotle on attention. *Archiv für Geschichte der Philosophie, 103*(4), 602-633.

Findley, L. J., Suratt, P. M., & Dinges, D. F. (1999). Time-on-task decrements in “steer clear” performance of patients with sleep apnea and narcolepsy. *Sleep*, *22*(6), 804-809.

Finomore, V., Matthews, G., Shaw, T., & Warm, J. (2009). Predicting vigilance: A fresh look at an old problem. *Ergonomics*, *52*(7), 791-808.

Finomore, V. S., Shaw, T. H., Warm, J. S., Matthews, G., Weldon, D., & Boles, D. B. (2009, October). On the Workload of Vigilance: Comparison of the NASA-TLX and the MRQ. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 53, No. 17, pp. 1057-1061). Sage CA: Los Angeles, CA: Sage Publications.

Fisher, A. V. (2019). Selective sustained attention: A developmental foundation for cognition. *Current opinion in psychology*, *29*, 248-253.

Fisk, A. D., & Scerbo, M. W. (1987). Automatic and control processing approach to interpreting vigilance performance: A review and reevaluation. Human Factors, 29(6), 653-660.

Fisk, A. D., & Schneider, W. (1981). Control and automatic processing during tasks requiring sustained attention: A new approach to vigilance. *Human factors*, *23*(6), 737-750.

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, *111*(3), 345-355.

Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences*, *1396*(1), 70-91.

Fox, J. R., Park, B., & Lang, A. (2007). When available resources become negative resources: The effects of cognitive overload on memory sensitivity and criterion bias. *Communication Research, 34*(3), 277-296.

Fuermaier, A. B., Tucha, L., Guo, N., Mette, C., Müller, B. W., Scherbaum, N., & Tucha, O. (2022). It takes time: Vigilance and sustained attention assessment in adults with ADHD. *International Journal of Environmental Research and Public Health*, *19*(9), 5216.

Gamboz, N., Zamarian, S., & Cavallero, C. (2010). Age-related differences in the attention network test (ANT). *Experimental aging research*, *36*(3), 287-305.

Gartenberg, D., Gunzelmann, G., Hassanzadeh-Behbaha, S., & Trafton, J. G. (2018). Examining the role of task requirements in the magnitude of the vigilance decrement. *Frontiers in psychology*, *9*, 1504.

Gastaldi, M., Rossi, R., & Gecchele, G. (2014). Effects of driver task-related fatigue on driving performance. Procedia-Social and Behavioral Sciences, 111, 955-964.

Ghylin, K. M., Drury, C. G., Batta, R., & Lin, L. (2007, October). Temporal effects in a security inspection task: Breakdown of performance components. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 51, No. 2, pp. 93-97). Sage CA: Los Angeles, CA: SAGE Publications.

Gilden, D. L., & Hancock, H. (2007). Response variability in attention-deficit disorders. *Psychological Science*, *18*(9), 796-802.

Grech, M. R., Neal, A., Yeo, G., Humphreys, M., & Smith, S. (2009). An examination of the relationship between workload and fatigue within and across consecutive days of work: Is the relationship static or dynamic?. *Journal of occupational health psychology*, *14*(3), 231.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1, pp. 1969-12). New York: Wiley.

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2018). Driver vigilance in automated vehicles: Hazard detection failures are a matter of time. Human factors, 60(4), 465-476.

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human factors*, *45*(3), 349-359.

Gutzwiller, R. S., & Sitzman, D. M. (2017, September). Examining task priority effects in multi-task management. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 61, No. 1, pp. 762-766). Sage CA: Los Angeles, CA: SAGE Publications.

Gutzwiller, R. S., Fugate, S., Sawyer, B. D., & Hancock, P. A. (2015, September). The human factors of cyber network defense. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 59, No. 1, pp. 322-326). Sage CA: Los Angeles, CA: SAGE publications.

Hahn, M., Wild-Wall, N., & Falkenstein, M. (2011). Age-related differences in performance and stimulus processing in dual task situation. *Brain research*, *1414*, 66-76.

Hall, C. L., Valentine, A. Z., Groom, M. J., Walker, G. M., Sayal, K., Daley, D., & Hollis, C. (2016). The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: A systematic review. *European child & adolescent psychiatry*, *25*, 677-699.

Hancock, P. A. (1989). The effect of performance failure and task demand on the perception of mental workload. Applied Ergonomics, 20(3), 197-205.

Hancock, P. A. (2013). In search of vigilance: the problem of iatrogenically created psychological phenomena. American Psychologist, 68(2), 97.

Hancock, P. A., & Hart, S. G. (2002). Defeating terrorism: What can human factors/ergonomics offer?. *Ergonomics in design*, *10*(1), 6-16.

Hancock, P. A., & Matthews, G. (2019). Workload and performance: Associations, insensitivities, and dissociations. *Human factors*, *61*(3), 374-392.

Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In Proceedings of the human factors and ergonomics society annual meeting (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage publications.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In Advances in psychology (Vol. 52, pp. 139-183). North-Holland.

He, J., & McCarley, J. S. (2011). Effects of cognitive distraction on lateral lane keeping performance. Proceedings of Human Factors and Ergonomics Society, CA.He, J., Becic, E., Lee, Y. C., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human factors*, *53*(1), 13-21.

He, J., McCarley, J. S., & Kramer, A. F. (2014). Lane keeping under cognitive load: Performance changes and mechanisms. *Human factors*, *56*(2), 414-426.

Head, H. (1923). The conception of nervous and mental energy (II). *British Journal of Psychology*, *14*(2), 126.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychological review*, *62*(4), 243.

Heikoop, D. D., de Winter, J. C., van Arem, B., & Stanton, N. A. (2017). Effects of platooning on signal-detection performance, workload, and stress: A driving simulator study. *Applied ergonomics, 60*, 116-127.

Helton, W. S. (2004, September). Validation of a short stress state questionnaire. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 48, No. 11, pp. 1238-1242). Sage CA: Los Angeles, CA: Sage Publications.

Helton, W. S. (2009). Impulsive responding and the sustained attention to response task. *Journal of Clinical and Experimental Neuropsychology*, *31*(1), 39-47.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental brain research*, *212*(3), 429-437.

Helton, W. S., & Russell, P. N. (2017). Rest is still best: The role of the qualitative and quantitative load of interruptions on vigilance. Human Factors, 59(1), 91–100. https://doi.org/10.1177/0018720816683509

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. Acta psychologica, 129(1), 18-25.

Helton, W. S., Kern, R. P., & Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and cognition*, *18*(3), 600-607.

Henderson, J. M. (2017). Gaze control as prediction. *Trends in cognitive sciences*, *21*(1), 15-23.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic bulletin & review*, *16*(5), 850-856.

Hommel, B., Chapman, C. S., Cisek, P., Neyedli, H. F., Song, J. H., & Welsh, T. N. (2019). No one knows what attention is. *Attention, Perception, & Psychophysics*, *81*, 2288-2303.

Hoonakker, M., Doignon-Camus, N., & Bonnefond, A. (2017). Sustaining attention to simple visual tasks: a central deficit in schizophrenia? A systematic review. *Annals of the New York Academy of Sciences*, *1408*(1), 32-45.

Hope, A. T., Woolman, P. S., Gray, W. M., Asbury, A. J., & Millar, K. (1998). A system for psychomotor evaluation; design, implementation and practice effects in volunteers. Anaesthesia, 53(6), 545-550.

Huang-Pollock, C. L., Karalunas, S. L., Tam, H., & Moore, A. N. (2012). Evaluating vigilance deficits in ADHD: a meta-analysis of CPT performance. *Journal of abnormal psychology*, *121*(2), 360.

Innes, H. E. (1973). *Subjective and physiological indicators of fatigue in a vigilance task* (Doctoral dissertation, Monterey, California. Naval Postgraduate School).

Ishigami, Y., & Klein, R. M. (2009). Are individual differences in absentmindedness correlated with individual differences in attention?. *Journal of Individual Differences*, *30*(4), 220-237.

Ishigami, Y., & Klein, R. M. (2010). Repeated measurement of the components of attention using two versions of the Attention Network Test (ANT): Stability, isolability, robustness, and reliability. *Journal of neuroscience methods*, *190*(1), 117-128.

Ishigami, Y., Eskes, G. A., Tyndall, A. V., Longman, R. S., Drogos, L. L., & Poulin, M. J. (2016). The Attention Network Test-Interaction (ANT-I): reliability and validity in healthy older adults. *Experimental Brain Research*, *234*, 815-827.

James, W. (1890). The principles of psychology (vol. 1) New York: Henry Holt & Co. Inc. http://dx. doi. org/10.1037/11059-000.

Johnson, K. A., Robertson, I. H., Barry, E., Mulligan, A., Dáibhis, A., Daly, M., ... & Bellgrove, M. A. (2008). Impaired conflict resolution and alerting in children with ADHD: evidence from the Attention Network Task (ANT). *Journal of Child Psychology and Psychiatry*, *49*(12), 1339-1347.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual review of psychology*, *37*(1), 43-75.

Kahneman, D. (1973). Attention and effort (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

Karpinsky, N. D., Chancey, E. T., Palmer, D. B., & Yamani, Y. (2018). Automation trust and attention allocation in multitasking workspace. *Applied ergonomics, 70,* 194-201.

Kassambara, A. (2021). rstatix: pipe-friendly framework for basic statistical tests. R package version 0.7. 0. *Computer software]. https://CRAN. R-proje ct. org/packa ge= rstatix*.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human–Computer Interaction*, *12*(4), 391-438.

Kieras, D. E., Meyer, D. E., Ballas, J. A., & Lauber, E. J. (2000). Modern computational perspectives on executive mental processes and cognitive control: Where to from here. *Control of cognitive processes: Attention and performance XVIII*, 681-712.

Kinnear, N., Kelly, S. W., Stradling, S., & Thomson, J. (2013). Understanding how drivers learn to anticipate risk on the road: A laboratory experiment of affective anticipation of road hazards. *Accident Analysis & Prevention*, *50*, 1025-1033.

Kirlin, K. A. (2002). *Inattentive and impulsive profiles of the CPT-II and their relationship with DSM-IV ADHD subtypes*. University of Montana.

