THE NATURE AND MEASUREMENT OF SUSTAINING ATTENTION OVER TIME: THE INFLUENCE OF COGNITIVE ABILITY, INTERNAL DISTRACTION, AROUSAL, AND MOTIVATION ON SUSTAINED ATTENTION

A Dissertation
Presented to
The Academic Faculty

by

Jason S. Tsukahara

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Psychology
College of Sciences

Georgia Institute of Technology
December 2022

THE NATURE AND MEASUREMENT OF SUSTAINING ATTENTION OVER TIME: THE INFLUENCE OF COGNITIVE ABILITY, INTERNAL DISTRACTION, AROUSAL, AND MOTIVATION ON SUSTAINED ATTENTION

Approved by:

Dr. Randall Engle, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Mark Wheeler
School of Psychology
Georgia Institute of Technology

Dr. Eric Schumacher
School of Psychology
Georgia Institute of Technology

Dr. Michael Beran
Department of Psychology
Georgia State University

Dr. Dobromir Rahnev
School of Psychology
Georgia Institute of Technology

Date Approved: November 28, 2022

ACKNOWLEDGEMENTS

I would like to thank my family and friends, who have provided support and encouragement as I earned my doctorate degree. Especially my mother, who has shown immense patience and understanding as I pursue my dreams—and at my own pace. I cannot express the gratitude I have towards Randy Engle, my advisor, for truly believing in me and providing the opportunities for me to develop into a competent scientist. I have had the fortune of working with incredible lab mates and peers over the years—which can make all the difference for one's mental well-being and emotional balance during the challenges that a doctorate program brings. For that I want to thank Cody Mashburn, Chris Draheim, Jessie Martin, Alex Burgoyne, Tyler Harrison, and Kenny Hicks for making my time at Georgia Tech unforgettable. Similarly, I would like to thank my other peers in the program, Sophie Kay, Ben Perrodin, Ben Jones, Gina Bufton, Gabe Plummer, Zack Tidler, Reba Koenan, Corey Tatel, Sibley Lyndgaard, Paulina Maxim, Julie Harrison, and Maugan Lloyd—all of whom I have immense respect for. I would not have been able to complete my dissertation project without the hard and tireless work from Jess Campbell, our lab manager, to manage a team of research assistants and overseeing data collection for my dissertation study. Finally, I would like to thank my dissertation committee—Eric Shumacher, Doby Rahnev, Mark Wheeler, and Mike Beran—for taking their time to serve on my committee, evaluate my dissertation work, and provide valuable feedback. The achievement of this milestone would not have been possible without the gratitude of those mentioned here, as well as others not mentioned—and for that, with much love, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

SACT	Sustained attention-to-cue task
SACT-v1	Sustained attention-to-cue task – version 1
SACT-v2	Sustained attention-to-cue task – version 2
AC	Attention control
WMC	Working memory capacity
Gf	Fluid intelligence
StroopDL	Stroop with an adaptive response deadline
VAorient-S	Selective visual arrays task with orientation judgement
SymSpan	Symmetry span
RotSpan	Rotation span
RAPM	Raven's advanced progressive matrices
SD	Standard deviation
ms	Milliseconds
px	Pixels
cm	Centimeters
Hz	Hertz
χ^2	Chi-square statistic
CFI	Comparative fit index
RMSEA	Root mean square error of approximation
SRMR	Standardized root mean square residual
η_p^2	Partial eta-squared effect size estimate

SUMMARY

It is evident that it takes a great deal of effort to sustain our attention on any one thing over a period of minutes or even seconds. This ability to sustain attention is critical for many everyday tasks and is often seen as a fundamental factor underlying differences in cognitive ability. Therefore, it is important to understand the factors that determine how long we can voluntarily sustain our attention. Across two studies I used a novel task, the sustained attention-to-cue task (SACT), to assess sustained attention. The critical element of the task is to sustain attention at a cued location for a variable amount of time (0 – 12 seconds). In Study 1, I investigated how individual differences in cognitive ability are related to sustained attention. I found that those higher on attention control showed less of a decline in performance the longer attention had to be sustained. However, sustained attention performance was not related to working memory capacity or fluid intelligence. In Study 2, I investigated how susceptibility to distraction, changes in arousal, and motivation are related to sustained attention performance on the SACT. Overall, there was a large decline in attention on a shorter timescale based on performance, eye gaze, pupil size, and mind wandering measures. There were no changes in attention at a longer timescale, however there was strong evidence that arousal declined over the course of the task. Reward and motivation lead to improvements in attention overall and motivation led to improvements in sustained attention at a shorter timescale. In general, these findings suggest that attention can fluctuate and wane over a relatively short time scale of around 10 seconds or less and that this is related to individual differences in attention control, distractibility, arousal, and motivation.

CHAPTER 1. GENERAL INTRODUCTION

Anecdotally and empirically, it is evident that it takes a great deal of effort to *voluntarily sustain attention* on any one thing over the course of minutes or even seconds. Our minds tend to drift away from a task at hand to other events in our environment or to internally generated thoughts related to our recent concerns and daydreams. It has been estimated that we spend almost 50% of our waking experience caught up in task-unrelated wandering thoughts (Killingsworth & Gilbert, 2010). This is concerning given that many everyday tasks require us to continuously focus the mind on a particular activity. The ability to voluntarily sustain attention has been one of the defining aspects that is thought to contribute to individual differences in cognitive abilities such as working memory capacity (Engle & Kane, 2004; Unsworth & Robison, 2020) and attention control (Engle, 2018; Unsworth & Miller, 2021). Therefore, it is important to understand **the factors that determine how long we can voluntarily sustain our attention.**

In a general sense, we can think of sustained attention as the *continuous and purposeful placement of our attention to one task or event over other things for a period of time*. The focus of our attention on can vary in complexity, from simple objects in our environment to entire task schemas that contain task rules, stimulus-response mappings, speed-accuracy tradeoffs, etc. At longer time-scales, sustained attention is characterized as a decrement in performance over time (minutes to hours), known as the *vigilance decrement*. At shorter time-scales, sustained attention is characterized as changes in attentional state over time (seconds to minutes), and includes several different constructs under the umbrella of *momentary fluctuations of attention* (e.g., lapses of attention, mind wandering, distraction, task focus). Vigilance decrements are demonstrated as declines in performance over long durations—typically at least 20 minutes—and tend to be more pronounced in unchallenging and monotonous tasks (Kahneman, 1973; Langner & Eickhoff, 2013; Manly et al., 2003; Parasuraman, 1984; Poffenberger, 1928; Robertson & O'Connell, 2010). Momentary fluctuations of attention refer to how attention fluctuates between a more optimal task-focused state and suboptimal task-unfocused states on a moment-to-moment basis. For instance, some researchers categorize momentary fluctuations of attention into on-task vs. off-task thoughts (Christoff et al., 2011, 2016; Seli et al., 2018), or into "in the zone" vs. "out of the zone" (Esterman et al., 2013). Although these two characteristics are often assessed in different ways and time scales, it may be that they represent an underlying continuum of sustained attention. For instance, it could be that early on in task performance the focus of attention is more frequently in a task-focused mode with occasional attentional lapses (failures of sustained attention at a short time scale) but later in task performance the focus of attention wanes to a more task-disengaged mode with attentional lapses being more frequent and of longer duration. Similarly, momentary fluctuations of attention can be seen as mini-vigilance decrements over a short period of time.

Understanding the factors that differ between individuals provides an important perspective on why some individuals are able to sustain their attention longer than others. Therefore, in Study 1, I investigated the role of cognitive abilities in determining how long attention can be sustained. In general, sustained attention is seen as a fundamental dimension of attention control (Cohen, 2014; Engle & Kane, 2004; Mirsky et al., 1991; Stuss et al., 1995; Unsworth & Miller, 2021) to maintain activation of the attended object (e.g. task schema) and inhibit distracting objects (e.g., conflicting schema) that may interfere with performance (Norman & Shallice, 1986; Stuss et al., 1995). As such, the concept of sustained attention has become a central feature to understanding individual differences in cognitive abilities like working memory capacity and attention control (Engle & Kane, 2004; Unsworth & Miller, 2021; Unsworth & Robison, 2020).

Another important question is how factors that change within the individual can explain how long one can sustain attention. Therefore, in Study 2, I will investigate how 1) resource depletion, 2) changes in task utility, 3) susceptibility to internal distraction, and 4) changes in arousal are all related to how long attention can be sustained within the individual. A novel task I recently developed— the sustained attention-to-cue task (Draheim et al., 2021)—will be used to assess sustained attention over time in the two studies.

1.1 The Sustained Attention-to-Cue Task

In the sustained attention-to-cue task (Draheim et al., 2021), *the* critical element is the wait time delay in which attention must be sustained at a cued location for a variable amount of time. At the end of the wait time delay, a to-be-identified target is briefly presented among non-targets. The purpose of the brief target presentation is that, if attention has shifted from the cued location, then the target location will not be in the focus of attention and thereby not accurately identified due to a failure to sustain attention. In an original version of the sustained attention-to-cue task (SACT-v1), the circle cue remained on the screen for the entire wait time and a distractor was presented briefly before the target (Figure 1).

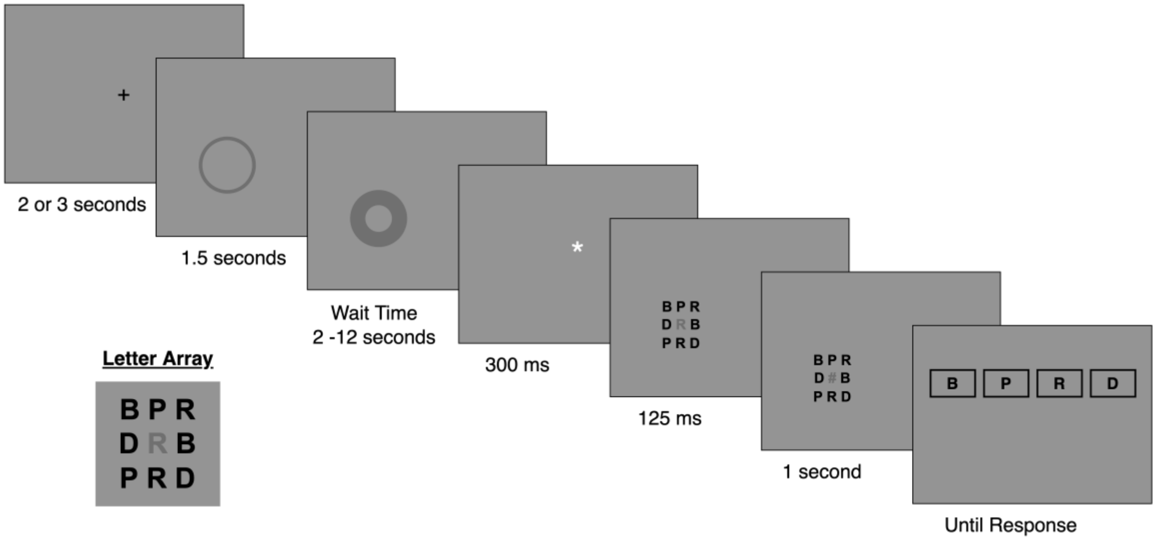

Figure 1 - Trial Procedure for SACT-v1

The broader motivation underlying the development of the sustained attention-to-cue task was to develop new measures of attention control that did not rely on reaction time nor reaction time difference scores (Draheim et al., 2019, 2021). The reliance on traditional measures that use reaction time difference scores has hindered theoretical and practical importance of individual differences in attention control (Draheim et al., 2019). An exploratory study was conducted to modify and develop existing new attention control tasks that 1) do not rely on a difference score and 2) make reaction time irrelevant to performance (not just the scoring) of the task and rely on accuracy as the dependent measure (Draheim et al., 2021). A number of new and modified tasks were developed and tested as measures of attention control (Draheim et al., 2021). The sustained attention-to-cue task was one of those, and aggregating across several metrics, it was the third best attention control task in the test battery (see Table 7 in Draheim et al., 2021).