Klösch, G., Zeitlhofer, J., & Ipsiroglu, O. (2022). Revisiting the concept of vigilance. *Frontiers in psychiatry*, *13*.

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological bulletin*, *144*(6), 557.

Koelega, H. S. (1992). Extraversion and vigilance performance: 30 years of inconsistencies. Psychological bulletin, 112(2), 239.

Koelega, H. S., Brinkman, J. A., Hendriks, L., & Verbaten, M. N. (1989). Processing demands, effort, and individual differences in four different vigilance tasks. *Human Factors*, *31*(1), 45-62.

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing*, *3*, 2403-2409.

Krauzlis, R. J., Wang, L., Yu, G., & Katz, L. N. (2023). What is attention?. *Wiley Interdisciplinary Reviews: Cognitive Science*, *14*(1), e1570.

Kristjansson, S. D., Kircher, J. C., & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: An introduction to growth curve modeling. Psychophysiology, 44(5), 728-736.

Kubose, T. T., Bock, K., Dell, G. S., Garnsey, S. M., Kramer, A. F., & Mayhugh, J. (2006). The effects of speech production and speech comprehension on simulated driving performance. *Applied Cognitive Psychology: THe Official Journal of the Society for Applied Research in Memory and Cognition*, *20*(1), 43-63.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and brain sciences*, *36*(6), 661-679.

Lam, L. T. (2002). Distractions and the risk of car crash injury: The effect of drivers' age. *Journal of safety research*, *33*(3), 411-419.

Lamble, D., Kauranen, T., Laakso, M., & Summala, H. (1999). Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving. *Accident Analysis & Prevention*, *31*(6), 617-623.

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. *Psychological bulletin*, *139*(4), 870.

Lara, T., Madrid, J. A., & Correa, Á. (2014). The vigilance decrement in executive function is attenuated when individual chronotypes perform at their optimal time of day. *PloS one*, *9*(2), e88820.

Large, D. R., Burnett, G., Antrobus, V., & Skrypchuk, L. (2018). Driven to discussion: engaging drivers in conversation with a digital assistant as a countermeasure to passive task-related fatigue. *IET Intelligent Transport Systems*, *12*(6), 420-426.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of experimental psychology: General*, *133*(3), 339.

Lee, A., Cerisano, S., Humphreys, K. R., & Watter, S. (2017). Talking is harder than listening: The time course of dual-task costs during naturalistic conversation. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *71*(2), 111.

Lemay, S., Bédard, M. A., Rouleau, I., & Tremblay, P. L. (2004). Practice effect and test-retest reliability of attentional and executive tests in middle-aged to elderly subjects. The Clinical Neuropsychologist, 18(2), 284-302.

Levy, J., Pashler, H., & Boer, E. (2006). Central interference in driving: Is there any stopping the psychological refractory period?. *Psychological science*, *17*(3), 228-235.

Levy, F., Pipingas, A., Harris, E. V., Farrow, M., & Silberstein, R. B. (2018). Continuous performance task in ADHD: Is reaction time variability a key measure?. *Neuropsychiatric disease and treatment*, 781-786.

Lichstein, K. L., Riedel, B. W., & Richman, S. L. (2000). The mackworth clock test: A computerized version. *The Journal of psychology*, *134*(2), 153-161.

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences*, *1129*(1), 305-322.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 339-346.

Logan, G. D., & Crump, M. J. (2009). The left hand doesn't know what the right hand is doing: The disruptive effects of attention to the hands in skilled typewriting. *Psychological Science*, *20*(10), 1296-1300.

Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in human neuroscience*, *13*, 57.

López-Ramón, M. F., Castro, C., Roca, J., Ledesma, R., & Lupiañez, J. (2011). Attentional networks functioning, age, and attentional lapses while driving. *Traffic injury prevention*, *12*(5), 518-528.

Lord, R. G. (1985). Accuracy in behavioral measurement: An alternative definition based on raters' cognitive schema and signal detection theory. *Journal of Applied Psychology*, *70*(1), 66.

Los, S. A., Knol, D. L., & Boers, R. M. (2001). The foreperiod effect revisited: Conditioning as a basis for nonspecific preparation. *Acta Psychologica*, *106*(1-2), 121-145.

Luna, F. G., Barttfeld, P., Martín-Arévalo, E., & Lupiáñez, J. (2021). The ANTI-Vea task: Analyzing the executive and arousal vigilance decrements while measuring the three attentional networks. *Psicol´ogica, 42*(1), 1–26. https://doi.org/10.2478/psicolj-2021-0001

Luna, F. G., Barttfeld, P., Martín-Arévalo, E., & Lupiáñez, J. (2022). Cognitive load mitigates the executive but not the arousal vigilance decrement. *Consciousness and cognition*, *98*, 103263.

Luna, F. G., Marino, J., Roca, J., & Lupiáñez, J. (2018). Executive and arousal vigilance decrement in the context of the attentional networks: The ANTI-Vea task. *Journal of Neuroscience Methods*, *306*, 77–87. https://doi.org/10.1016/j.jneumeth.2018.05.011

Luna, F. G., Roca, J., Martín-Arévalo, E., & Lupiáñez, J. (2021). Measuring attention and vigilance in the laboratory vs. online: The split-half reliability of the ANTIVea. *Behavior Research Methods, 53*(3), 1124–1147. https://doi.org/10.3758/s13428-020-01483-4

Lynn, S. K., and Barrett, L. F. (2014). "Utilizing" signal detection theory. Psychol. Sci. 25, 1663–1673. doi: 10.1016/j.neuron.2009.10.017.A

Mackie, R. R. (1987). Vigilance research—Are we ready for countermeasures?. *Human factors*, *29*(6), 707-723.

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly journal of experimental psychology*, *1*(1), 6-21.

Mackworth, J. F. (1968). Vigilance, arousal, and habituation. *Psychological review*, *75*(4), 308.

MacLean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, & Psychophysics*, *71*(5), 1042-1058.

MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, *24*(5), 637.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and" nonparametric" indexes. *Psychological bulletin*, *107*(3), 401.

Mahr, A., Feld, M., Moniri, M. M., & Math, R. (2012). The contre (continuous tracking and reaction) task: A flexible approach for assessing driver cognitive workload with high sensitivity. Automotive user interfaces and interactive vehicular applications, 88-91.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind:: further investigations of sustained attention to response. *Neuropsychologia*, *37*(6), 661-670.

Marcum, J. I. (1947). *A statistical theory of target detection by pulsed radar*. RAND CORP SANTA MONICA CA.

Masoudian, M., & Razavi, H. (2019). An investigation of the required vigilance for different occupations. *Safety Science*, *119*, 353-359.

Math, R., Mahr, A., Moniri, M. M., & Müller, C. (2013). OpenDS: A new open-source driving simulator for research. GMM-Fachbericht-AmE 2013, 2.

Mathis, J., & Hess, C. W. (2009). Sleepiness and vigilance tests. *Swiss Med Wkly*, *139*(15-16), 214-219.

Matthews, G., Davies, D. R., & Holley, P. J. (1993). Cognitive predictors of vigilance. *Human factors*, *35*(1), 3-24.

Matthews, G., Warm, J. S., & Smith, A. P. (2017). Task engagement and attentional resources: Multivariate models for individual differences and stress factors in vigilance. Human factors, 59(1), 44-61.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three. *Personality psychology in Europe*, *7*, 335-350.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., ... & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. Emotion, 2(4), 315.

McBain, W. N. (1970). Arousal, monotony, and accidents in line driving. *Journal of Applied Psychology*, *54*(6), 509.

McBride, S. A., Merullo, D. J., Johnson, R. F., Banderet, L. E., & Robinson, R. T. (2007). Performance during a 3-hour simulated sentry duty task under varied work rates and secondary task demands. *Military psychology*, *19*(2), 103-117.

McCarley, J. S., & Yamani, Y. (2021). Psychometric curves reveal three mechanisms of vigilance decrement. *Psychological science*, *32*(10), 1675-1683.

McWilliams, T., & Ward, N. (2021). Undermulti-task on the road: measuring vigilance decrements during partially automated driving. Frontiers in psychology, 12, 631364.

Medeiros-Ward, N., Cooper, J. M., & Strayer, D. L. (2014). Hierarchical control and driving. *Journal of experimental psychology: General*, *143*(3), 953.

Mehrabi, E., & Kim, J. E. (2022, September). Physiological Measurements of Vigilance: A Systematic Review. In *Proceedings of the Human Factors and Ergonomics*

Society Annual Meeting (Vol. 66, No. 1, pp. 823-827). Sage CA: Los Angeles, CA: SAGE Publications.

Mehler, B., Reimer, B., Pohlmeyer, A. E., & Coughlin, J. F. (2008). The association between heart rate reactivity and driving performance under dual task demands in late middle age drivers. *Advances in Transportation Studies, An International Journal, Special Issue, 5370*.

Miall, R. C., Weir, D. J., & Stein, J. F. (1993). Intermittency in human manual tracking tasks. *Journal of motor behavior, 25*(1), 53-63.

Mirsky, A. F., Anthony, B. J., Duncan, C. C., Ahearn, M. B., & Kellam, S. G. (1991). Analysis of the elements of attention: A neuropsychological approach. *Neuropsychology review, 2*, 109-145.

Monsell, S. (2003). Task switching. Trends in cognitive sciences, 7(3), 134-140.

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: a review. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 67(1), 11.

Moran, T., Hughes, S., Hussey, I., Vadillo, M. A., Olson, M. A., Aust, F., ... & De Houwer, J. (2020). Incidental attitude formation via the surveillance task: A registered replication report of Olson and Fazio (2001).

Moray, N., & Haudegond, S. (1998, October). An absence of vigilance decrement in a complex dynamic task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, No. 3, pp. 234-236). Sage CA: Los Angeles, CA: SAGE Publications.

Muhrer, E., & Vollrath, M. (2011). The effect of visual and cognitive distraction on driver's anticipation in a simulated car following scenario. Transportation research part F: traffic psychology and behaviour, 14(6), 555-566.

Munnik, A., Näswall, K., Woodward, G., & Helton, W. S. (2020). The quick and the dead: A paradigm for studying friendly fire. *Applied ergonomics*, *84*, 103032.

Murray, D. J., & Ross, H. E. (1982). Vives (1538) on memory and recall. *Canadian Psychology/Psychologie canadienne*, *23*(1), 22.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological review*, *104*(1), 3.

Neigel, A. R., Claypoole, V. L., & Szalma, J. L. (2019). Effects of state motivation in overload and underload vigilance task scenarios. *Acta psychologica*, *197*, 106-114.

Neigel, A. R., Claypoole, V. L., Smith, S. L., Waldfogle, G. E., Fraulini, N. W., Hancock, G. M., ... & Szalma, J. L. (2020). Engaging the human operator: a review of the theoretical support for the vigilance decrement and a discussion of practical applications. *Theoretical issues in ergonomics science*, *21*(2), 239-258.

Nicholas, J. M. (1979). Leibniz: Apperception, Perception, and Thought. *Journal of the History of Philosophy*, *17*(1), 96-98.

Nichols, S. L., & Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of ADHD. *Child Psychiatry and Human Development*, *34*, 297-315.

Nickerson, R. S. (1968). Response time to the second of two signals following varied vs constant intersignal intervals. *Perception & Psychophysics*, *4*, 78-80.

Nicolae, F., Cotorcea, A., Ristea, M., & Atodiresei, D. (2016, January). Human Reliability Using the Fault Tree Analysis. A Case Study of a Military Accident Investigation. In *International conference Knowledge-Based Organization* (Vol. 22, No. 1, pp. 215-219).

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual sustained attention: Image degradation produces rapid sensitivity decrement over time. *Science*, *220*(4594), 327-329.

Ogundele, M. O., Ayyash, H. F., & Banerjee, S. (2011). Role of computerised continuous performance task tests in ADHD. *Progress in Neurology and Psychiatry*, *15*(3), 8-13.

Oken, B. S., Salinsky, M. C., & Elsas, S. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical neurophysiology*, *117*(9), 1885-1901.

Ord, A. S., Miskey, H. M., Lad, S., Richter, B., Nagy, K., & Shura, R. D. (2021). Examining embedded validity indicators in Conners continuous performance test-3 (CPT-3). *The Clinical Neuropsychologist*, *35*(8), 1426-1441.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. Science, 205(4409), 924-927.

Parasuraman, R., & Davies, D. R. (1977). A taxonomic analysis of vigilance performance. In vigilance (pp. 559-574). Springer, Boston, MA.

Parasuraman, R., & Giambra, L. (1991). Skill development in vigilance: effects of event rate and age. *Psychology and aging*, *6*(2), 155.

Parasuraman, R., & Mouloua, M. (1987). Interaction of signal discriminability and task type in vigilance decrement. Perception & Psychophysics, 41(1), 17-22.

Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: boredom or cognitive fatigue?. *Physiology & behavior*, *93*(1-2), 369-378.

Pavlov, I. P. (1902). *The Work of the Digestive Glands: Lectures by Professor JP Pavlow Tr. Into English by WH Thompson*. C. Griffin, Limited.

Peugh, J. L. (2010). A practical guide to multilevel modeling. Journal of school psychology, 48(1), 85-112.

Pylyshyn, Z. (1994). Some primitive mechanisms of spatial attention. *Cognition*, *50*(1-3), 363-384.

Pylyshyn, Z. W.; Storm, R. W. (1988). "Tracking multiple independent targets: Evidence for a parallel tracking mechanism". Spatial Vision. 3 (3): 179–197. doi:10.1163/156856888X00122.Poljac, E., Kiesel, A., Koch, I., & Müller, H. (2018). New perspectives on human multitasking. *Psychological Research*, *82*(1), 1-3.

Popp, R. F., Maier, S., Rothe, S., Zulley, J., Crönlein, T., Wetter, T. C., ... & Hajak, G. (2015). Impact of overnight traffic noise on sleep quality, sleepiness, and vigilant attention in long-haul truck drivers: results of a pilot study. *Noise & Health*, *17*(79), 387.

Posner, M. I. (1980). Orienting of attention. Quarterly journal of experimental psychology, 32(1), 3-25.

Posner, M. I. (2008). Measuring alertness. *Annals of the New York Academy of Sciences*, *1129*(1), 193-199.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological review*, *78*(5), 391.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual review of neuroscience*, *13*(1), 25-42.

Poudel, G. R., Jones, R. D., & Innes, C. R. (2008). A 2-D pursuit tracking task for behavioural detection of lapses. *Australasian Physical & Engineering Sciences in Medicine*, *31*(4), 528.

Psychology Software Tools, Inc. [E-Prime 3.0]. (2020). Retrieved from https://support.pstnet.com/.

Rakauskas, M. E., Gugerty, L. J., & Ward, N. J. (2004). Effects of naturalistic cell phone conversations on driving performance. *Journal of safety research*, *35*(4), 453-464.

Rajan, R., Selker, T., & Lane, I. (2016, March). Task load estimation and mediation using psycho-physiological measures. In Proceedings of the 21st international conference on intelligent user interfaces (pp. 48-59).

Ralph, B. C., Onderwater, K., Thomson, D. R., & Smilek, D. (2017). Disrupting monotony while increasing demand: Benefits of rest and intervening tasks on vigilance. *Psychological research*, *81*, 432-444.

Rann, J. C., & Almor, A. (2022). Effects of verbal tasks on driving simulator performance. Cognitive Research: Principles and Implications, 7(1), 1-26.

Razavi, T. (2001). Self-report measures: An overview of concerns and limitations of questionnaire use in occupational stress research.

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis & Prevention*, *43*(5), 1771-1781.

Reinerman-Jones, L., Matthews, G., & Mercado, J. E. (2016). Detection tasks in nuclear power plant operation: Vigilance decrement and physiological workload monitoring. *Safety science*, *88*, 97-107.

Reynolds, B., Penfold, R. B., & Patak, M. (2008). Dimensions of impulsive behavior in adolescents: laboratory behavioral assessments. *Experimental and clinical psychopharmacology*, *16*(2), 124.

Riccio, C. A., & Reynolds, C. R. (2001). Continuous performance tests are sensitive to ADHD in adults but lack specificity: A review and critique for differential diagnosis. *Annals of the New York academy of sciences*, *931*(1), 113-139.

Riccio, C. A., Reynolds, C. R., Lowe, P., & Moore, J. J. (2002). The continuous performance test: a window on the neural substrates for attention?. *Archives of clinical neuropsychology*, *17*(3), 235-272.

Riccio, C. A., Waldrop, J. J., Reynolds, C. R., & Lowe, P. (2001). Effects of stimulants on the continuous performance test (CPT) implications for CPT use and interpretation. *The Journal of neuropsychiatry and clinical neurosciences*, *13*(3), 326-335.

Roach, G. D., Dawson, D., & Lamond, N. (2006). Can a shorter psychomotor vigilance task be usedas a reasonable substitute for the ten-minute psychomotor vigilance task?. *Chronobiology international*, *23*(6), 1379-1387.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia, 35(6), 747-758.

Roca, J., Castro, C., López-Ramón, M. F., & Lupiánez, J. (2011). Measuring vigilance while assessing the functioning of the three attentional networks: The ANTI-Vigilance task. *Journal of neuroscience methods*, *198*(2), 312-324.

Roebuck, H., Freigang, C., & Barry, J. G. (2016). Continuous performance tasks: Not just about sustaining attention. *Journal of speech, language, and hearing research*, *59*(3), 501-510.

Rosenberg, M., Noonan, S., DeGutis, J., & Esterman, M. (2013). Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. *Attention, Perception, & Psychophysics*, *75*, 426-439.

Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., & Chun, M. M. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature neuroscience*, *19*(1), 165-171.

Ross, H. A., Russell, P. N., & Helton, W. S. (2014). Effects of breaks and goal switches on the vigilance decrement. *Experimental brain research*, *232*, 1729-1737.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome Jr, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. Journal of consulting psychology, 20(5), 343.

Rubinstein, J. S. (2020). Divergent response-time patterns in vigilance decrement tasks. *Journal of experimental psychology: human perception and performance*, *46*(10), 1058.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of experimental psychology: human perception and performance*, *27*(4), 763.

Salvucci, D. D., & Taatgen, N. A. (2011). Toward a unified view of cognitive control. *Topics in cognitive science*, *3*(2), 227-230.

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain research reviews*, *35*(2), 146-160.

Saxby, D. J., Matthews, G., Warm, J. S., Hitchcock, E. M., & Neubauer, C. (2013). Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *Journal of experimental psychology: applied*, *19*(4), 287.

Scerbo, M. W. (1998, October). Sources of stress and boredom in vigilance. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 42, No. 10, pp. 764-768). Sage CA: Los Angeles, CA: SAGE Publications.

Scerbo, M. W., Greenwald, C. Q., & Sawin, D. A. (1992, October). Vigilance: It's boring, it's difficult, and I can't do anything about it. In *Proceedings of the Human Factors Society Annual Meeting* (Vol. 36, No. 18, pp. 1508-1512). Sage CA: Los Angeles, CA: SAGE Publications.

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis & Prevention*, *41*(5), 1087-1093.

Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological science, 12*(2), 101-108.

Scullin, M. K., & Bugg, J. M. (2013). Failing to forget: prospective memory commission errors can result from spontaneous retrieval and impaired executive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 965.

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological bulletin*, *117*(2), 230.

See, J. E., Warm, J. S., Dember, W. N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors*, *39*(1), 14-29.

Seli, P., Risko, E. F., & Smilek, D. (2016). On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychological science*, *27*(5), 685-691.

Seli, P., Jonker, T. R., Solman, G. J., Cheyne, J. A., & Smilek, D. (2013). A methodological note on evaluating performance in a sustained-attention-to-response task. *Behavior research methods, 45*(2), 355-363.

Shallice, T., Stuss, D. T., Alexander, M. P., Picton, T. W., & Derkzen, D. (2008). The multiple dimensions of sustained attention. Cortex, 44(7), 794–805. https://doi.org/10.1016/j.cortex.2007.04.002

Shappell, S. A., & Wiegmann, D. A. (2004, November). HFACS analysis of military and civilian aviation accidents: A North American comparison. In *Proceedings of the Annual Meeting of the International Society of Air Safety Investigators* (pp. 2-8). Australia: Gold Coast.