The more specific motivation behind the development of the sustained attention-to-cue task was to develop an accuracy analog of the psychomotor vigilance task. The psychomotor vigilance task is a prominent paradigm for measuring sustained attention, and in particular for studying the effects of fatigue and sleep deprivation (Dinges & Powell, 1985; Doran et al., 2001). Some individual differences researchers have largely relied on this task as one indicator of a latent attention control construct (Unsworth & McMillan, 2014). In line with the broader motivation, I decided to develop a purely accuracy analog of the psychomotor vigilance task. The point was not to develop an identical version of the psychomotor vigilance task but to include the critical feature of a variable wait time.

In terms of the psychometrics of the task, the SACT-v1 displayed good reliability and validity as an attention control measure (Draheim et al., 2021). The SACT-v1 correlated well with the other attention control tasks in the battery (on average; r = .32), particularly the antisaccade (r = .40)—our gold standard attention control task. The SACT-v1 also cohered well as an indicator on a latent attention control construct. On a three-indicator factor with the three best attention control tasks (antisaccade, visual arrays, and SACT-v1), the SACT-v1 had a loading of .61 in comparison to .74 (antisaccade) and .63 (visual arrays). This three-indicator factor provided higher and more balanced factor loadings on attention control compared to previous research (Friedman & Miyake, 2004; Rey-Mermet et al., 2018; Shipstead et al., 2014). Modified versions of the SACT-v1 will be used in Study 1 to improve on the validity of the task as a measure of sustained attention and in Study 2 to investigate the influence of resource depletion, changes in task utility, susceptibility to internal distraction, and changes in arousal on sustained attention.

CHAPTER 2. STUDY 1

2.1 Introduction

Sustained attention is an important aspect of what is referred to more broadly as the control of attention, an ability that is an important determinant of individual differences in general cognitive abilities, such as working memory capacity and fluid intelligence (Burgoyne et al., 2022). More so, working memory concepts such as maintenance of information and distractor interference are closely related to the concept of sustained attention. For instance, Engle et al. (1999) state "we assume that "working memory capacity" is not really about storage or memory per se, but about the *capacity for controlled,* ***sustained attention*** [emphasis added] *in the face of interference or distraction*… working memory capacity reflects the ability to apply activation to memory representations, to either bring them into focus or **maintain them in focus** [emphasis added], particularly in the face of interference or distraction." That is, individual differences in working memory capacity are seen as differences in the ability to sustain or maintain the focus of attention. And sustaining attention implies resistance to distraction—such that a failure to prevent distraction *is* a failure to sustain attention.

Another aspect of sustained attention that is related to working memory capacity is "maintaining sufficient access to the current task goals so that they, rather than habit, control responding." (McVay & Kane, 2009)—that is, active goal maintenance. This aspect of sustained attention is more about preventing lapses of attention to a task and is often reflected in slow responding, large reaction time variability, and frequent mind wandering. In fact, research has shown that individuals with higher working memory capacity show less extreme reaction times, less reaction time variability, and less mind wandering (McVay & Kane, 2009, 2012; Robison & Unsworth, 2018; Schmiedek et al., 2007). Whether active goal maintenance and resistance to distraction represent the same underlying sustained attention ability is not well studied. However, some work has shown that mind wandering and efficiency of filtering out distractors are uncorrelated, and each explain unique variance in a working memory task (Unsworth & Robison, 2016).

Unsworth and colleagues have characterized differences in working memory capacity as being largely driven by fluctuations of attention (Unsworth & Robison, 2017, 2020). Specifically, they propose that low working memory capacity individuals display more fluctuations of attention due to a dysregulation of the locus coeruleus-norepinephrine system (Unsworth & Robison, 2017). The locus coeruleus-norepinephrine system is central neuromodulatory system in the brain associated with modulating arousal levels and attentional filtering (Aston-Jones & Cohen, 2005; Sara, 2009). Similarly, I have proposed that fluid intelligence is related to the locus coeruleus-norepinephrine system at essentially all levels of brain function, and in particular regulating the organization of functional brain networks related to fluid intelligence and the control of attention (Tsukahara & Engle, 2021).

More recently, Unsworth has made a distinction between the intensity and consistency of attention as two sources of variation in attention control abilities (Unsworth & Miller, 2021). Intensity refers to how much attention is allocated to a given task or stimulus at a given moment in time, whereas consistency refers to how consistently attention is allocated over time. The use of pupillometry may be a particularly useful way of capturing variation in both the intensity and consistency of attention. The amount of pupil dilation has been shown to be related to attentional effort and thereby may be an indicator of the intensity of attention—a larger pupil dilation suggests greater allocation of attention to a stimulus (Unsworth & Robison, 2015, 2016). The variability in pupil size has been shown to be an indicator of the consistency of attention—greater pupil variability suggests less consistency of attention to a task over time (Robison & Unsworth, 2019; Unsworth & Robison, 2015, 2018). In fact, these pupillary indicators of the intensity and consistency of attention are related to individual differences in working memory capacity and attention control (Robison & Unsworth, 2019; Unsworth & Robison, 2015, 2016, 2018).

Although researchers have associated sustained attention, fluctuations of attention, attentional lapses, and mind wandering with individual differences in working memory capacity, this relationship is better characterized as the control of attention and not working memory per se (Engle, 2018). There is little research on associating performance in sustained attention tasks with different cognitive abilities at the latent construct level; however, some research has shown that individual sustained attention tasks—such as the psychomotor vigilance task and the sustained attention to response task—do correlate well with other attention control tasks and load well onto a latent and broad attention control factor, and this latent factor does correlate strongly with working memory capacity (McVay & Kane, 2009; Robison & Unsworth, 2018; Unsworth & McMillan, 2014).

Even though tasks like the psychomotor vigilance task (Dinges & Powell, 1985) and the sustained attention to response task (Robertson et al., 1997) are commonly used as sustained attention or vigilance tasks, they are not without concerns as to what is actually being measured. For instance, the primary feature of the sustained attention to response task is to withhold a habitual and prepotent response that is built up over the duration of the task—getting at the active goal maintenance component by McVay and Kane (2009) that I highlighted earlier. However, some have suggested that it is not clear how much the sustained attention to response task measures sustained attention vs. speed-accuracy tradeoffs and response inhibition (Seli, Jonker, Cheyne, et al., 2013; Seli, Jonker, Solman, et al., 2013).

In the psychomotor vigilance task, like the SACT-v1, there is a variable wait time delay (2 – 10 seconds) with a stimulus on the display (a row of zero's) and a response needs to be made as quickly as possible once the stimulus changes (the zero's start counting up in milliseconds). This task demonstrates large vigilance decrements such that reaction times are longer and more extreme after at least 20 minutes of performance on the task and after sleep deprivation (Doran et al., 2001; Dorrian et al., 2004). These classic findings have validated the psychomotor vigilance task as a measure of sustained attention and vigilance.

However, there is another common—yet less frequently reported—finding in this task (Matthews et al., 2017); the longest and most extreme reaction times occur on the shortest wait times (2 – 3 seconds) and the fastest reaction times occur on the more moderate to longer wait times (5 – 10 seconds). If sustained attention, or vigilance, over the variable wait time was placing demands on sustained attention then it would be expected that the longest and most extreme reaction times would occur at the longest wait times. But the exact opposite is found.

This is particularly concerning in individual differences research when a common dependent measure of “attentional lapses” and extreme reaction times (20% slowest) are used to correlate with other cognitive ability measures (Matthews et al., 2017; Unsworth & Robison, 2016). However, it could be that longer and more extreme reaction times still reflect vigilance decrements over the long term. In fact, if this was the case it would be expected that the waning of sustained attention over the course of the task would be reflected in the shorter wait times because there is less time to recover and orient attention back to the task at the start of the trial compared to longer wait times (e.g., 10 seconds). In fact, Unsworth et al. (Unsworth & Robison, 2020) demonstrated that working memory capacity differences on the psychomotor vigilance task are larger at the shortest wait times. These findings suggest that, although the psychomotor vigilance task is capturing differences in sustained attention over a longer time scale (the vigilance decrement) it is not able to reflect differences in sustained attention over a shorter time scale (momentary fluctuations of attention).

One question, then, is how does the sustained attention-to-cue task compare to the psychomotor vigilance task in capturing differences in sustained attention at both long and short time-scales? If sustained attention over the wait time was driving performance differences on the sustained attention-to-cue task, then it would be expected that performance should get worse with longer wait times. In a re-analysis of the SACT-v1 task from Draheim et al. (2021), I found the opposite (Figure 2); accuracy was the worst at the shortest wait time (2 seconds), $F(3, 4210) = 19.39$, $p < .05$. There was also a vigilance decrement ($F(2, 14.470) = 70.157$, $p < .05$) and a vigilance decrement × wait time interaction ($F(6, 4211) = 72.93$, $p < .05$) such that the shortest wait time (2 seconds) showed the largest decrease in performance with time-on-task. Although consistent with findings from the psychomotor vigilance task, the finding that performance and the vigilance decrement was worst at the shortest wait time in the SACT-v1 questions whether we can truly consider this task to be measuring variability in sustained attention ability over the wait time and not some other aspect of attention control.

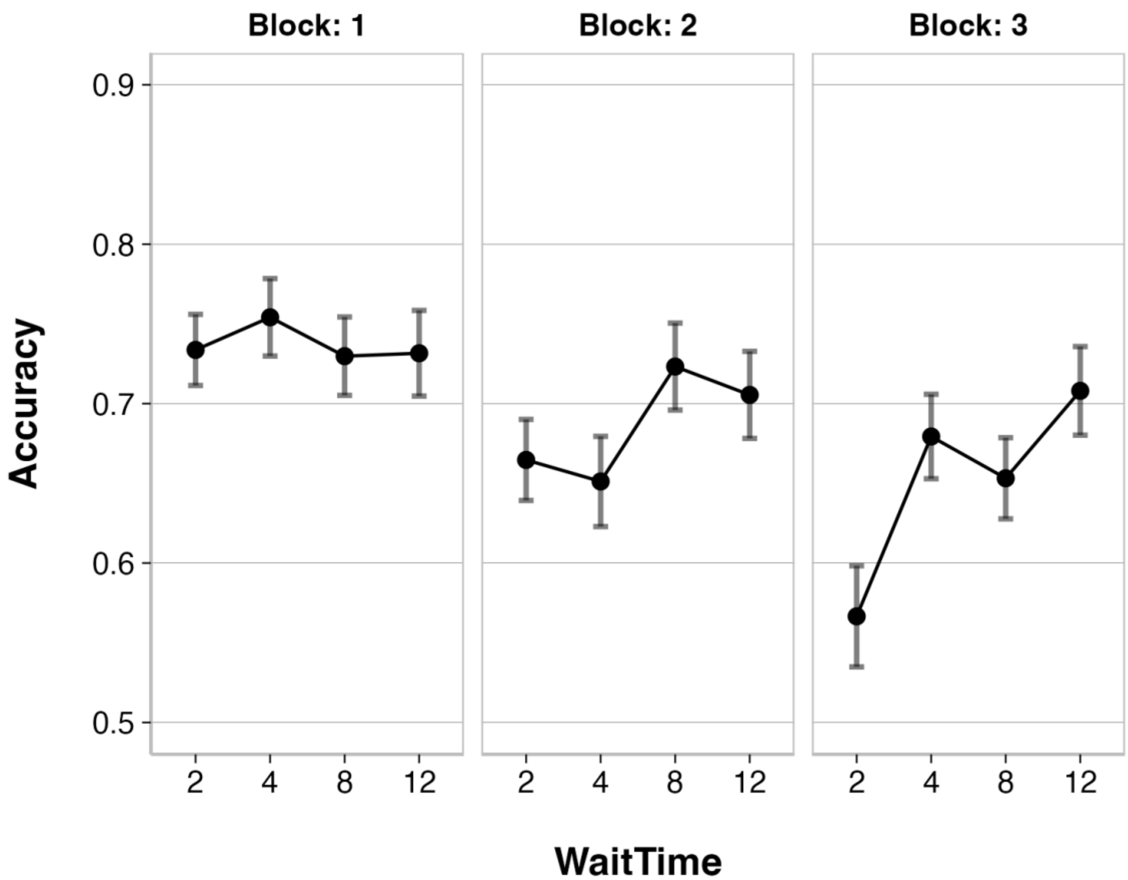

Figure 2 - Wait Time × Block Interaction on Performance in the SACT-v1

Note. Estimated marginal means are plotted. Error bars represent 95% confidence intervals.