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.

Silverstein, M. L., Weinstein, M., & Turnbull, A. (2004). Nonpatient CPT performance varying target frequency and interstimulus interval on five response measures. *Archives of Clinical Neuropsychology*, *19*(8), 1017-1025.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. Psychological bulletin, 132(6), 946.

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. Consciousness and cognition, 13(4), 657-690.

Smid, H. G. O. M., De Witte, M. R., Homminga, I., & Van Den Bosch, R. J. (2006). Sustained and transient attention in the continuous performance task. *Journal of clinical and experimental neuropsychology*, *28*(6), 859-883.

Smith, S. L., Helton, W. S., Matthews, G., & Funke, G. J. (2021). Performance, hemodynamics, and stress in a two-day vigilance task: practical and theoretical implications. *Human Factors*, 00187208211011333.

Smith, V., Maslovat, D., Drummond, N. M., & Carlsen, A. N. (2019). A timeline of motor preparatory state prior to response initiation: evidence from startle. *Neuroscience*, *397*, 80-93.

St. John, M., Harris, W. C., & Osga, G. (1997, October). Designing for multi-tasking environments: Multiple monitors vs. multiple windows. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 41, No. 2, pp. 1313-1317). Sage CA: Los Angeles, CA: SAGE Publications.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. Behavior research methods, instruments, & computers, 31(1), 137-149.

Stearman, E. J., & Durso, F. T. (2016). Vigilance in a dynamic environment. *Journal of experimental psychology: applied*, *22*(1), 107.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta psychologica*, *30*, 276-315.

Stevenson, H., Russell, P. N., & Helton, W. S. (2011). Search asymmetry, sustained attention, and response inhibition. *Brain and cognition*, *77*(2), 215-222.

Strayer, D. L., & Fisher, D. L. (2016). SPIDER: A framework for understanding driver distraction. Human factors, 58(1), 5-12.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of experimental psychology: Applied*, *9*(1), 23.

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. In Psychology of learning and motivation (Vol. 54, pp. 29-58). Academic Press.

Strayer, D. L., Cooper, J. M., McCarty, M. M., Getty, D. J., Wheatley, C. L., Motzkus, C. J., et al. (2019). Visual and cognitive demands of carplay, android auto, and five native infotainment systems. Hum. Factors J. Hum. Factors Ergon. Soc. 61, 1371–1386. doi: 10.1177/0018720819836575

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: their ROCs and implied models. *Psychological bulletin*, *99*(1), 100.

Tatham, A. J., Boer, E. R., Rosen, P. N., Della Penna, M., Meira-Freitas, D., Weinreb, R. N., ... & Medeiros, F. A. (2014). Glaucomatous retinal nerve fiber layer thickness loss is associated with slower reaction times under a divided attention task. *American journal of ophthalmology*, *158*(5), 1008-1017.

Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human factors*, *16*(4), 339-352.

Thomas, M. L., & Russo, M. B. (2007). Neurocognitive monitors: toward the prevention of cognitive performance decrements and catastrophic failures in the operational environment. Aviation, space, and environmental medicine, 78(5), B144-B152.

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on psychological science*, *10*(1), 82-96.

Thomson, D. R., Besner, D., & Smilek, D. (2016). A critical examination of the evidence for sensitivity loss in modern vigilance tasks. *Psychological review*, *123*(1), 70.

Tillman, G., Strayer, D., Eidels, A., & Heathcote, A. (2017). Modeling cognitive load effects of conversation between a passenger and driver. *Attention, Perception, & Psychophysics, 79*(6), 1795-1803.

Tiwari, T., Singh, A. L., & Singh, I. L. (2009). Task demand and workload: Effects on vigilance performance and stress. *Journal of the Indian Academy of Applied Psychology*, *35*(2), 265-275.

Tombu, M., & Seiffert, A. E. (2008). Attentional costs in multiple-object tracking. *Cognition*, *108*(1), 1-25.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological review*, *76*(3), 282.

Tsang, S. N., & Chan, A. H. (2018). Tracking and discrete dual task performance for different visual spatial stimulus-response mappings with focal and ambient vision. *Applied ergonomics*, *67*, 39-49.

Tucha, L., Tucha, O., Walitza, S., Sontag, T. A., Laufkötter, R., Linder, M., & Lange, K. W. (2009). Vigilance and sustained attention in children and adults with ADHD. *Journal of attention disorders*, *12*(5), 410-421.

van Schie, M. K., Lammers, G. J., Fronczek, R., Middelkoop, H. A., & van Dijk, J. G. (2021). Vigilance: discussion of related concepts and proposal for a definition. *Sleep Medicine*, *83*, 175-181.

Vardaki, S., Yannis, G., & Papageorgiou, S. G. (2014). Assessing selected cognitive impairments using a driving simulator: a focused review. *Advances in Transportation Studies,* (34).

Vater, C., Kredel, R., & Hossner, E. J. (2017). Disentangling vision and attention in multiple-object tracking: How crowding and collisions affect gaze anchoring and dual-task performance. *Journal of vision*, *17*(5), 21-21.

Warm, J. S., Matthews, G., & Finomore Jr, V. S. (2018). Vigilance, workload, and stress. In Performance under stress (pp. 131-158). CRC Press.

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human factors*, *50*(3), 433-441.

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior research methods*, *50*, 642-661.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, *3*(2), 159-177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human factors*, *50*(3), 449-455.

Wiener, E. L., Curry, R. E., & Faustina, M. L. (1984). Vigilance and task load: In search of the inverted U. *Human factors*, *26*(2), 215-222.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*(2), 229-235.

Wijayaratna, K. P., Cunningham, M. L., Regan, M. A., Jian, S., Chand, S., & Dixit, V. V. (2019). Mobile phone conversation distraction: Understanding differences in impact between simulator and naturalistic driving studies. *Accident Analysis & Prevention*, *129*, 108-118.

Williams, P. G., Rau, H. K., Suchy, Y., Thorgusen, S. R., & Smith, T. W. (2017). On the validity of self-report assessment of cognitive abilities: Attentional control scale

associations with cognitive performance, emotional adjustment, and personality. *Psychological Assessment*, *29*(5), 519.

Wilson, K. M. (2015). Friendly fire and the Sustained Attention to Response Task: Using basic laboratory research to investigate a real-world problem.

Wilson, K. M., Finkbeiner, K. M., De Joux, N. R., Russell, P. N., & Helton, W. S. (2016). Go-stimuli proportion influences response strategy in a sustained attention to response task. *Experimental brain research*, *234*, 2989-2998.

Winter, B. (2013). A very basic tutorial for performing linear mixed effects analyses. arXiv preprint arXiv:1308.5499.

Winter, B., & Wieling, M. (2016). How to analyze linguistic change using mixed models, Growth Curve Analysis and Generalized Additive Modeling. Journal of Language Evolution, 1(1), 7-18.

Wodka, E. L., Simmonds, D. J., Mahone, E. M., & Mostofsky, S. H. (2009). Moderate variability in stimulus presentation improves motor response control. *Journal of clinical and experimental neuropsychology*, *31*(4), 483-488.

Wundt, W. (1880). Grundzüge der physiologischen Psychologie [Principles of physiological psychology] (2nd ed., Vol. 1). Engelmann. http://vlp.mpiwg-berlin.mpg.de/references?idlit575

Wyatt, S., & Langdon, J. N. (1932). Inspection Processes in Industry.(A Preliminary Report). *Inspection Processes in Industry.(A Preliminary Report).*, (63).

Yanko, M. R., & Spalek, T. M. (2014). Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human factors*, *56*(2), 260-269.

Zhang, Y., & Kumada, T. (2017). Relationship between workload and mind-wandering in simulated driving. *PloS one*, *12*(5), e0176962.

Zuercher, J. D. (1965). The effects of extraneous stimulation on vigilance. *Human factors*, *7*(2), 101-105.

APPENDIX A: GROWTH CURVE MODELS

Table A.1. Growth Curve Models for Fitting CPT, SDT, Tracking measures for Participant *i* at Block Time-Point *j* for E1 and E2.

Model	Equation
1. Base	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1$ $\beta_1 = \zeta_2$ $\beta_2 = \zeta_3$
2. Intercept	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1 * \text{Session} * \text{ISI}$ $\beta_1 = \zeta_2$ $\beta_2 = \zeta_3$
3. Linear	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1 * \text{Session} * \text{ISI}$ $\beta_1 = \zeta_2 * \text{Session} * \text{ISI}$ $\beta_2 = \zeta_3$
4. Quadratic	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1 * \text{Session} * \text{ISI}$ $\beta_1 = \zeta_2 * \text{Session} * \text{ISI}$ $\beta_2 = \zeta_3 * \text{Session} * \text{ISI}$

Table A.2. Growth Curve Models for Fitting CPT, SDT, Tracking measures for Participant *i* at Block Time-Point *j* for E3.