2.1.1 The Sustained Attention-to-Cue Task – Version 2

What are some possible reasons why the SACT-v1 failed to capture sustained attention across the wait time? One possibility is that the circle cue remaining on the display over the wait time produced less of a demand on sustained attention—such that if there was a failure of sustained attention over the wait time, attention to the cued location can be restored by redirecting the gaze back to the circle cue. There were also only four possible wait times, perhaps leading to some individuals being able to predict the onset of the target. Additionally, even though the interval between the onset of the distractor and the cued target location was designed to minimize the distractor as a cue when to focus attention at the cued location, there is a possibility that some found a way to utilize this strategy anyways. If so, this would capture variability due more to the orientation of attention rather than sustained attention. Therefore, to address these potential issues, in Study 1 I developed a second version of the sustained attention-to-cue task to hopefully improve on the validity of this task as a measure of sustained attention.

In a second version of the sustained attention-to-cue task (SACT-v2), the circle cue no longer remained on the display during the wait time (Figure 3). This change was made to prevent the subject from shifting attention away from the cued location and using the circle to cue attention back to the cued location. Additionally, the distractor was removed from the task as it created a strategic possibility to be used as a cue for when the target was about to occur. The display of the letter array containing the target letter was also modified to make it less predictable what the expected array of stimuli would look like to hopefully create more interference between target and non-target letters. The easier the target is to identify the less demand there is to keep attention sustained at the cued location. Finally, a more continuous set of wait time delays was used; from 2 to 12 seconds in 500 ms intervals in addition to a 0 second wait time delay as a baseline comparison in which there should be no demand on sustained attention.

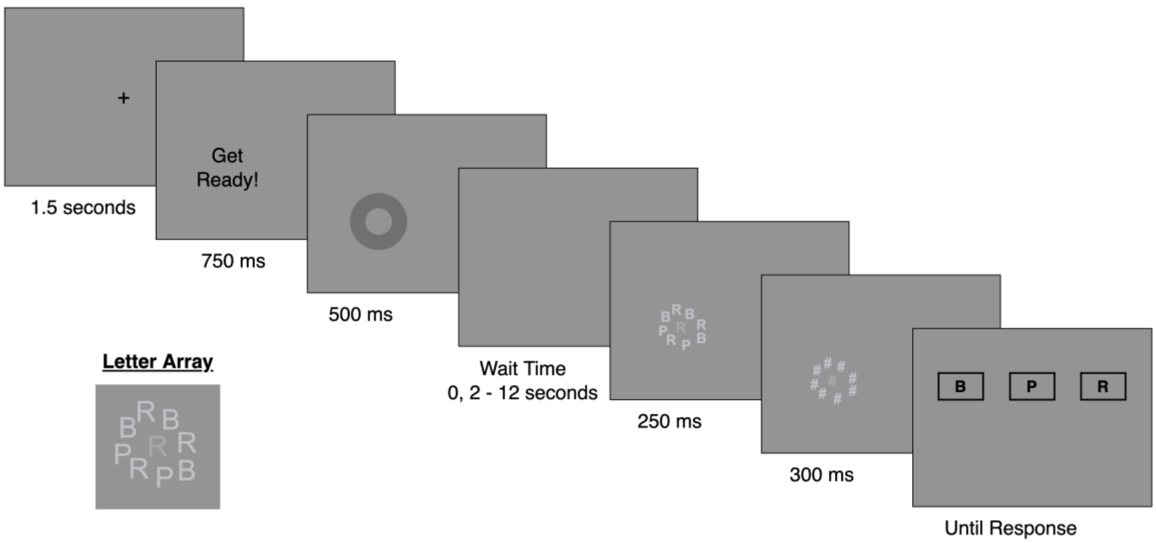

Figure 3 - Trial Procedure for the SACT-v2

Data collection on the SACT-v2 was part of a larger study with a battery of tasks to measure attention control, working memory capacity, fluid intelligence, and multitasking (https://osf.io/qbwem). The primary analyses focused on 1) the relationship of the SACT-v2 with other measures of attention control and 2) the effect of wait time (0 – 12 seconds) on performance and if this effect of wait time is correlated with attention control. If the SACT-v2 is a valid measure of sustained attention, then it would be expected that there should be an effect of wait time on accuracy and that effect will depend on attention control such that those with a higher attention control ability will show a smaller effect of wait time on accuracy. That is, those with a higher attention control ability should be better at sustaining their attention over longer wait times and therefore show a smaller difference in performance between short and long wait times.

2.2 Method

The data analyzed in this study was part of a larger data collection effort that occurred between November 2020 – April 2022. The following link has a summary of the larger data collection procedure and a reference list of all publications to come out of this sample https://osf.io/qbwem.

2.2.1 Subjects

The study was conducted at the Georgia Institute of Technology in Atlanta, GA, USA. The study consisted of five 2.5-hour sessions. All subjects were required to be native English speakers and 18-35 years of age. No screening on vision or hearing was performed. We recruited subjects from Georgia Tech, other surrounding colleges in Atlanta, and the broader Atlanta community. Georgia Tech students enrolled in an undergraduate psychology course were given the option to receive 2.5 hours of course credit or monetary compensation for each session. We compensated subjects with an average of $40 for each session. This study was approved by the Georgia Institute of Technology's Institutional Review Board under Protocol H20165.

A total of 327 subjects completed at least four sessions[1] and a few subjects were completely removed from the data as a result of data cleaning procedure discussed below for a final sample of 324. Monte-Carlo simulations suggest that for stable estimates of correlations sample sizes should approach 250 (Schönbrodt & Perugini, 2013).

2.2.2 Tasks and Procedures

On each session, subjects performed a battery of cognitive tasks in individual testing rooms. A single research assistant administered each task and monitored all subjects. The tasks were performed on either a Windows 10 or Windows 7 computer with an LED-backlit LCD monitor, and subjects wore headphones during all tasks. The tasks were programmed in E-Prime 3.0 and E-Prime 2.0.10 software ([E-Prime 2.0] Psychology Software Tools, 2016; Inc. [E-P. 3. 0] Psychology Software Tools, 2016). Of the constructs relevant to the current study, we administered tasks to measure *attention control*, *working memory capacity*, and *fluid intelligence*.

2.2.2.1 <u>Sustained attention-to-cue task – version 2 (SACT-v2)</u>

In this task subjects needed to sustain their attention at a visually cued location for a variable amount of time and ultimately identify a target letter presented briefly at the center of the cued location. The stimuli were presented against a grey background. Each trial started with a central black fixation for 1 second. After the fixation, a 750 ms interval occurred in which the words "Get Ready!" were displayed at the to-be cued location along

[1] The tasks administered on the fifth session were part of a specific project we were collaborating on with the Navy Research Lab. The main set of tasks, therefore, were from sessions 1-4. Because of this, across all our projects we considered our total sample as those subjects that completed at least session 4.

with an auditory beep. This interval was to prepare and focus the subject for the upcoming trial. A circle cue was then displayed. To orient the subject on the circle cue, the large circle began to immediately shrink in size until it reached a fixed size. The entire duration of the circle cue on the display was approximately 500 ms. Once the cue reached the fixed size, it was removed from the display. The display remained blank over the entire variable wait time. The wait time lasted either 0 seconds or 2 – 12 seconds in 500 ms intervals (e.g., 2, 2.5, 3, 3.5… seconds). After the variable wait time, a cloud array of letters was displayed at the cued location. At the center of the cloud and the cued location was the target letter displayed in "dark gray" font. The surrounding non-target letters were displayed in "silver" font. The target letter was either a B, P, or R and the non-target letters were three of each letter B, P, and R for a total of 9 non-target letters. The array of letters was displayed for 250 ms, after which the target letter was masked for 300 ms with a "#". After the mask, a response screen with B, P, and R response options were displayed and a mouse was used to select which one was the target letter. After a response, there was a blank buffer display presented for 500 ms.

The location of the cued location was semi-randomly determined such that there were an equal number of locations in the top-left, top-right, bottom-left, and bottom-right quadrants of the display. In the array of letters, the target letter always occurred at the center of the cued location. The non-target letters were semi-randomly dispersed within a 96×96 pixel square around the target letter with a minimum stimulus distance of 24 pixels to prevent any overlap.

The task consisted of 6 practice trials in which feedback was provided and a criterion of getting 3 out of the 6 practice trials correct before moving on to the real trials. The task had 3 blocks of 22 trials for a total of 66 trials without feedback. Each wait time delay occurred once per block for a total of 3 trials for each wait time across the entire task. There was a self-timed break given after the first and second block of trials. The task was scored based on accuracy as the proportion of correct responses.

2.2.2.2 Attention control tasks

We measured attention control with the antisaccade, a selective visual arrays, and a Stroop task with an adaptive response deadline. See Draheim et al. (2021) and Martin et al. (2021) for the reliability and validity of the attention control measures.

Antisaccade (Hallett, 1978; Hutchison, 2007). Participants identified a “Q” or “O” that appeared briefly on the opposite side of the screen as a distractor stimulus. After a central fixation cross appeared for 1000 ms or 2000 ms, an asterisk (*) flashed at 12.3° visual angle to the left or right of the central fixation for 100 ms. Afterward, the letter “Q” or “O” was presented on the opposite side at 12.3° visual angle of the central fixation for 100 ms, immediately followed by a visual mask (##). Participants indicated whether the letter was a “Q” or an “O”. They completed 16 slow practice trials during which letter duration was set to 750 ms, followed by 72 test trials. The task was scored based on accuracy as the proportion of correct responses.

Selective Visual Arrays with orientation judgement – VAorient-S (Draheim et al., 2021; Luck & Vogel, 1997; Shipstead et al., 2014). After a central fixation of 1000 ms, a cue word (“RED” or “BLUE”) appeared instructing the participant to attend to either red or blue rectangles. Next, a target array of red and blue rectangles of different orientations (horizontal, left diagonal, right diagonal, and vertical) was presented for 250 ms, followed by a blank screen for 900 ms. Next, a probe array with only the cued-color rectangles was presented, with one rectangle highlighted by a white dot. The orientation of the highlighted rectangle was either the same as it was in the target array, or different, with equal likelihood. The participant indicated with the keyboard whether the orientation of the highlighted rectangle had changed or stayed the same. The target array contained either 3 or 5 rectangles per color (10 and 14 total). There were 48 trials per array set size. Capacity scores (*k*) for each set size were calculated using the single-probe correction (Cowan et al., 2005; Shipstead et al., 2014): set size * (hit rate + correction rejection rate – 1). The task was scored as the mean *k* estimate across the two set sizes.

Stroop task with an adaptive response deadline – version 2 (StroopDL-v2). This task was a modified version of the adaptive response deadline Stroop task (StroopDL-v1) used in Draheim et al. (2021). The task consisted of a color Stroop task in which the words “RED”, “GREEN”, and “BLUE” were presented one at a time in red, green, or blue font colors. The words were either congruent with the color (e.g., the word “RED” in red font color) or incongruent with the color (e.g., the word “RED” in blue font color). The instructions were to indicate the font color by pressing 1, 2, or 3 on the keyboard number pad for (green, blue, and red, respectively). To assist with response mapping, the keys had a colored piece of paper with the corresponding color taped onto them. There was a 2:1 ratio of congruent to incongruent trials with 96 incongruent trials and a total of 288 trials overall. The task was administered over 4 blocks of 72 trials each with an optional rest break between blocks. Practice trials were administered in different blocks, with 24 response mapping practice trials, 18 standard Stroop no deadline practice trials, and 18 non-adaptive response deadline practice trials.

An adaptive procedure was used to estimate the subject's response deadline threshold at about 75 % accuracy. The adaptive procedure based only on the incongruent trials. On each incongruent trial, if an incorrect response was made *or* the reaction time was longer than the response deadline then the response deadline increased (more time to respond) on the next trial. If a correct response was made *and* the reaction time was shorter than the response deadline, then the response deadline decreased (less time to respond) on the next trial. The response deadline started at a relatively easy level of 1.5 seconds. A 3:1 up-to-down ratio was used for the step sizes such that the step size (change in response deadline) for incorrect/too slow of trials was three times larger than the step size for correct/deadline met trials. The step size started at 240:80 ms, decreased to 120:40 ms after 17 incongruent trials, decreased to 60:20 ms after 33 incongruent trials, decreased to 30:10 ms after 49 incongruent trials, decreased to 15:5 ms after 65 incongruent trials, and finally settled at 9:3 ms after 81 incongruent trials. Feedback was given in the form of an audio tone and the words "TOO SLOW! GO FASTER!" presented in red font when the response deadline was not met.

2.2.2.3 Working memory capacity tasks

We measured working memory capacity with two different sets of working memory tasks. We used two spatial complex-span tasks, the *Advanced Symmetry Span* and *Advanced Rotation Span* tasks. The complex-span tasks consist of alternating memory storage and processing sub-tasks (Conway et al., 2005; Unsworth et al., 2005). The advanced versions of the tasks include larger set-sizes of memory items (Draheim et al., 2018). In all complex-span tasks, the edit-distance scoring method was used to calculate span scores (Gonthier, 2022). The edit-distance scoring method is an alternative and more accurate scoring method to the traditional partial score that requires items to be correctly recalled in their absolute serial position (Conway et al., 2005).

Advanced Symmetry Span. This task required subjects to remember a series of spatial locations in a 4x4 matrix presented in alternation with a pattern of squares which they had to decide whether the pattern was symmetrical on the vertical midline. On each trial, subjects were first presented with a 16 × 16 matrix of black and white squares and were required to decide whether the pattern was symmetric on the vertical midline. Followed by the symmetry judgment, a 4 × 4 matrix of squares with one square highlighted in red were displayed. The location of the red-square was the to-be-remembered spatial location. This alternation continued until a variable set-size of spatial locations had been presented. Then, on the recall screen the same 4 × 4 matrix of squares was presented but with no squares highlighted in red. Subjects had to recall the spatial locations in the correct order by clicking the mouse on the appropriate squares in the matrix. There was a total of 12 trials (2 blocks of 6 trials), set-sizes ranged from 2-7, and each set-size occurred twice (once in each block). Scores on the advanced symmetry span task were calculated using the edit-distance scoring method.

Advanced Rotation Span. This task required subjects to remember a series of directional arrows of varying size in alternation with a mental rotation task in which they had to mentally rotate and decide if a letter was mirror reversed or not. On each trial, subjects first solved a mental rotation problem followed by the presentation of a single arrow with a specific direction (8 possible directions; the four cardinal and four ordinal directions) and specific size (small or large). Both the direction and size of the arrow were the to-be-remembered features. This alternation continued until a variable set-size of arrows had been presented. Then, on the recall screen all possible arrow directions and sizes were presented. Subjects had to recall the direction and size of the arrows in the correct order by clicking the mouse on the appropriate arrow. There was a total of 12 trials (2 blocks of 6 trials), set-sizes ranged from 2-7, and each set-size occurred twice (once in each block). Scores on the advanced rotation span task were calculated using the edit-distance scoring method.

2.2.2.4 Fluid intelligence tasks

Raven's Advanced Progressive Matrices (Raven et al., 1998). In this task subjects were presented with a matrix of figures that follow a logical pattern across rows and columns. For each problem in this task, a 3×3 matrix of 8 abstract figures was presented with the bottom-right element missing. Subjects had to identify the logical pattern and select one of eight answer choices that fits the logical pattern of the matrix. Subjects were given 10 minutes to solve 18 of the odd numbered problems from the full test. Scores on this task were calculated as the total number of problems solved correctly.

Letter Sets (Ekstrom et al., 1976). Subjects were presented with 5 sets of 4-letter sequences (e.g., NOPQ DEFL ABCD HIJK UVWX). Subjects had to identify a common pattern among 4 of the sets and select the set of letters that did not follow the pattern (e.g., the letter sets are all in consecutive alphabetical order except for DEFL). Subjects were given 10 minutes to solve 30 problems. Scores on this task were calculated as the total number of problems solved correctly.

Number Series (Thurstone, 1938). For each problem in this task, a series of numbers were presented that progressed in a particular logical fashion. Subjects had to identify the rule and select the next number, out of 5 answer choices, that should occur next in the series of numbers to be consistent with the logical rule. Subjects were given 5 minutes to complete 15 problems. Scores on this task were calculated as the total number of problems solved correctly.

2.2.3 Data processing

All data processing, cleaning, scoring, and analyses were conducted in R statistical software (R Core Team, 2020). For all of the cognitive tasks, except for the fluid intelligence tasks, problematic subjects were detected as having an overall accuracy equal to or less than chance performance and their scores for that task were set to missing. For the complex-span tasks, overall accuracy was assessed based on the processing task (e.g., symmetry judgments for the symmetry span task). For all cognitive tasks, a two-pass outlier method was used on the task scores. On each pass, z-scores were computed and univariate outliers were identified as having scores +/- 3.5 standard deviations or greater from the mean score on that pass and outlier scores were replaced with missing data. Therefore, missing data was present due to data cleaning and other factors such as a subject not having enough time to complete a task on a given session, and the task program crashing during administration.

2.2.4 Planned Data Analysis

Descriptives and Reliability. Descriptive statistics were calculated for each task to assess the distribution of scores including the mean, standard deviation, skewness, and kurtosis. A split-half reliability method was used to estimate reliability for each task. The tasks were split into even/odd trials and the scores were calculated for each half just as they are for the whole task. Correlations between even and odd scores were corrected with the spearman-brown prophecy formula. Additionally, for tasks in which it was appropriate, Cronbach's alpha was calculated as an estimate of reliability.

Relationship of SACT-v2 with Cognitive Abilities. The correlations between the SACT-v2 and composite factor scores of Attention Control, Working Memory Capacity, and Fluid Intelligence were assessed. Given the findings from Draheim et al. (2021), the hypothesis is that the SACT-v2 will correlate more strongly with Attention Control than the other cognitive ability constructs. A confirmatory factor analysis was conducted to test for 1) the latent factor loading of the SACT-v2 on an Attention Control latent factor and 2) the correlation between the latent constructs. The hypothesis for this analysis is that the SACT-v2 will have a moderate to strong latent factor loading on Attention Control and there will be moderate to strong correlations between the latent factors.

Wait Time Analysis. The most critical analysis is on the effect of wait time in the SACT-v2 and its interaction with cognitive ability. Because there were only a few trials for each wait time, the wait times were grouped to increase power and interpretability of the analyses. Wait times at 0 seconds were considered as "none", 2 – 5 seconds were considered as "short", 5.5 – 8 seconds as "medium", and 8.5 – 12 seconds as "long" wait times. The "none" group only had 3 trials while the other groups had 7 trials each. In addition, the effect of time-on-task was also tested by breaking the task into 3 blocks of trials. A 4 (Wait Time) ×3 (Time-on-Task) within-subjects ANOVA was conducted to test for the effects of wait time and time-on-task on SACT-v2 performance.

If the modifications of the second version of the sustained attention-to-cue task do indeed make this version a more valid measure of sustained attention compared to the first version, then it is expected that there will be an effect of wait time on performance in the SACT-v2—such that performance will be worse at longer wait times than shorter wait times. A within-subjects ANOVA was conducted to test for the effect of wait time on accuracy in the SACT-v2. In addition, it is expected that those who score higher on the SACT-v2 will show less of a decrease in performance as wait time gets longer. To test for this interaction, a within-subjects ANCOVA was conducted with SACT-v2 score as a covariate. Another critical analysis is the interaction between wait time and cognitive ability. If sustained attention is an important aspect of individual differences in attention control, then it is expected that there will be a significant interaction between wait time on the SACT-v2 and attention control. To test this, a within-subjects ANCOVA was conducted with attention control as a covariate. The general hypothesis is that those with higher attention control ability will show less of a change in accuracy due to wait time compared to those with lower attention control ability. A similar model was conducted for each cognitive ability with no specific predictions as to whether there will be a significant interaction or not.

2.3 Results

Descriptive and reliability statistics are presented in Table 1. In general, all tasks had acceptable levels of skew, kurtosis, and reliability. All tasks had less than 6% of missing data. The SACT-v2 had a slight negative skew (-1.11), though it had good internal consistency (α = .87).

2.3.1 Correlations with cognitive abilities

The SACT-v2 correlated well with other measures of attention control, on average $r = .28$. The full correlation matrix for all tasks is shown in Table 2. To assess the correlations that SACT-v2 had with attention control, working memory capacity, and fluid intelligence a composite of each construct was created using exploratory factor analysis specifying a one factor solution with only the respective tasks for the construct included. The SACT-v2 correlated strongly with attention control ($r = .35$, $t(275) = 6.28$, $p < .001$), had a small but significant correlation with working memory capacity ($r = .15$, $t(287) = 2.49$, $p = .027$), and a small and non-significant correlation with fluid intelligence ($r = .11$, $t(290) = 1.85$, $p = .065$).

Table 1 - Descriptives and Reliabilities for Study 1

Task	Mean (SD)	Range	Skewness	Kurtosis	Reliability	% Missing
SACT-v2	.89 (.10)	.53 – 1.00	-1.11	.71	.87	0
Antisaccade	.81 (.12)	.51 – 1.00	-.60	-.68	.87	5.54
VAorient-S	2.50 (0.66)	.67 – 3.78	-.34	-.41	.81[a]	1.63
StroopDL	1009.56 (475.42)	433 – 2715	1.72	2.95	.87	4.23
SymSpan	35.01 (8.76)	7 – 53	-.46	-.15	.74	2.61
RotSpan	29.59 (8.49)	3 – 48	-.39	.03	.76	3.58
RAPM	11.36 (2.82)	3 – 18	-.40	-.24	.77	.65
NumSeries	10.01 (2.90)	2 – 15	-.19	-.73	.73	1.30
LetterSets	16.42 (4.38)	5 – 26	-.17	-.69	.85	2.93

[a] Spearman-brown corrected split-half reliability. All other reliabilities were computed as Cronbach's alpha

Table 2 - Correlation Matrix for Study 1

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. SACT									
2. Antisaccade	**0.30**								
3. VAorient_S	**0.32**	**0.39**							
4. StroopDL	**-0.21**	**-0.32**	**-0.24**						
5. SymSpan	**0.12**	**0.24**	**0.42**	**-0.16**					
6. RotSpan	**0.12**	**0.25**	**0.31**	**-0.20**	**0.55**				
7. RAPM	0.11	**0.30**	**0.40**	-0.11	**0.34**	**0.27**			
8. LetterSets	0.09	**0.26**	**0.28**	**-0.19**	**0.30**	**0.17**	**0.38**		
9. NumSeries	0.06	**0.27**	**0.40**	**-0.28**	**0.32**	**0.23**	**0.43**	**0.57**	

Note. Pearson-method with pairwise-deletion was used to compute correlations. Correlations in bold font are statistically significant ($p < .05$).

A confirmatory factor analysis was conducted to test how the SACT-v2 loads onto an attention control latent factor. A model with attention control, working memory capacity, and fluid intelligence was specified with each task loading onto their respective construct (Figure 4). The model had good fit; $\chi^2(24) = 59.09$, $p < .05$, CFI = .94, RMSEA [95% CI] = .07 [.05, .09], SRMR = .047. The SACT-v2 had a moderate loading on the attention control factor (.38) and there was a strong correlation between the cognitive ability constructs. These results largely replicated the findings from the first version of the task (Draheim et al., 2021) and provides further evidence for the reliability and validity of the SACT as a measure of individual differences in the ability to control attention.