Model	Equation
1. Base	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1$ $\beta_1 = \zeta_2$ $\beta_2 = \zeta_3$ $\beta_3 = \zeta_4$ $\beta_4 = \zeta_5$
2. Intercept	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1$ * Session * ISI * SegType $\beta_1 = \zeta_2$ $\beta_2 = \zeta_3$ $\beta_3 = \zeta_4$ $\beta_4 = \zeta_5$
3. Linear	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1$ * Session * ISI * SegType $\beta_1 = \zeta_2$ * Session * ISI * SegType $\beta_2 = \zeta_3$ $\beta_3 = \zeta_4$ $\beta_4 = \zeta_5$
4. Quadratic	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1$ * Session * ISI * SegType $\beta_1 = \zeta_2$ * Session * ISI * SegType $\beta_2 = \zeta_3$ * Session * ISI * SegType $\beta_3 = \zeta_4$ $\beta_4 = \zeta_5$
5. Cubic	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1$ * Session * ISI * SegType $\beta_1 = \zeta_2$ * Session * ISI * SegType $\beta_2 = \zeta_3$ * Session * ISI * SegType $\beta_3 = \zeta_4$ * Session * ISI * SegType $\beta_4 = \zeta_5$

6. Quartic	$Y_{ij} = \beta_{0i} + \beta_1 * \text{Block}_j + \beta_2 * \text{Block}_j^2 + \varepsilon_i$ $\beta_{0i} = \beta_0 + \zeta_{0i}$ $\beta_0 = \zeta_1 * \text{Session} * \text{ISI} * \text{SegType}$ $\beta_1 = \zeta_2 * \text{Session} * \text{ISI} * \text{SegType}$ $\beta_2 = \zeta_3 * \text{Session} * \text{ISI} * \text{SegType}$ $\beta_3 = \zeta_4 * \text{Session} * \text{ISI} * \text{SegType}$ $\beta_4 = \zeta_5 * \text{Session} * \text{ISI} * \text{SegType}$

APPENDIX B: E1 TARGET HIT RATE

Table B.1 Model Comparison for Target Hit Rate in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	3767.2	3795.3	-1878.6	3757.2				
Intercept	10	3766.8	3822.9	-1873.4	3746.8	10.3823	5	6.51E-02	.
Linear	15	3773.4	3857.5	-1871.7	3743.4	3.4269	5	0.6345	
Quadratic	20	3781.8	3894	-1870.9	3741.8	1.6092	5	0.9001	

Table B.2 Intercept Model Coefficients for Target Hit Rate in E1.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	11.8125	0.05262	118.55	224.47	<2e-16	***
time^1	0.07612	0.03234	1960	2.35	0.0187	*
time^2	-0.03929	0.03234	1960	-1.22	0.2246	
sessionNameMulti-Task	-0.05952	0.04574	1960	-1.30	0.1933	
ISI2000	0.03571	0.04574	1960	0.78	0.435	
ISI4000	0.05952	0.04574	1960	1.30	0.1933	
sessionNameMulti-Task:ISI2000	0.0119	0.06468	1960	0.18	0.854	
sessionNameMulti-Task:ISI4000	0.05952	0.06468	1960	0.92	0.3576	

APPENDIX C: E1 OMISSIONS

Table C.1 Model Comparison for Omissions in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*
Base	5	3644.5	3672.6	-1817.2	3634.5			
Intercept	10	3649.2	3705.3	-1814.6	3629.2	5.2921	5	3.81E-01
Linear	15	3654.5	3738.6	-1812.2	3624.5	4.7275	5	0.45
Quadratic	20	3662.2	3774.3	-1811.1	3622.2	2.3227	5	0.8029

Table C.2 Base Model Coefficients for Omissions in E1.

Fixed Effects	Estimate	Std. Error	df	t value	p	
(Intercept)	0.15774	0.04273	55.999	3.692	0.000507	***
$time^1$	-0.07541	0.03145	1960	-2.398	0.01657	*
$time^2$	0.04189	0.03145	1960	1.332	0.182975	

APPENDIX D: E1 COMMISSIONS

Table D.1 Model Comparison for Commissions in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	4383.6	4411.6	-2186.8	4373.6				
Intercept	10	4337.8	4393.8	-2158.9	4317.8	55.8369	5	8.78E-11	***
Linear	15	4340.4	4424.6	-2155.2	4310.4	7.3042	5	0.19898	
Quadratic	20	4340.8	4453	-2150.4	4300.8	9.6733	5	0.08504	.

Table D.2 Quadratic Model Coefficients for Commissions in E1.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	6.79E-01	7.06E-02	9.36E+01	9.612	1.28E-15	***
time^1	-1.02E-01	9.03E-02	1.96E+03	-1.134	0.25686	
sessionNameMulti-Task	1.67E-01	5.22E-02	1.96E+03	3.196	0.00142	**
ISI2000	-5.95E-03	5.22E-02	1.96E+03	-0.114	0.90914	
ISI4000	-9.52E-02	5.22E-02	1.96E+03	-1.826	0.06797	.
time^2	1.23E-01	9.03E-02	1.96E+03	1.359	0.17434	
time^1:sessionNameMulti-Task	-9.39E-02	1.28E-01	1.96E+03	-0.735	0.46234	
time^1:ISI2000	1.37E-01	1.28E-01	1.96E+03	1.069	0.28506	
time^1:ISI4000	1.96E-01	1.28E-01	1.96E+03	1.537	0.12442	
sessionNameMulti-Task:ISI2000	3.87E-02	7.38E-02	1.96E+03	0.525	0.59993	
sessionNameMulti-Task:ISI4000	1.16E-01	7.38E-02	1.96E+03	1.574	0.1157	
sessionNameMulti-Task:time^2	2.53E-02	1.28E-01	1.96E+03	0.198	0.84285	
ISI2000:time^2	-6.62E-02	1.28E-01	1.96E+03	-0.519	0.60412	
ISI4000:time^2	-1.48E-01	1.28E-01	1.96E+03	-1.159	0.24653	
time^1:sessionNameMulti-Task:ISI2000	7.47E-02	1.81E-01	1.96E+03	0.413	0.67929	
time^1:sessionNameMulti-Task:ISI4000	-1.13E-01	1.81E-01	1.96E+03	-0.626	0.53129	
sessionNameMulti-Task:ISI2000:time^2	-1.13E-01	1.81E-01	1.96E+03	-0.626	0.5317	
sessionNameMulti-Task:ISI4000:time^2	3.04E-01	1.81E-01	1.96E+03	1.682	0.09264	.

APPENDIX E: E1 CORRECT REJECTIONS

Table E.1 Model Comparison for Correct Rejections in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	4395.5	4423.5	-2192.8	4385.5				
Intercept	10	4348.7	4404.8	-2164.4	4328.7	56.7467	5	5.70E-11	***
Linear	15	4351.1	4435.2	-2160.5	4321.1	7.6733	5	0.17518	
Quadratic	20	4350.7	4462.8	-2155.3	4310.7	10.4181	5	0.06422	.

Table E.2 Quadratic Model Coefficients for Correct Rejections in E1.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	2.32E+00	7.10E-02	9.33E+01	32.667	2.00E-16	***
$time^1$	1.09E-01	9.05E-02	1.96E+03	1.202	0.22942	
sessionNameMulti-Task	-1.67E-01	5.23E-02	1.96E+03	-3.188	0.00145	**
ISI2000	5.95E-03	5.23E-02	1.96E+03	0.114	0.90935	
ISI4000	9.82E-02	5.23E-02	1.96E+03	1.879	0.06041	.
$time^2$	-1.21E-01	9.05E-02	1.96E+03	-1.334	0.1823	
$time^1$:sessionNameMulti-Task	9.82E-02	1.28E-01	1.96E+03	0.767	0.44332	
$time^1$:ISI2000	-1.37E-01	1.28E-01	1.96E+03	-1.067	0.28619	
$time^1$:ISI4000	-2.03E-01	1.28E-01	1.96E+03	-1.584	0.11346	
sessionNameMulti-Task:ISI2000	-3.87E-02	7.39E-02	1.96E+03	-0.523	0.60078	
sessionNameMulti-Task:ISI4000	-1.22E-01	7.39E-02	1.96E+03	-1.651	0.09898	.
sessionNameMulti-Task:$Time^2$	-3.70E-02	1.28E-01	1.96E+03	-0.289	0.77253	
ISI2000:$time^2$	6.62E-02	1.28E-01	1.96E+03	0.517	0.60497	
ISI4000:$time^2$	1.46E-01	1.28E-01	1.96E+03	1.141	0.25391	
$time^1$:sessionNameMulti-Task:ISI2000	-8.32E-02	1.81E-01	1.96E+03	-0.46	0.6458	
$time^1$:sessionNameMulti-Task:ISI4000	1.13E-01	1.81E-01	1.96E+03	0.625	0.53225	
sessionNameMulti-Task:ISI2000:$time^2$	1.31E-01	1.81E-01	1.96E+03	0.721	0.47106	
sessionNameMulti-Task:ISI4000:$time^2$	-3.00E-01	1.81E-01	1.96E+03	-1.657	0.09768	.

APPENDIX F: E1 REACTION TIME

Table F.1 Model Comparison for Reaction Time in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-2365.6	-2337.5	1187.8	-2375.6				
Intercept	10	-3374.5	-3318.5	1697.3	-3394.5	1018.97	5	< 2e-16	***
Linear	15	-3376.2	-3292	1703.1	-3406.2	11.6346	5	0.0402	*
Quadratic	20	-3372.8	-3260.6	1706.4	-3412.8	6.6411	5	0.2487	

Table F.2 Linear Model Coefficients for Reaction Time in E1.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	5.90E+00	1.75E-02	6.60E+01	338.031	<2e-16	***
time^1	1.16E-02	1.31E-02	1.96E+03	0.885	3.76E-01	
sessionNameMulti-Task	1.15E-01	7.58E-03	1.96E+03	15.158	<2e-16	***
ISI2000	6.32E-02	7.58E-03	1.96E+03	8.333	<2e-16	***
ISI4000	1.21E-01	7.58E-03	1.96E+03	15.932	<2e-16	***
time^2	-1.36E-02	5.36E-03	1.96E+03	-2.533	1.14E-02	*
time^1:sessionNameMulti-Task	1.92E-02	1.86E-02	1.96E+03	1.033	0.3015	
time^1:ISI2000	9.25E-03	1.86E-02	1.96E+03	0.498	0.6185	
time^1:ISI4000	3.71E-02	1.86E-02	1.96E+03	1.999	0.0457	*
sessionNameMulti-Task:ISI2000	2.21E-02	1.07E-02	1.96E+03	2.058	0.0397	*
sessionNameMulti-Task:ISI4000	5.20E-03	1.07E-02	1.96E+03	0.485	0.6275	
time^1:sessionNameMulti-Task:ISI2000	2.67E-03	2.63E-02	1.96E+03	0.101	0.9192	
time^1:sessionNameMulti-Task:ISI4000	-9.56E-04	2.63E-02	1.96E+03	-0.036	0.971	

APPENDIX G: E1 SENSITIVITY

Table G.1 Model Comparison for Sensitivity in E1.

Model		AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	3634.3	3662.3	-1812.1	3624.3				
Intercept	10	3582.1	3638.2	-1781.1	3562.1	62.1324	5	4.40E-12	***
Linear	15	3586.7	3670.9	-1778.4	3556.7	5.4105	5	0.36786	
Quadratic	20	3585.1	3697.2	-1772.5	3545.1	11.6753	5	0.03952	*

Table G.2 Quadratic Model Coefficients for Sensitivity in E1.