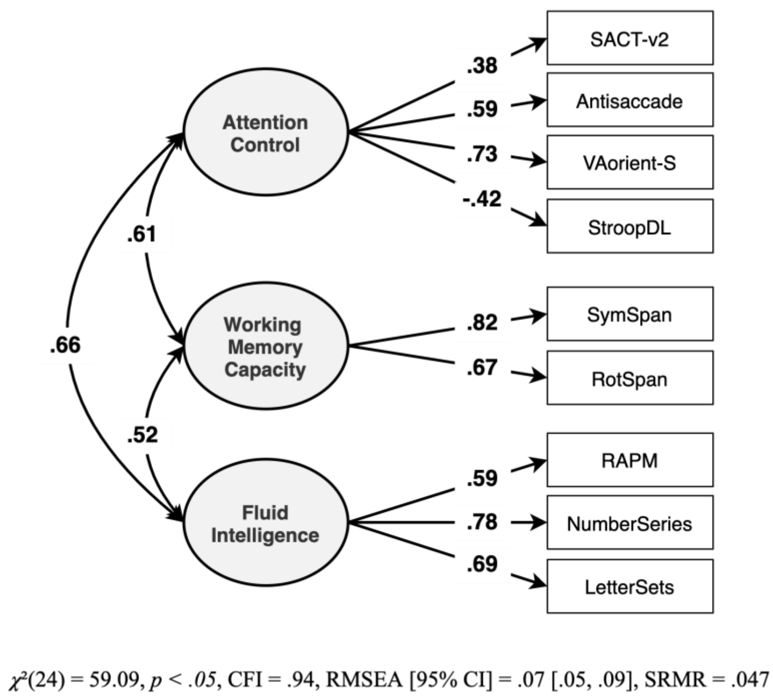

Figure 4 - Confirmatory Factor Analysis with SACT-v2

2.3.2 Validity of the SACT-v2 as a measure of sustained attention

The SACT was designed to specifically capture differences in the ability to sustain attention over time. As a reminder, the first version of the task was not able to capture performance differences across varying demands to sustain attention. Mainly, accuracy did not decrease as the amount of time attention had to be sustained over the wait time delay increased. The second version of the task was developed to hopefully improve on the validity of the SACT as a measure of sustained attention. The critical features of the task that I changed in the second version was removing the circle cue from the display during the wait interval and no longer presenting a distractor stimulus just before the target. The idea is that if the gaze shifts away from the cued location, due to a lapse in attention, then they cannot easily reorient attention back to the cued location or use the distractor as a cue for when the target is about to appear. Therefore, it is expected that performance and the vigilance decrement will get worse with longer wait times on the SACT-v2.

To test for the effects of wait time, time-on-task, and their interaction on accuracy, a 4 (Wait Time: none, short, medium, long) × 3 (Bock: 1 – 3) within-subjects ANOVA was conducted. There was a small effect of time-on-task such that accuracy decreased on later blocks; $F(1.91, 584.26) = 13.816$, $p < .001$, $\eta_p^2 = .043$. There was a large effect of wait time such that accuracy decreased at longer wait times; $F(2.093, 640.363) = 88.251$, $p < .001$, $\eta_p^2 = .224$. There was also a small time-on-task × wait time interaction such that the effect of wait time on accuracy was larger on later blocks (Figure 5); $F(4.407, 1348.663) = 14.793$, $p < .001$, $\eta_p^2 = .046$.

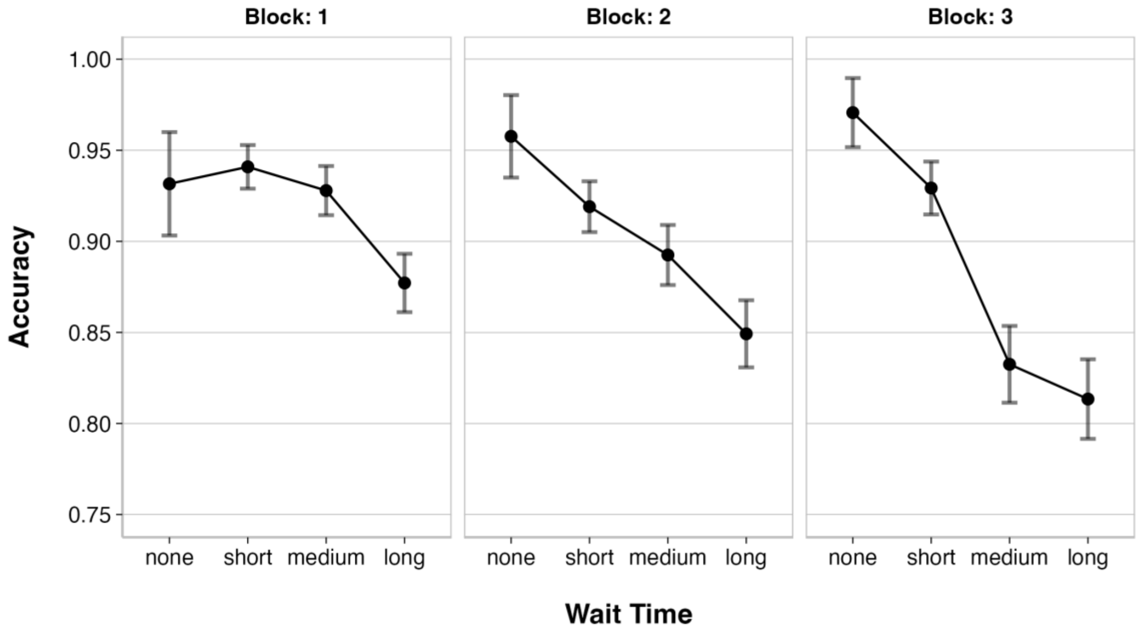

Figure 5 - Wait Time × Time-on-Task Interaction in the SACT-v2.

Note. Estimated marginal means at +1 SD above the mean, the mean, and -1 SD below the mean on SACT-v2 scores. Error bars represent 95% confidence intervals.

As seen in Figure 5, interaction contrasts indicated that the decrease in accuracy from short to long wait times was larger on block 3 compared to block 1; $t(306) = -4.38$, $p < .001$. Similarly, the decrease in accuracy from short to medium long wait times was larger on block 3 compared to block 1; $t(306) = -6.85$, $p < .001$. Another way to interpret these interaction contrasts is that there was a vigilance decrement for the long and medium wait times but not for the short (and none) wait time condition. This suggests that it is not just overall levels of performance that is decreasing with time-on-task but that it is being able to sustain attention over the wait time that is getting worse with time-on-task. These effects show, unlike the SACT-v1, that performance on the SACT-v2 declined as wait time and time-on-task increases—providing an important piece of validity of this task as capturing differences in sustained attention.

2.3.3 Individual Differences in Sustained Attention

From an individual differences perspective, do scores (overall accuracy) on the SACT-v2 task reflect differences in the ability to sustain attention over the wait time and vigilance decrements over blocks? One possibility is that everyone demonstrates equivalent effects of wait time and time-on-task on performance and that those with high SACT-v2 scores are just showing overall better performance—that is, a lack of any interactions of wait time or time-on-task with overall task scores. Alternatively, if scores on the SACT-v2 do reflect differences in sustained attention ability, then it would be expected that those with high SACT-v2 scores will show a smaller decrease in performance at longer wait times and with time-on-task—that is, there should be two-way interactions of wait time and / or time-on-task with SACT-v2 scores.

To test for this, SACT-v2 scores were added as a covariate with wait time and time-on-task in a within-subjects ANCOVA. There was a large SACT-v2 × wait time interaction; $F(2.225, 678.584) = 48.127$, $p < .001$, $\eta_p^2 = .136$. Those with higher overall scores on the SACT-v2 showed less of a decline in accuracy with longer wait times (Figure 6). There was also a small SACT-v2 × time-on-task interaction; $F(1.924, 586.931) = 11.970$, $p < .001$, $\eta_p^2 = .036$. Those with higher overall scores on the SACT-v2 showed less of a decline in accuracy with time-on-task. These results provide evidence that those who score higher on the SACT-v2 do so partly because they are better able to sustain attention over the wait time delay and with time-on-task.

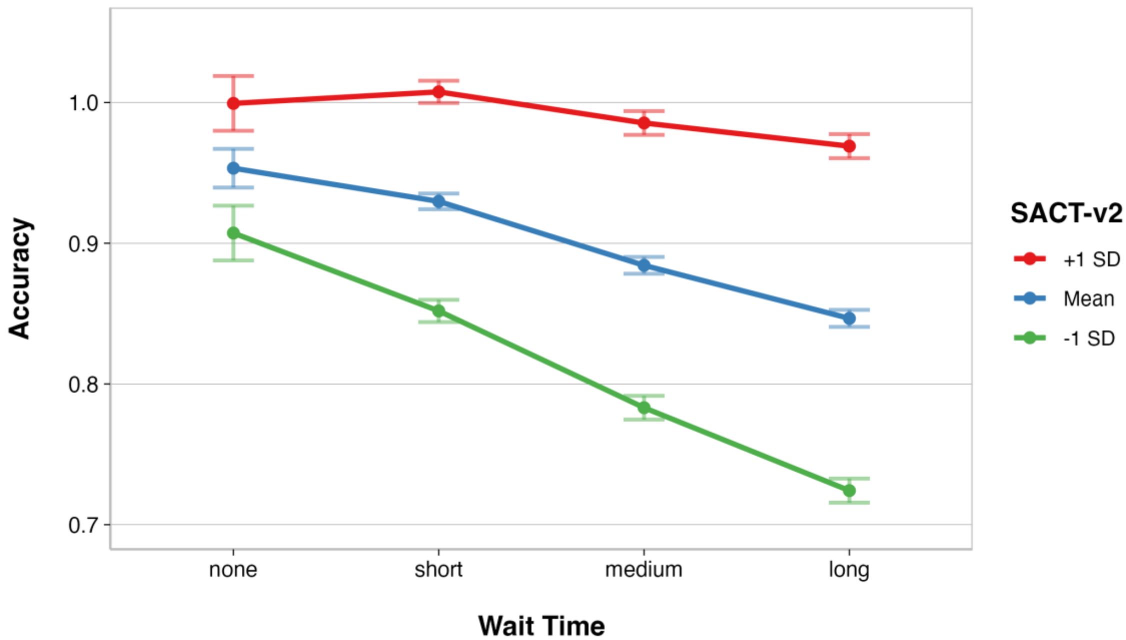

Figure 6 - SACT-v2 Scores × Wait Time Interaction

Note. Estimated marginal means at +1 SD above the mean, the mean, and -1 SD below the mean on SACT-v2 scores. Error bars represent 95% confidence intervals.

Is this true for other cognitive abilities? That is, do those higher on cognitive ability show a smaller decline in performance with longer wait times and time-on-task? There was a small attention control × wait time interaction; $F(2.137, 587.664) = 7.090$, $p = .001$, $\eta_p^2 = .025$. Those higher on attention control showed less of a decline in accuracy at longer wait times in the SACT-v2 (Figure 7). There was also a small attention control × time-on-task interaction; $F(1.922, 528.578) = 4.812$, $p = .009$, $\eta_p^2 = .017$. Those higher on attention control showed less of a decline in accuracy with time-on-task in the SACT-v2.

This was not the case with the other cognitive ability measures (Figure 8 and Figure 9). There was no interaction of working memory capacity with wait time; *F*(2.082, 597.479) = 1.026, p = .361, η_p^2 = .004; or time-on-task; $F(1.903, 546.276) = 0.002$, p = .997, η_p^2 = .000. There was no interaction of fluid intelligence with wait time; $F(2.135, 619.030) = 1.491$, p = .225, η_p^2 = .005; or time-on-task; $F(1.916, 555.515) = 1.855$, p = .159, η_p^2 = .006. However, those higher on working memory capacity did show overall better accuracy on the SACT-v2; $F(1.000, 287.000) = 6.051$, p = .014, η_p^2 = .021. Those higher on fluid intelligence did not show overall better accuracy; $F(1.000, 290.000)$ = 2.342, p = .127, η_p^2 = .008. These findings suggest that the ability to sustain attention over the wait time delay is related to attention control ability but not to individual differences in working memory capacity or fluid intelligence.

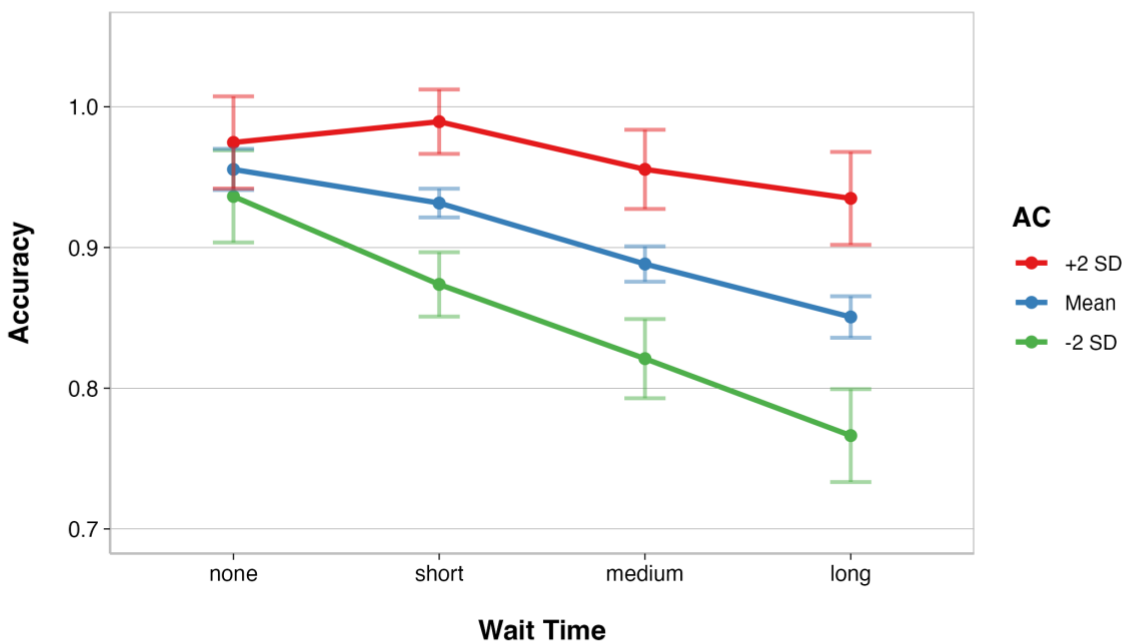

Figure 7 - Attention Control × Wait Time Interaction in the SACT-v2

Note. Estimated marginal means at +2 SD above the mean, the mean, and -2 SD below the mean on Attention Control (AC). Error bars represent 95% confidence intervals.

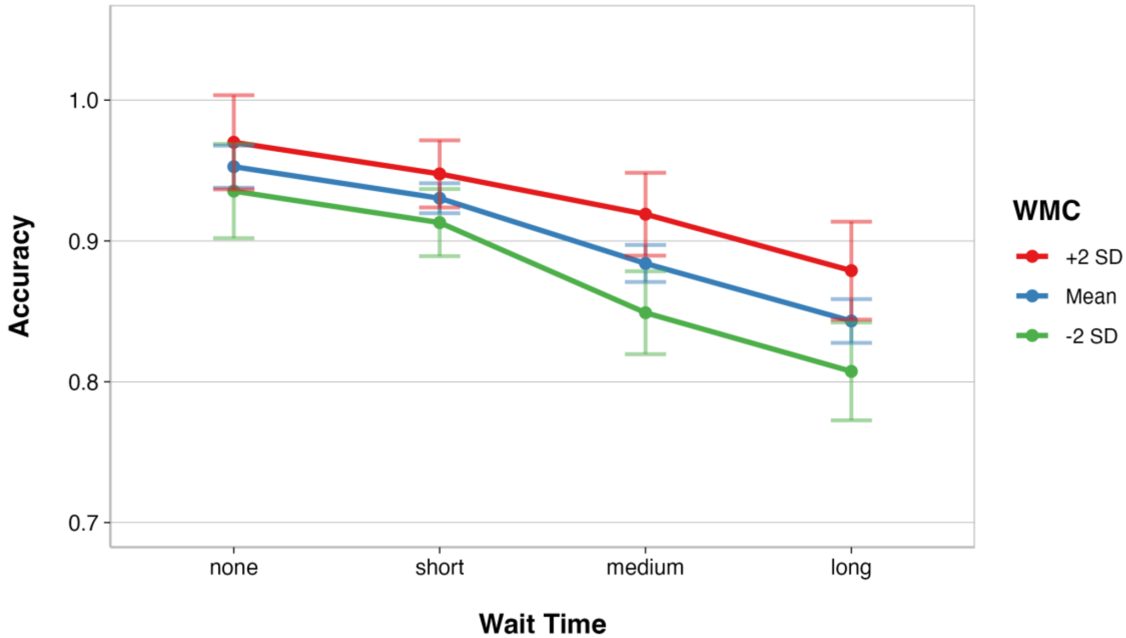

Figure 8 - No Working Memory Capacity × Wait Time Interaction in the SACT-v2

Note. Estimated marginal means at +2 SD above the mean, the mean, and -2 SD below the mean on working memory capacity (WMC). Error bars represent 95% confidence intervals.

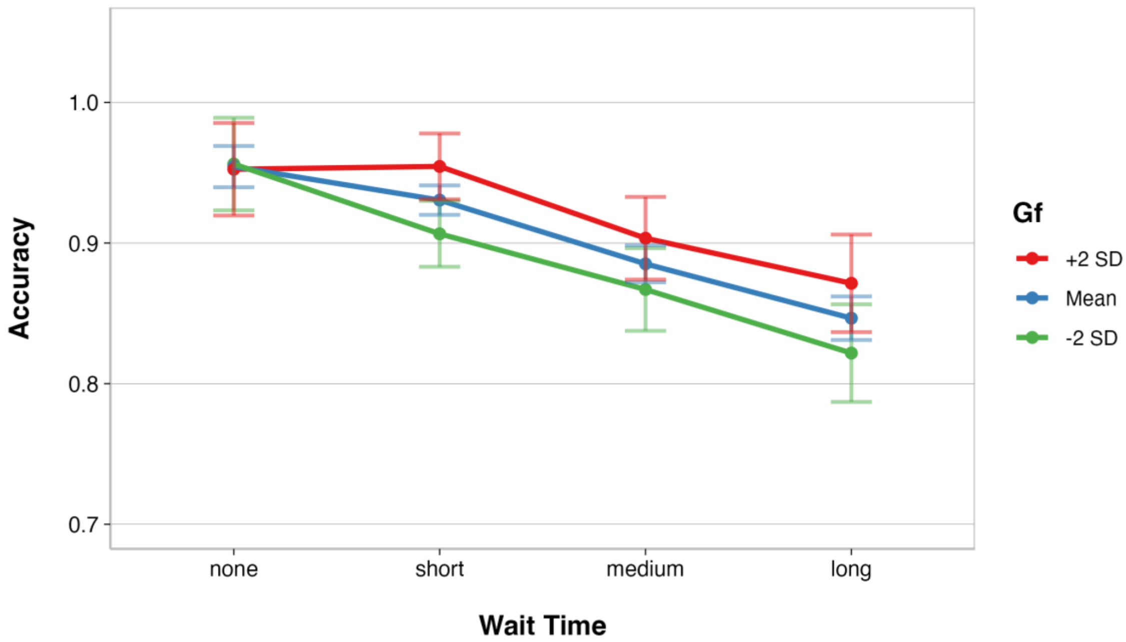

Figure 9 - No Fluid Intelligence × Wait Time Interaction in the SACT-v2

Note. Estimated marginal means at +2 SD above the mean, the mean, and -2 SD below the mean on fluid intelligence (Gf). Error bars represent 95% confidence intervals.

2.4 Discussion

The SACT-v2, like the SACT-v1, demonstrated to be a reliable and valid measure of individual differences in the ability to control attention. Because this task was designed to specifically capture differences due to being able to sustain attention over time it is important to provide evidence that performance differences on this task are due to the varying wait time delays on the task. The first version of the SACT-v1 did not show any performance differences across the variable wait times and therefore is questionable whether it was capturing variance due to sustained attention. In the second version of the task, I removed the circle cue from the wait time delay and got rid of the attention capturing distractor stimulus. The purpose of this was to place more demands on sustaining attention to the cued location over the wait time delay.

These changes were effective in producing performance differences across the varying wait time delays. First, accuracy on the SACT-v2 got worse as the wait time got longer. This effect of wait time got larger as time-on-task increased (Figure 5). Importantly, performance got worse in the medium and long wait times as time-on-task increased but performance did not change for the none and short wait times. This interaction suggests that what was waning over time was the ability to sustain attention and not the ability to engage the focus of attention for a short period of time.

These effects are all within-subject changes in performance; however, this task was designed primarily to capture between-subject differences in sustained attention. As an individual differences measure, performance on the SACT is simply computed as the proportion of correct responses across all trials. Therefore, I tested whether between-subject differences on SACT-v2 scores were capturing differences in performance across the varying wait times. I found that those who scored higher on the SACT-v2 showed less of a decline in performance as wait time got longer (Figure 6). This is a critical piece of evidence for validity that between-subject differences on the second version of the SACT are, to some extent, reflecting differences in the ability to sustain attention over the wait time.

Similarly, those who scored higher on attention control also showed less of a decline in performance as wait time got longer on the SACT-v2 (Figure 7). This is an expected result if part of what we mean by the ability to control attention includes sustaining attention over time. Curiously, however, there were no performance differences across the wait times on the SACT-v2 due to working memory capacity (Figure 8) or fluid intelligence (Figure 9). Given that sustained attention—and similar constructs such as goal-maintenance—is a frequently used explanation for differences in working memory capacity this result was unexpected. This suggests some differentiation of how attention control relates to sustained attention and working memory capacity. One possibility is that the ability to sustain attention to a simple object in the environment is different from the ability to maintain more complex relations, task-sets, and goals in working memory.

2.4.1 Limitations

The primary limitation with the SACT-v2 is the overall high rates of accuracy on the task. The SACT-v1 had a mean accuracy of 70% (Draheim et al., 2021), whereas the SACT-v2 had a mean accuracy of 89%. This difference can also be seen in comparing Figure 2 and Figure 5. This could possibly be due to the attention capturing distractor stimulus being removed and the duration of the target being on for an extra 125 ms in the SACT-v2. Although these changes likely resulted in greater demands to sustain attention over the wait time relative to resisting attention capture and detecting a briefly presented target, it would be more ideal to have accuracy be lower overall. One possibility would be to add longer wait times, even up to 30 seconds. The downside, however, is that this would result in a considerably longer administration time. Another possibility is to first use an adaptive threshold procedure to determine the target duration at a certain level of accuracy that should be used on an individual basis. This would bring accuracy rates down while also controlling for differences in the ability to quickly identify a briefly presented target item (e.g., the intensity of attention).

2.4.2 Considerations for Future Research

These results demonstrate that the sustained attention-to-cue task can be a valid measure of individual differences in sustained attention ability. They also suggest that this task can be used to study sustained attention at both a shorter time-scale (over seconds during the wait time) and a longer time-scale (over minutes during the course of the task). Most sustained attention tasks rely on behavioral performance metrics (dependent variables) such as reaction time variability, extremely long reaction times, or commission / omission errors on vigilance tasks to index attentional states related to sustained attention (e.g., stability of attention over time or a lapse of attention). However, direct manipulations (independent variables) on the demands of sustained attention are not common in the sustained attention literature. Instead, what is more common is to manipulate variables like cognitive load, stress, reward, sleep deprivation and other factors to see how these variables relate to dependent measures of sustained attention. To the contrary, the sustained attention-to-cue task, itself, includes a manipulation on the demands of sustained attention by varying the wait time delay.

The point is not that these other approaches are flawed, but that the sustained attention-to-cue task provides a novel and more direct manipulation of sustained attention. That is, if a variable—such as those mentioned above— has an impact on sustained attention then there are divergent predictions on how it should affect performance at shorter wait times vs. longer wait times. To my knowledge, there is no other task paradigm that accomplishes this.

Another unique feature of the sustained attention-to-cue task is the blank wait time delay in which nothing exogenously is happening and thus there are only demands on endogenous attention during this time. Long intervals (such as 12 seconds at the longest wait time) in which no exogenous stimulation is occurring is rare in cognitively demanding tasks. Therefore, the use of physiological measures such as eye gaze, pupil size, heart rate, and galvanic skin response may be highly informative as to what is occurring during the wait time and how that relates to the ability to successfully sustain endogenous attention. Study 2 was an experimental study that utilized these novel features of the sustained attention-to-cue task to investigate how resource depletion, task utility, internal distraction, and arousal are related to sustained attention.