Fixed Effects	Estimate	Std. Error	df	t value	p	
(Intercept)	2.30405	0.05972	91.39759	38.583	2.00E-16	***
$time^1$	0.12162	0.07483	1960	1.625	0.10428	
sessionNameMulti-Task	-0.15443	0.04321	1960	-3.574	0.00036	***
ISI2000	0.02291	0.04321	1960	0.53	0.59602	
ISI4000	0.09538	0.04321	1960	2.208	0.02739	*
$time^2$	-0.12735	0.07483	1960	-1.702	0.08897	.
$time^1$:sessionNameMulti-Task	0.04608	0.10583	1960	0.435	0.6633	
$time^1$:ISI2000	-0.10078	0.10583	1960	-0.952	0.34106	
$time^1$:ISI4000	-0.15912	0.10583	1960	-1.504	0.13287	
sessionNameMulti-Task:ISI2000	-0.03157	0.0611	1960	-0.517	0.6054	
sessionNameMulti-Task:ISI4000	-0.07442	0.0611	1960	-1.218	0.22339	
sessionNameMulti-Task:$time^2$	-0.01859	0.10583	1960	-0.176	0.86057	
ISI2000:$time^2$	0.07111	0.10583	1960	0.672	0.50172	
ISI4000:$time^2$	0.15675	0.10583	1960	1.481	0.13873	
$time^1$:sessionNameMulti-Task:ISI2000	-0.07093	0.14967	1960	-0.474	0.63562	
$time^1$:sessionNameMulti-Task:ISI4000	0.06007	0.14967	1960	0.401	0.68819	
sessionNameMulti-Task:ISI2000:$time^2$	0.11334	0.14967	1960	0.757	0.44898	
sessionNameMulti-Task:ISI4000:$time^2$	-0.27168	0.14967	1960	-1.815	0.06965	.

APPENDIX H: E1 RESPONSE BIAS

Table H.1 Model Comparison for Response Bias in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	df	p	
Base	5	-1410.5	-1382.5	710.26	-1420.5				
Intercept	10	-1411.6	-1355.5	715.82	-1431.6	11.1237	5	4.90E-02	*
Linear	15	-1403.8	-1319.7	716.9	-1433.8	2.1669	5	0.8256	
Quadratic	20	-1396.5	-1284.3	718.23	-1436.5	2.6558	5	0.75287	

Table H.2 Intercept Model Coefficients for Response Bias in E1.

Fixed Effects	Estimate	Std. Error	df	t value	p	
(Intercept)	-4.80E-01	1.19E-02	2.05E+02	-40.471	<2e-16	***
time^1	8.16E-03	9.04E-03	1.96E+03	0.902	0.3672	
time^2	-1.74E-03	9.04E-03	1.96E+03	-0.192	0.8479	
sessionNameMulti-Task	1.43E-02	1.28E-02	1.96E+03	1.121	0.2625	
ISI2000	-9.03E-03	1.28E-02	1.96E+03	-0.706	0.4803	
ISI4000	5.54E-03	1.28E-02	1.96E+03	0.433	0.6652	
sessionNameMulti-Task:ISI2000	-7.18E-03	1.81E-02	1.96E+03	-0.397	0.6914	
sessionNameMulti-Task:ISI4000	-4.51E-02	1.81E-02	1.96E+03	-2.492	0.0128	*

APPENDIX I: E1 TRACKING DISTANCE

Table I.1 Model Comparison for Tracking Distance in E1.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-2278.6	-2254	1144.3	-2288.6				
Intercept	7	-2314.7	-2280.2	1164.3	-2328.7	40.0739	2	1.99E-09	***
Linear	9	-2317.1	-2272.9	1167.5	-2335.1	6.4473	2	0.03981	*
Quadratic	11	-2313.3	-2259.2	1167.7	-2335.3	0.1864	2	0.91103	

Table I.2 Linear Model Coefficients for Tracking Distance in E1.

Fixed Effects	Estimate	Std. Error	df	*t* value	*p*	
(Intercept)	0.294836	0.010679	66.808	27.609	< 2e-16	***
$time^1$	-0.032734	0.009323	952	-3.511	0.000467	***
ISI2000	-0.021692	0.005383	952	-4.03	6.02E-05	***
ISI4000	-0.034138	0.005383	952	-6.342	3.49E-10	***
$time^2$	0.015569	0.005383	952	2.892	0.003909	**
$time^1$:ISI2000	0.031073	0.013185	952	2.357	0.018637	*
$time^1$:ISI4000	0.026458	0.013185	952	2.007	0.045062	*

APPENDIX J: E2 TARGET HIT RATE

Table J.1 Model Comparison for Target Hit Rate in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	2645.9	2672.2	-1317.9	2635.9				
Intercept	8	2626.5	2668.7	-1305.2	2610.5	25.4082	3	1.27E-05	***
Linear	11	2629.2	2687.2	-1303.6	2607.2	3.3253	3	0.3441	
Quadratic	14	2630.7	2704.5	-1301.3	2602.7	4.4461	3	0.2171	

Table J.2 Intercept Model Coefficients for Target Hit Rate in E2.

Fixed Effects	Estimate	Std. Error	df	*t* value	*p*	
(Intercept)	1.59E+01	4.91E-02	1.16E+02	323.89	< 2e-16	***
time1	1.64E-02	3.71E-02	1.38E+03	0.443	0.6576	
time2	-8.64E-02	3.71E-02	1.38E+03	-2.33	0.0199	*
sessionNameMulti-Task	-1.86E-01	4.28E-02	1.38E+03	-4.348	1.48E-05	***
ISI4000	2.78E-03	4.28E-02	1.38E+03	0.065	0.9483	
sessionNameMulti-Task:ISI4000	1.31E-01	6.05E-02	1.38E+03	2.157	0.0312	*

APPENDIX K: E2 OMISSIONS

Table K.1 Model Comparison for Omissions in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	2479.6	2506	-1234.8	2469.6				
Intercept	8	2467.1	2509.3	-1225.6	2451.1	18.47	3	0.0003517	***
Linear	11	2470.4	2528.4	-1224.2	2448.4	2.763	3	0.429586	
Quadratic	14	2471	2544.8	-1221.5	2443	5.388	3	0.1454808	

Table K.2 Base Model Coefficients for Omissions in E2.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	7.50E-02	4.49E-02	1.24E+02	1.669	0.097615	.
time1	-3.19E-02	3.52E-02	1.38E+03	-0.907	0.364695	
time2	9.00E-02	3.52E-02	1.38E+03	2.561	0.010547	*
sessionNameMulti-Task	1.56E-01	4.06E-02	1.38E+03	3.833	0.000133	***
ISI4000	8.33E-03	4.06E-02	1.38E+03	0.205	0.837357	
sessionNameMulti-Task:ISI4000	-1.11E-01	5.74E-02	1.38E+03	-1.936	0.053106	.

APPENDIX L: E2 COMMISSIONS

Table L.1 Model Comparison for Commissions in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	3362.2	3388.6	-1676.1	3352.2				
Intercept	8	3357.1	3399.3	-1670.5	3341.1	11.1337	3	1.10E-02	*
Linear	11	3362.5	3420.5	-1670.2	3340.5	0.5998	3	0.89648	
Quadratic	14	3366.4	3440.2	-1669.2	3338.4	2.1368	3	0.54451	

Table L.2 Quadratic Model Coefficients for Commissions in E2

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	6.89E-01	6.94E-02	1.02E+02	9.933	< 2e-16	***
time^1	-2.01E-01	4.75E-02	1.38E+03	-4.224	2.56E-05	***
time^2	1.47E-01	4.75E-02	1.38E+03	3.09	0.00204	**
sessionNameMulti-Task	3.06E-02	5.49E-02	1.38E+03	0.557	0.57769	
ISI4000	-9.72E-02	5.49E-02	1.38E+03	-1.772	0.07663	.
sessionNameMulti-Task:ISI4000	1.47E-01	7.76E-02	1.38E+03	1.897	0.058	.

APPENDIX M: E2 CORRECT REJECTIONS

Table M.1 Model Comparison for Correct Rejections in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	3376.1	3402.5	-1683	3366.1				
Intercept	8	3370.4	3412.6	-1677.2	3354.4	11.7122	3	8.44E-03	**
Linear	11	3375.8	3433.8	-1676.9	3353.8	0.614	3	0.893217	
Quadratic	14	3379.7	3453.5	-1675.8	3351.7	2.0767	3	0.556641	

Table M.2 Quadratic Model Coefficients for Correct Rejections in E2.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	3.31E+00	7.02E-02	1.01E+02	47.16	< 2e-16	***
$time^1$	1.94E-01	4.77E-02	1.38E+03	4.07	4.96E-05	***
$time^2$	-1.52E-01	4.77E-02	1.38E+03	-3.182	0.00149	**
sessionNameMulti-Task	-3.33E-02	5.51E-02	1.38E+03	-0.605	0.54531	
ISI4000	9.72E-02	5.51E-02	1.38E+03	1.764	0.07788	.
sessionNameMulti-Task:ISI4000	-1.50E-01	7.79E-02	1.38E+03	-1.925	0.05444	.