CHAPTER 3. STUDY 2

3.1 Introduction

Coming back to the primary research question posed at the beginning—what are the factors that determine how long we can voluntarily sustain our attention? We found in Study 1 that individuals with a greater ability to control their attention were more successful at sustaining their attention over a relatively short period of time (12 seconds was the longest condition) and demonstrated less of a vigilance decrement. However, other cognitive abilities such as working memory capacity and fluid intelligence were not related to these effects of sustained attention.

Although individual differences in cognitive ability provide an important perspective and practical relevance to real-world performance on the job, they do not necessarily offer as much insight into factors that might change *within* the individual over time and context. Additionally, *within*-individual factors potentially have more practical and direct relevance for individuals to implement in their daily-life that do not require the more long-term and arduous goal of improving one's trait-level cognitive ability. It is important to also keep in mind, however, that *between*-individual and *within*-individual factors are not entirely independent and may even interact with one another.

Let us now turn our attention to some of the *within*-individual factors that are thought to play a role in determining how long attention can be sustained. Namely, I will focus on the factors of resource depletion, task utility, internal distraction, and arousal. Theories of sustained attention need to account for why performance changes (gets worse) over long *and* short time-scales. Most theories of sustained attention have focused on why performance decreases over long time-scales—the vigilance decrement. There are two main factors that models of sustained attention have highlighted to explain the vigilance decrement. One is the availability of a limited pool of resources—I will refer to theories that emphasize resource availability as resource depletion models (Grier et al., 2003; Warm et al., 1996). Another factor is the valuation of costs and rewards for continuing to place attention on one thing over other things—I will refer to theories that emphasize cost-reward valuation as task utility models (Esterman & Rothlein, 2019; Fortenbaugh et al., 2017; Kurzban et al., 2013).

3.1.1 Resource Depletion

Resource depletion models fall into what are known as overload accounts of the vigilance decrement (Grier et al., 2003; Warm et al., 1996). These models presume that to the extent a task is effortful, it will consume a limited pool of cognitive resources needed to successfully perform the task (Caggiano & Parasuraman, 2004; Grier et al., 2003). Thus, the longer an effortful task is performed, availability of resources to perform the task decreases. These models have created some tension in the literature on the vigilance decrement because many vigilance tasks are developed to be monotonous and under-stimulating. Nevertheless, vigilance tasks are subjectively reported to be of high workload and even stressful—a key piece of evidence that overload accounts of the vigilance decrement rely on (Hitchcock et al., 1999; Temple et al., 2000; Warm et al., 1996). There is also support for resource depletion models based on findings that greater cognitive load in traditional vigilance tasks leads to larger vigilance decrements (Helton & Russell, 2011, 2013; Smit et al., 2004). Given the emphasis on effort and resource availability, resource depletion models predict that increases in required effort should more quickly deplete available resources and performance should decline faster over time.

3.1.2 Task Utility

Task utility models fall into what are known as underload accounts of the vigilance decrement. Underload accounts presume that, because vigilance tasks are monotonous and under-stimulating, they do not require a high degree of effort. These accounts explain the vigilance decrement as due to mindlessness, boredom, disinterest, and/or under-arousal—factors associated with motivation, reward, and arousal (Esterman & Rothlein, 2019; Manly et al., 1999; Robertson et al., 1997). While task utility models recognize the limitations of resources and capacity for information processing, they explain the vigilance decrement (and performance changes more generally) as a result of changes in the valuation of a cost-reward tradeoff such that task utility decreases with time-on-task (Esterman et al., 2014, 2016; Kurzban et al., 2013). As task utility decreases it becomes more likely that resources get deployed away from the task and towards other potential sources of reward. The decrease in task utility with time-on-task can potentially be offset by valuations of potential reward (which can come from intrinsic or extrinsic sources) or even valuations of cost associated with failing to maintain engagement (e.g., performance-related cost in the form of negative feedback or monetary loss). These cost-reward valuations jointly inform the utility of continuing to engage one's resources to a task—that is, continuing to place attention on one thing over other potential sources of reward. Given the emphasis on task utility, task utility models predict that motivational and reward-driven factors that increase effort should increase task engagement and thereby reduce the vigilance decrement. That is, factors that increase effort should reduce vigilance decrement over time.

Note that the predictions made by resource depletion and task utility models on the effects of increased effort on the vigilance decrement are in opposite directions (Pattyn et al., 2008). Some researchers have recently taken advantage of this and used monetary reward manipulations in vigilance tasks to tease apart these two models (Esterman et al., 2014, 2016). In general, the findings are somewhat mixed but so far, they tend to be more consistent with a task utility model of the vigilance decrement. In the current study, a monetary incentive will be used to test resource depletion and task utility models in the sustained attention-to-cue task.

It may be that behavioral performance data on the vigilance decrement will not be enough to distinguish between resource depletion and task utility models. Both models agree that the vigilance decrement is due to fewer resources overall being applied to the task. However, they disagree on what is happening to those resources. Resource depletion models suggest that they are simply being depleted and are thereby not available for effortful task processing. Task utility models claim that they are still available, they simply are deployed to other activities. A major question for task utility models then is, where are those resources being redeployed? In real-world scenarios this could be to a large number of alternative activities. For instance, a lonely graduate student working in his lab can stop writing his dissertation and go on his phone to browse social media, check email, pester some other graduate student, daydream, and so on. For a subject participating in an experiment, however, the number of alternative activities is very small. One activity that is still prominently available for the research subject is engaging in self-generated thoughts, such as daydreaming, mind wandering, and rumination. Therefore, examining changes in thought content with changes in the vigilance decrement may provide a useful way to distinguish resource depletion and task utility models.

3.1.3 Internal Distraction

In our daily life we are met with the demands of not only navigating through noisy and stimulating environments but also of managing the maelstrom of internal thoughts, desires, and emotions. In the midst of carrying out important personal or work-related tasks, it is not uncommon for those tasks to get interrupted by thoughts unrelated to the task at hand (Killingsworth & Gilbert, 2010). Cognitive psychologists refer to these events as task-unrelated thoughts or mind wandering, and have been characterized as lapses of attention and executive attention failures to maintain task goals (Kane & McVay, 2012; Smallwood, 2013; Smallwood & Schooler, 2006). The distraction from internal thoughts is thought to be a major contributor to differences in overall task performance as well as vigilance decrements. This evidence comes from several studies showing individual differences in mind wandering frequency in relation to cognitive abilities and studies showing that mind wandering increases with time-on-task (McVay & Kane, 2009; Robison & Unsworth, 2015; Unsworth & McMillan, 2014).

The resource-control model makes explicit the relationship of mind wandering to sustained attention and vigilance decrements (Thomson et al., 2015). Typical resource depletion models explain vigilance decrements as being a result of the depletion of available attentional and cognitive resources to perform the task and thereby a reduction in task performance with increasing time-on-task. The resource-control model modifies this by assuming that the amount of available resources remains constant over the course of a task, but that resources allocated to the task decreases while the allocation of resources to mind wandering increases over the course of a task. There are several reasons for this. The resource-control model posits that mind wandering is a default state such that there is a continuous bias for attentional resources to be pulled towards mind wandering. This is supported from brain imaging studies showing that a set of brain regions dubbed the "default-mode" network is more active during a passive resting state and instances of mind wandering than during task engagement (Raichle, 2015; Raichle et al., 2001). Additionally, the default-mode network and task-positive networks tend to be anti-correlated with one another (Fox et al., 2005). This push and pull between task engagement (task-positive networks) and mind wandering (default-mode network) necessitates the need to engage in executive control to maintain task engagement even in relatively simple tasks. An important part of the resource-control model is that executive control is reduced with time-on-task resulting in fewer resources being devoted to the task and failures of goal maintenance more likely. Specifically, this model predicts that the vigilance decrement can be partially explained by a reduction in executive control and an associated increase in mind wandering.

The resource-control model can be considered as a hybrid between resource depletion and task utility models. The resource-control model specifies that it is not a general pool of resources that depletes over time, but rather it is the ability to engage executive control to continuously allocate resources to a task that becomes diminished. Although it can accommodate changes in task utility, it places more emphasis on a natural tendency to shift towards a mind wandering, default-mode state rather than a change in utility (though that natural tendency itself might simply be a utility tradeoff because self-referential thoughts likely provide a signal of other potential sources of reward). Therefore, even though the resource-control model is based more on resource depletion models the predictions it makes about the nature of resource allocation and the vigilance decrement are the same as task utility models.

The resource-control model, however, does a better job of connecting our findings from Study 1 on the nature of individual differences in attention control and sustained attention. Resource depletion and task utility models do not necessarily make any predictions about that nature of cognitive abilities and/or executive control processes and the vigilance decrement. Based on the resource-control model, one prediction might be that individuals with a greater ability to control their attention will show less of a decrement in executive control over the course of a task and therefore less of a vigilance decrement. This prediction is consistent with what we found in Study 1.

Researchers have shown that the nature of mind wandering is not a unidimensional construct but can be divided into different types of mind wandering (Seli et al., 2018). One important distinction for the current work is between intentional and unintentional mind wandering (Seli, Risko, Smilek, et al., 2016). Intentional mind wandering is when one deliberately disengages from a task to engage in self-generated thoughts. By its very definition this type of mind wandering reflects a change in task utility where mind wandering becomes of a higher utility than engaging in the task. Unintentional mind wandering, on the other hand, is when internal thoughts capture one’s attention and thus an unintentional shift away from the task and towards mind wandering occurs. These reflect slips or lapses of attention and executive control failures to maintain attention to the task.

Research has shown that these types of mind wandering are dissociable at both the individual differences and experimental level. Unintentional mind wandering is positively associated with symptomology of attention deficit hyperactivity disorder and obsessive compulsive disorder, whereas, intentional mind wandering is not (Seli, Smallwood, et al., 2015; Seli et al., 2017). Additionally, an individual's level of self-reported motivation predicts intentional but not unintentional mind wandering (Seli, Cheyne, et al., 2015; Seli, Wammes, et al., 2016). At the experimental level, intentional mind wandering is more frequent under low demanding tasks compared to high demanding tasks but rates of unintentional mind wandering do not change (Seli, Risko, & Smilek, 2016).

The distinction between intentional and unintentional mind wandering may help to delineate whether the vigilance decrement is due to a shift in task utility (more intentional mind wandering), due to a diminishing of executive control (more unintentional mind wandering), or both (more intentional and unintentional mind wandering). Similarly, it can be helpful to distinguish between resource depletion models and task utility models. Resource depletion models would predict that there should be no change in intentional mind wandering with time-on-task because the resources to do so will be depleted. In contrast, task utility models make a rather strong prediction that intentional mind wandering should increase with time-on-task as task utility decreases.

3.1.4 Arousal

One aspect of underload accounts of the vigilance decrement that is not captured in some task utility models is the factor of arousal. Under-arousal and mindlessness models explain the vigilance decrement as being due to a decrease in arousal or mindful attentiveness (Manly et al., 1999; Robertson et al., 1997). Therefore, rather than an overload of cognitive resources or changes in task utility, these models suggest that a general inattentiveness, due to low arousal, is a major factor leading to the vigilance decrement. The adaptive gain theory of locus coeruleus function provides a convenient framework for understanding the relationship among levels of arousal, task utility, and task performance (Aston-Jones & Cohen, 2005).