APPENDIX N: E2 REACTION TIME

Table N.1 Model Comparison for Reaction Time in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-1329.2	-1302.8	669.59	-1339.2				
Intercept	8	-2432.2	-2390	1224.09	-2448.2	1108.994	3	< 2.2e-16	***
Linear	11	-2445.8	-2387.8	1233.89	-2467.8	19.6127	3	0.0002042	***
Quadratic	14	-2441.7	-2367.9	1234.85	-2469.7	1.9225	3	0.5886372	

Table N.2 Linear Model Coefficients for Reaction Time in E2.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	5.92E+00	1.57E-02	7.04E+01	377.28	<2e-16	***
$time^1$	2.79E-02	1.23E-02	1.38E+03	2.269	2.34E-02	*
sessionNameMulti-Task	1.28E-01	7.11E-03	1.38E+03	17.957	<2e-16	***
ISI4000	1.77E-01	7.11E-03	1.38E+03	24.972	<2e-16	***
$time^2$	-1.16E-02	6.15E-03	1.38E+03	-1.887	5.94E-02	.
$time^1$:sessionNameMulti-Task	-2.66E-02	1.74E-02	1.38E+03	-1.526	1.27E-01	
$time^1$:ISI4000	1.72E-02	1.74E-02	1.38E+03	0.989	0.323	
sessionNameMulti-Task:ISI4000	-1.42E-02	1.01E-02	1.38E+03	-1.411	0.1585	
$time^1$:sessionNameMulti-Task:ISI4000	5.81E-02	2.46E-02	1.38E+03	2.361	0.0183	*

APPENDIX O: E2 SENSITIVITY

Table O.1 Model Comparison for Sensitivity in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	2389.3	2415.7	-1189.7	2379.3				
Intercept	8	2378.9	2421.1	-1181.5	2362.9	16.4283	3	9.26E-04	***
Linear	11	2384.5	2442.5	-1181.3	2362.5	0.3587	3	0.9486287	
Quadratic	14	2390.2	2464	-1181.1	2362.2	0.3049	3	0.9591043	

Table O.2 Intercept Model Coefficients for Sensitivity in E2.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	2.66E+00	5.02E-02	9.99E+01	52.974	< 2e-16	***
time^1	1.40E-01	3.38E-02	1.38E+03	4.128	3.87E-05	***
time^2	-1.20E-01	3.38E-02	1.38E+03	-3.563	0.00038	***
sessionNameMulti-Task	-6.95E-02	3.90E-02	1.38E+03	-1.779	0.07539	.
ISI4000	5.98E-02	3.90E-02	1.38E+03	1.531	0.12594	
sessionNameMulti-Task:ISI4000	-6.86E-02	5.52E-02	1.38E+03	-1.243	0.21422	

APPENDIX P: E2 RESPONSE BIAS

Table P.1 Model Comparison for Response Bias in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-796.04	-769.68	403.02	-806.04				
Intercept	8	-799.72	-757.54	407.86	-815.72	9.6771	3	0.02152	*
Linear	11	-796.11	-738.11	409.06	-818.11	2.3912	3	0.49527	
Quadratic	14	-794.31	-720.5	411.16	-822.31	4.2024	3	0.24042	

Table P.2 Intercept Model Coefficients for Response Bias in E2.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	-5.38E-01	1.32E-02	1.50E+02	-40.782	< 2e-16	***
time1	1.85E-02	1.14E-02	1.38E+03	1.628	0.1037	
time2	9.09E-03	1.14E-02	1.38E+03	0.799	0.4245	
sessionNameMulti-Task	3.11E-02	1.31E-02	1.38E+03	2.369	0.01798	*
ISI4000	1.15E-02	1.31E-02	1.38E+03	0.873	0.38294	
sessionNameMulti-Task:ISI4000	-4.98E-02	1.86E-02	1.38E+03	-2.678	0.00748	**

APPENDIX Q: E2 TRACKING DISTANCE

Table Q.1 Model Comparison for Tracking Distance in E2.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-1618.2	-1595.3	814.12	-1628.2				
Intercept	6	-1660.7	-1633.3	836.37	-1672.7	44.5007	1	2.54E-11	***
Linear	7	-1658.8	-1626.8	836.4	-1672.8	0.0656	1	0.7979	
Quadratic	8	-1656.8	-1620.2	836.41	-1672.8	0.0095	1	0.9224	

Table Q.2 Intercept Model Coefficients for Tracking Distance in E2.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	0.275055	0.010906	66.83378	25.22	< 2e-16	***
$time^1$	0.022824	0.006129	660.01	3.724	2.13E-04	***
$time^2$	0.005139	0.006129	660.01	0.839	4.02E-01	
ISI4000	-0.033953	0.005004	660.01	-6.785	2.59E-11	***

APPENDIX R: E3 TARGET HIT RATE

Table R.1 Model Comparison for Target Hit Rate in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	4665.2	4691.6	-2327.6	4655.2				
Intercept	8	4444.6	4486.8	-2214.3	4428.6	226.56	3	< 2.2e-16	***
Linear	11	4197.2	4255.2	-2087.6	4175.2	253.425	3	< 2.2e-16	***
Quadratic	14	4155.3	4229.1	-2063.7	4127.3	47.908	3	2.23E-10	***

Table R.2 Intercept Model Coefficients for Target Hit Rate in E3.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	1.59E+01	8.26E-02	1.18E+02	192.158	< 2e-16	***
$time^1$	3.98E-03	1.26E-01	1.38E+03	0.032	0.9747	
sessionNameMulti-Task	-1.09E+00	7.25E-02	1.38E+03	-15.016	< 2e-16	***
ISI4000	1.94E-02	7.25E-02	1.38E+03	0.268	7.89E-01	
$time^2$	6.55E-02	1.26E-01	1.38E+03	0.521	0.602301	
$time^1$:sessionNameMulti-Task	-2.60E+00	1.78E-01	1.38E+03	-14.646	< 2e-16	***
$time^1$:ISI4000	-3.79E-02	1.78E-01	1.38E+03	-0.213	0.831296	
sessionNameMulti-Task:ISI4000	6.86E-01	1.03E-01	1.38E+03	6.69	3.23E-11	***
sessionNameMulti-Task:$time^2$	-1.11E+00	1.78E-01	1.38E+03	-6.245	5.63E-10	***
ISI4000:$time^2$	-7.82E-02	1.78E-01	1.38E+03	-0.44	0.659848	
$time^1$:sessionNameMulti-Task:ISI4000	2.12E+00	2.51E-01	1.38E+03	8.43	< 2e-16	***
sessionNameMulti-Task:ISI4000: $time^2$	8.49E-01	2.51E-01	1.38E+03	3.381	0.000743	***

APPENDIX S: E3 OMISSIONS

Table S.1 Model Comparison for Omissions in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	4532.6	4559	-2261.3	4522.6				
Intercept	8	4330.1	4372.3	-2157.1	4314.1	208.467	3	< 2.2e-16	***
Linear	11	4090.3	4148.3	-2034.2	4068.3	245.838	3	< 2.2e-16	***
Quadratic	14	4046.3	4120.1	-2009.1	4018.3	50.035	3	7.85E-11	***

Table S.2 Quadratic Model Coefficients for Omissions in E3.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	1.17E-01	8.12E-02	1.14E+02	1.437	0.153578	
$time^1$	-2.39E-02	1.21E-01	1.38E+03	-0.198	0.843135	
sessionNameCPT+Tracking	1.01E+00	6.97E-02	1.38E+03	14.421	< 2e-16	***
ISI4000	-5.56E-03	6.97E-02	1.38E+03	-0.08	0.93651	
$time^2$	-5.46E-02	1.21E-01	1.38E+03	-0.452	0.65156	
$time^1$:sessionNameCPT+Tracking	2.47E+00	1.71E-01	1.38E+03	14.485	< 2e-16	***
$time^1$:ISI4000	5.18E-02	1.71E-01	1.38E+03	0.303	0.76176	
sessionNameCPT+Tracking:ISI4000	-6.56E-01	9.86E-02	1.38E+03	-6.648	4.28E-11	***
sessionNameCPT+Tracking: $time^2$	1.09E+00	1.71E-01	1.38E+03	6.367	2.62E-10	***
ISI4000: $time^2$	6.91E-02	1.71E-01	1.38E+03	0.405	0.685857	
$time^1$:sessionNameCPT+Tracking:ISI4000	-2.00E+00	2.42E-01	1.38E+03	-8.263	3.29E-16	***
sessionNameCPT+Tracking:ISI4000:$time^2$	-8.20E-01	2.42E-01	1.38E+03	-3.395	0.000705	***

APPENDIX T: E3 COMMISSIONS

Table T.1 Model Comparison for Commissions in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	3500.4	3526.7	-1745.2	3490.4				
Intercept	8	3477.8	3520	-1730.9	3461.8	28.585	3	2.74E-06	***
Linear	11	3465.6	3523.6	-1721.8	3443.6	18.189	3	0.0004021	***
Quadratic	14	3455	3528.8	-1713.5	3427	16.623	3	0.0008446	***

Table T.2 Quadratic Model Coefficients for Commissions in E3.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	7.81E-01	6.91E-02	1.07E+02	11.302	< 2e-16	***
$time^1$	-2.29E-01	9.81E-02	1.38E+03	-2.335	0.019681	*
sessionNameMulti-Task	1.03E-01	5.66E-02	1.38E+03	1.815	0.069811	.
ISI4000	5.56E-03	5.66E-02	1.38E+03	0.098	0.92188	
$time^2$	1.27E-02	9.81E-02	1.38E+03	0.13	0.896781	
$time^1$:sessionNameMulti-Task	5.64E-01	1.39E-01	1.38E+03	4.063	5.11E-05	***
$time^1$:ISI4000	3.51E-01	1.39E-01	1.38E+03	2.527	0.011615	*
sessionNameMulti-Task:ISI4000	1.61E-01	8.01E-02	1.38E+03	2.011	0.044487	*
sessionNameMulti-Task:$time^2$	4.73E-01	1.39E-01	1.38E+03	3.408	0.000674	***
ISI4000:$time^2$	1.82E-03	1.39E-01	1.38E+03	0.013	0.989544	
$time^1$:sessionNameMulti-Task:ISI4000	-4.62E-01	1.96E-01	1.38E+03	-2.355	0.018642	*
sessionNameMulti-Task:ISI4000:$time^2$	-1.87E-01	1.96E-01	1.38E+03	-0.955	0.339943	

APPENDIX U: E3 CORRECT REJECTIONS

Table U.1 Model Comparison for Correct Rejections in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	3535.5	3561.9	-1762.8	3525.5				
Intercept	8	3508.8	3551	-1746.4	3492.8	32.757	3	3.63E-07	***
Linear	11	3493.5	3551.5	-1735.8	3471.5	21.279	3	9.21E-05	***
Quadratic	14	3481.7	3555.5	-1726.8	3453.7	17.788	3	0.0004864	***

Table U.2 Quadratic Model Coefficients for Correct Rejections in E3.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	3.22E+00	7.03E-02	1.06E+02	45.754	< 2e-16	***
$time^1$	2.31E-01	9.90E-02	1.38E+03	2.335	0.019696	*
sessionNameMulti-Task	-1.36E-01	5.71E-02	1.38E+03	-2.382	0.017352	*
ISI4000	-5.56E-03	5.71E-02	1.38E+03	-0.097	0.922561	
$time^2$	-5.46E-03	9.90E-02	1.38E+03	-0.055	0.956049	
$time^1$:sessionNameMulti-Task	-6.20E-01	1.40E-01	1.38E+03	-4.426	1.03E-05	***
$time^1$:ISI4000	-3.55E-01	1.40E-01	1.38E+03	-2.533	0.011407	*
sessionNameMulti-Task:ISI4000	-1.47E-01	8.08E-02	1.38E+03	-1.822	0.068693	.
sessionNameMulti-Task:$time^2$	-4.95E-01	1.40E-01	1.38E+03	-3.534	0.000423	***
ISI4000:$time^2$	-1.82E-03	1.40E-01	1.38E+03	-0.013	0.989636	
$time^1$:sessionNameMulti-Task:ISI4000	4.88E-01	1.98E-01	1.38E+03	2.466	0.013797	*
sessionNameMulti-Task:ISI4000:$time^2$	2.00E-01	1.98E-01	1.38E+03	1.011	0.312398	

APPENDIX V: E3 REACTION TIME

Table V.1 Model Comparison for Reaction Time in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*
Base	-777.63	-751.27	393.81	-787.63				
Intercept	-2061.45	-2019.27	1038.73	-2077.45	1289.825	3	< 2.2e-16	***
Linear	-2129.24	-2071.24	1075.62	-2151.24	73.788	3	6.59E-16	***
Quadratic	-2154.84	-2081.03	1091.42	-2182.84	31.598	3	6.36E-07	***

Table V.2 Quadratic Model Coefficients for Reaction Time in E3.

Fixed Effects	Estimate	Std. Error	df	*t* value	*p*	
(Intercept)	5.88E+00	1.43E-02	7.65E+01	410.89	< 2e-16	***
$time^1$	2.76E-02	1.37E-02	1.38E+03	2.01	4.47E-02	*
sessionNameMulti-Task	2.30E-01	7.92E-03	1.38E+03	29.018	< 2e-16	***
ISI4000	1.64E-01	7.92E-03	1.38E+03	20.72	< 2e-16	***
$time^2$	-3.16E-02	1.37E-02	1.38E+03	-2.308	2.12E-02	*
$time^1$:sessionNameMulti-Task	1.03E-01	1.94E-02	1.38E+03	5.284	1.47E-07	***
$time^1$:ISI4000	4.35E-02	1.94E-02	1.38E+03	2.244	0.02501	*
sessionNameMulti-Task:ISI4000	-1.98E-02	1.12E-02	1.38E+03	-1.769	7.71E-02	.
sessionNameMulti-Task:$time^2$	1.07E-01	1.94E-02	1.38E+03	5.502	4.47E-08	***
ISI4000: $time^2$	3.49E-02	1.94E-02	1.38E+03	1.8	0.07211	.
$time^1$:sessionNameMulti-Task:ISI4000	1.40E-02	2.74E-02	1.38E+03	0.51	6.10E-01	
sessionNameMulti-Task:ISI4000: $time^2$	-8.34E-02	2.74E-02	1.38E+03	-3.04	0.00241	**

APPENDIX W: E3 SENSITIVITY

Table W.1 Model Comparison for Sensitivity in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	2888	2914.3	-1439	2878				
Intercept	8	2775.1	2817.2	-1379.5	2759.1	118.904	3	< 2.2e-16	***
Linear	11	2649.4	2707.4	-1313.7	2627.4	131.644	3	< 2.2e-16	***
Quadratic	14	2605.9	2679.7	-1288.9	2577.9	49.558	3	9.92E-11	***

Table W.2 Quadratic Model Coefficients for Sensitivity in E3.

Fixed Effects	Estimate	Std. Error	df	t value	p	
(Intercept)	2.30405	0.05972	91.39759	38.583	2.00E-16	***
$time^1$	0.12162	0.07483	1960	1.625	0.10428	
sessionNameMulti-Task	-0.15443	0.04321	1960	-3.574	0.00036	***
ISI4000	0.02291	0.04321	1960	0.53	0.59602	
$time^2$	0.09538	0.04321	1960	2.208	0.02739	*
$time^1$:sessionNameMulti-Task	-0.12735	0.07483	1960	-1.702	0.08897	.
$time^1$:ISI4000	0.04608	0.10583	1960	0.435	0.6633	
sessionNameMulti-Task:ISI4000	-0.10078	0.10583	1960	-0.952	0.34106	
sessionNameMulti-Task: $time^2$	-0.15912	0.10583	1960	-1.504	0.13287	
ISI4000: $time^2$	-0.03157	0.0611	1960	-0.517	0.6054	
$time^1$:sessionNameMulti-Task:ISI4000	-0.07442	0.0611	1960	-1.218	0.22339	
sessionNameMulti-Task:ISI4000: $time^2$	-0.01859	0.10583	1960	-0.176	0.86057	

APPENDIX X: E3 RESPONSE BIAS

Table X.1 Model Comparison for Response Bias in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-75.978	-49.616	42.989	-85.98				
Intercept	8	-206.176	-163.997	111.088	-222.18	136.1978	3	<2e-16	***
Linear	11	-307.694	-249.698	164.847	-329.69	107.5178	3	<2e-16	***
Quadratic	14	-305.707	-231.894	166.854	-333.71	4.0135	3	0.26	

Table X.2 Intercept Model Coefficients for Response Bias in E3.

Fixed Effects	Estimate	Std. Error	df	*t* value	*p*	
(Intercept)	93.2741	0.47539	139.52625	196.204	< 2e-16	***
$time^1$	0.44729	0.79352	1380	0.564	5.73E-01	
sessionNameMulti-Task	-3.15398	0.45814	1380	-6.884	8.79E-12	***
ISI4000	2.72137	0.45814	1380	5.94	3.60E-09	***
$time^2$	-0.11326	0.79352	1380	-0.143	8.87E-01	
$time^1$:sessionNameMulti-Task	-14.58094	1.1222	1380	-12.993	< 2e-16	***
$time^1$:ISI4000	0.47924	1.1222	1380	0.427	0.6694	
sessionNameMulti-Task:ISI4000	3.89764	0.6479	1380	6.016	2.29E-09	***
sessionNameMulti-Task: $time^2$	-5.21421	1.1222	1380	-4.646	3.70E-06	***
ISI4000: $time^2$	0.08307	1.1222	1380	0.074	0.941	
$time^1$:sessionNameMulti-Task:ISI4000	13.30148	1.58703	1380	8.381	< 2e-16	***
sessionNameMulti-Task:ISI4000: $time^2$	3.95108	1.58703	1380	2.49	0.0129	*

APPENDIX Y: E3 TRACKING DISTANCE

Table Y.1 Model Comparison for Tracking Distance in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	5	-2088.6	-2065.7	1049.3	-2098.6				
Intercept	6	-2142.6	-2115.2	1077.3	-2154.6	56.0604	1	7.03E-14	***
Linear	7	-2142.7	-2110.7	1078.4	-2156.7	2.0955	1	0.1477	
Quadratic	8	-2140.7	-2104.1	1078.4	-2156.7	0.0023	1	0.9619	

Table Y.2 Intercept Model Coefficients for Tracking Distance in E3.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	0.286496	0.006732	69.753844	42.557	< 2e-16	***
time1	0.013508	0.004449	660	3.036	2.49E-03	**
time2	-0.004571	0.004449	660	-1.027	3.05E-01	
ISI4000	-0.027789	0.003633	660	-7.649	7.18E-14	***

APPENDIX Z: E3 TRACKING DISTANCE DURING VERBAL TASK

Table Z.1 Model Comparison for Tracking Distance during Verbal Task in E3.

Model	nPar	AIC	BIC	Loglik	Deviance	X^2	*df*	*p*	
Base	7	-4251.8	-4211.2	2132.9	-4265.8				
Intercept	10	-4253.5	-4195.5	2136.8	-4273.5	7.7451	3	0.05158	.
Linear	13	-4250.7	-4175.2	2138.3	-4276.7	3.1173	3	0.37389	
Quadratic	16	-4253.1	-4160.2	2142.5	-4285.1	8.3954	3	0.03851	*
Cubic	19	-4253.5	-4143.2	2145.8	-4291.5	6.4488	3	0.09171	.
Quartic	22	-4247.6	-4119.9	2145.8	-4291.6	0.0964	3	0.99226	

Table Z.2 Quadratic Model Coefficients for Tracking Distance during Verbal Task in E3.

Fixed Effects	Estimate	Std. Error	*df*	*t* value	*p*	
(Intercept)	2.73E-01	8.85E-03	2.12E+02	30.803	< 2e-16	***
$time^1$	-1.93E-02	2.71E-02	2.42E+03	-0.711	0.47721	
ISI4000	-5.03E-03	7.32E-03	2.39E+03	-0.688	0.49174	
segtypeRespond	1.38E-02	8.20E-03	2.39E+03	1.686	0.09193	.
$time^2$	-1.64E-02	2.01E-02	2.39E+03	-0.818	0.41328	
$time^3$	1.20E-02	1.01E-02	2.40E+03	1.179	0.2384	
$time^3$	3.07E-02	1.01E-02	2.39E+03	3.057	0.00226	**
$time^1$:ISI4000	1.77E-02	3.21E-02	2.42E+03	0.553	0.57999	
$time^1$:segtypeRespond	-1.48E-02	3.48E-02	2.39E+03	-0.425	0.67076	
ISI4000:segtypeRespond	-9.14E-03	9.38E-03	2.39E+03	-0.975	0.32981	
ISI4000: $time^2$	3.66E-02	2.59E-02	2.39E+03	1.414	0.15746	
segtypeRespond: $time^2$	-3.53E-02	2.77E-02	2.39E+03	-1.275	0.2023	
$time^1$:ISI4000:segtypeRespond	2.88E-02	4.07E-02	2.39E+03	0.709	0.47861	
ISI4000:segtypeRespond:$time^2$	1.00E-02	3.63E-02	2.39E+03	0.277	0.78196	