The locus coeruleus is a small cluster of neurons located in the pontine brainstem and is the sole source of norepinephrine to the cortex (Berridge & Waterhouse, 2003). The locus coeruleus-norepinephrine system has traditionally been associated with functions of arousal, similar to many brainstem nuclei (Sara, 2009). However, advancements in the study of this system have revealed that it plays a much larger role in perception, learning, memory, cognition, and behavior (Aston-Jones & Cohen, 2005; Berridge & Waterhouse, 2003; Sara, 2009; Tsukahara & Engle, 2021). Specifically, the adaptive gain theory proposes that the locus coeruleus-norepinephrine system acts as an attentional filter for optimal processing of task/goal-relevant stimuli and behavioral responding (Aston-Jones & Cohen, 2005). A crucial element of the adaptive gain theory is the role of two pre-frontal cortical areas in regulating locus coeruleus activation. The anterior cingulate cortex plays a role in the evaluation of performance-related cost and effort while the orbital frontal cortex plays a role in the evaluation of performance-related reward. These two areas conjointly influence locus coeruleus activation towards either an exploitative mode, moderate tonic and high phasic activation, or exploration mode, high tonic and low phasic activation. Exploitation mode is when task utility is high and therefore the locus coeruleus is tuned towards optimal task engagement. Exploration mode is when task utility is low and therefore locus coeruleus is tuned towards reward seeking and task disengagement.

So far this is largely consistent with task utility models of sustained attention. The added value of the adaptive gain theory is that these modes of locus coeruleus activation are associated with different levels of arousal. The exploitation mode, high task utility, is associated with moderate and optimal levels of arousal for task performance. The explorative mode, low task utility, is associated with high levels of arousal and states of distractibility. Less of a focus for the adaptive gain theory but still an important part of locus coeruleus function, are low arousal states in which the locus coeruleus shows low tonic and low phasic activation associated with inattentiveness and low responsiveness to stimulation. Presumably task utility will be low in such states—but due to low responsiveness, reward seeking and distractibility should be also be low.

Due to the location of the locus coeruleus, this system has not been easy to study in humans. Fortunately, there is considerable evidence showing the changes in pupil size correspond to changes in locus coeruleus activity (Joshi et al., 2016; Joshi & Gold, 2020; Laeng et al., 2012; Murphy et al., 2014; Rajkowski et al., 1993). This has allowed researchers a non-invasive and inexpensive way to study the locus coeruleus-norepinephrine system in humans. In the current study, the use of pupillometry will be informative to dissociate levels of arousal and changes in task utility. If changes in sustained attention at large *or* small time-scales is due to changes in arousal then pupil size should show either a decrease or increase in tonic pupil size (reflecting either too low or too high of tonic locus coeruleus activity) over a period of time and show a decreased phasic pupil dilation to critical task events. Increases in performance-related reward should lead to greater task engagement and larger pupil dilations. Resource depletion models would predict that pupil dilations should get smaller with time-on-task even with increased performance-related reward because the availability of resources for task disengagement will be diminished either way. If anything, they might predict even smaller pupil dilations with time-on-task under increased performance-related reward (reflecting more of a vigilance decrement due to higher effort and draining of resources). Task utility models predict performance-related reward will either increase or keep pupil dilations at the same level with increasing time-on-task (reflecting less of a vigilance decrement due to higher task utility).

3.1.5 Current Study

Above I have outlined some of the literature on how the factors of resource depletion, task utility, internal distraction, and arousal interact or in some cases lead to different predictions on sustained attention performance. In order to test how these different factors interact, in the current study I manipulated monetary incentives, obtained self-report measures of mind wandering and arousal, and measured physiological indicators of attention and arousal with eye gaze, pupil size, heart rate, and galvanic skin response in the sustained attention-to-cue task. The assumption is that monetary incentive will lead to a greater deployment of effort than no monetary incentive. This is critical because resource depletion and task utility models of sustained attention have opposite predictions as to what happens, in the long-run, when greater effort and thereby more resources are continuously devoted to a task. Based on these models, there are different hypotheses as to how performance (accuracy and eye gaze), mind wandering, arousal, and pupil size would differ with longer wait times and time-on-task effects across monetary incentive vs. no monetary incentive groups on the sustained attention-to-cue task.

3.1.5.1 Dependent Measures

Given that the sustained attention-to-cue task is essentially measuring how long one can sustain attention at a spatially cued location, eye gaze may be an even more direct measure of sustained attention than accuracy at the end of a trial. Two main metrics of eye gaze were considered in this study. One was how long it takes eye gaze to shift away from the cued location. Another was whether the eye gaze position was at the cued location at the end of the trial or not. Therefore, including accuracy, we had three measures of sustained attention performance.

Mind wandering was measured using thought probes similar to that administered by Unsworth and colleagues. The advantage to their method is that an option is included for on-task, off-task, and inattentive states of mind. The addition of the inattentive option allows for a test of arousal on sustained attention. Additionally, another thought probe was administered to get at the intentionality of mind wandering. Finally, a third thought probe was administered to get subjective ratings of arousal on a continuous scale. These thought probes occurred on one fourth of the trials. Finally, pupil size and heart rate were continuously recorded to provide two separate markers of arousal.

3.2 Method

3.2.1 Subjects

The proposed study was conducted at the Georgia Institute of Technology, Atlanta, GA, USA. All subjects were required to be native English speakers and 18-35 years of age. No screening on vision or hearing was performed. We recruited subjects from the undergraduate population at Georgia Tech enrolled in an undergraduate psychology course and were given 1 hour of SONA credit. This study's protocol H22301 was approved by the Georgia Institute of Technology's Institutional Review Board.

The critical manipulations and analyses were on reward, wait time, and time-on-task. Therefore, power analyses were performed on the 2 (Reward) × 3 (Wait Time) and the 2 (Reward) × 4 (Time-on-Task) interactions of interest with reward as the only between-subject factor. The method of the smallest effect size of interest was used to determine the adequate sample size to find an interaction effect with a partial $\eta^2 = .02$ ($f = .14$), power of 0.8, and alpha level of .05. Using G*Power 3.1, the reward × wait time interaction power analysis indicated that a total sample size of 82 is needed (Figure 10). The reward × time-on-task interaction power analysis indicated that a total sample size of 70 is needed. Therefore, I aimed for a target sample size of at least 82.

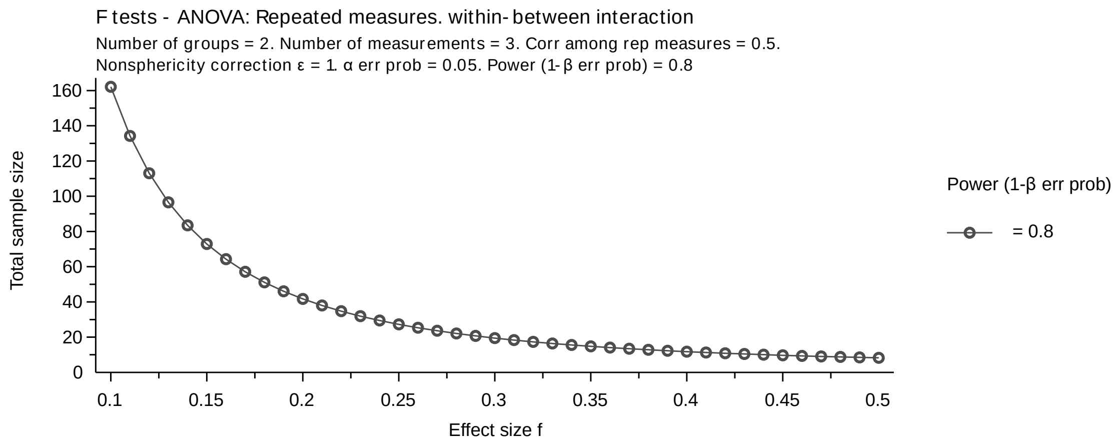

Figure 10 - Reward × Wait Time Power Analysis in G*Power

A total of 123 subjects completed the study, 63 in the no-reward group and 60 in the reward group. Several subjects were excluded from all analyses for various reasons such as program error/crashed, they decided to withdraw, they were caught "cheating" by placing their finger on the screen where the cued location was, or they thought the reward bonus was an experimental deception and therefore they performed assuming they would receive no reward bonus. This resulted in a final sample of 115 subjects, 59 in the no-reward group and 56 in the reward group.

3.2.2 Procedure

The study consisted of one session lasting approximately 45 minutes to 1 hour. All subjects performed a modified version of the SACT-v2. The design of the study was a 2 (Reward) × 3 (Wait Time) × 4 (Time-on-Task) mixed design with wait time and time-on-task as within-subject factors and reward as a between-subject factor. I chose a between-subject design to eliminate order effects of switching from a rewarded condition to a non-rewarded condition.

There were an equal number of short (1 – 3.5 seconds), medium (6 – 8.5 seconds), and long (11 – 13.5 seconds) wait times with 9 possible durations, separated by about 313 ms, for each category and a total of 27 different durations. The wait times were grouped and spaced out in such a way as to allow a greater number of trials at each category of wait times to increase power and at the same time reduce the chances that subjects can predict when the wait time delay will end. The entire task took about 42 minutes and was broken up into 4 blocks of trials with a short rest period between each block. The rest period included 7 seconds of block level feedback, 6 seconds of a blank rest screen, and 7 seconds of the reward value of the upcoming block. The block level feedback for the no-reward and performance-based reward groups displayed the percent of trials in which a wrong response was made. For the performance-based reward group, the amount lost on the block and the total amount of bonus left was also displayed. After the rest break and before the start of the next block, an eye tracker status screen appeared to make sure the subject was still positioned centrally, was 60 – 70 cm from the display, and both eyes could be detected. Additionally, after block 2 (half-way through) a 5-point re-calibration was performed.

For the reward manipulation, half of the subjects performed the task with the potential to earn a monetary reward based on their performance and the other half were not offered a monetary reward. Whether a subject was in the performance-based reward or no-reward group was determined semi-randomly. The experiment was conducted in a group running room with four subject stations; therefore, to reduce the chances that subjects would discover there are reward and no-reward groups, subjects scheduled on the same day were all assigned to the same group and the reward/no-reward group assignment alternated from one day to the next. For those in the performance-based reward group, they were told that they could earn a $50 bonus and that they lost a certain amount of that bonus every time they made an incorrect response. If the amount they will lose on every trial remained constant across the entire duration of the task, then this would still result in a considerable reduction of task utility over the duration of the task (Esterman et al., 2016) because the amount of money that is at stake would have been considerably smaller at the end of the task (e.g., $5) than at the beginning ($50). As an attempt to mitigate this considerable reduction in task utility, the potential for loss increased after each block. On the first block they could lose up to $5 (10% of the overall amount) resulting in a loss amount of $0.19 for every incorrect response; on the second block they could lose up to $10 (20% of the overall amount) resulting in a loss amount of $0.37 for every incorrect response; on the third block they could lose up to $15 (30% of the overall amount) resulting in a loss amount of $0.56 for every incorrect response; and on the fourth block they could lose up to $20 (40% of the overall amount) resulting in a loss amount of $0.74 for every incorrect response. This created a higher performance-related cost as time-on-task increased.

After completion of the study, all subjects were debriefed and were told that there were different experimental groups, without providing any details about the differences in reward, and were therefore asked to not talk about the nature of the study with their friends and peers to preserve the scientific merit of the study. This was done to ensure that the reward manipulation was not contaminated by subjects knowing about the reward ahead of time, particularly those in the no-reward group. They were also asked, at the completion of the study, as to whether they heard anything about the study before participating. No subjects reported hearing anything about the nature of the study or reward before participating.

3.2.2.1 Task Details

As in the sustained attention-to-cue task version 2, subjects need to sustain their attention for a variable amount of time at a visually cued location and ultimately identify a target letter presented briefly at the center of the cued location (see Figure 11). The stimuli were presented against a grey background (RGB: 128,128,128). Each trial started with a central black fixation for 1 second. After the fixation, a 750ms interval occurred in which the words “Get Ready!” were displayed at the to-be-cued location along with an auditory beep. This interval was to prepare and focus the subject for the upcoming trial. A circle cue was then displayed for approximately 500ms in dark gray font (RGB: 174,174,174). To orient the subject on the circle cue, the large circle began to immediately shrink in size until it reached a fixed size. Once the cue reached the fixed size, it was removed from the display. The display remained blank over the entire variable wait time delay. The wait time delay lasted for either a short, medium, or long duration (defined above). After the variable wait time delay, a cloud of 1 target letter and 8 non-target letters were displayed at the cued location. The center of the cloud was at the center of the cued location where a target letter was displayed in dark gray font (RGB: 152,152,152). The surrounding non-target letters were displayed in a slightly brighter gray font (RGB: 178,178,178). The target and non-target letters were randomly determined from a list of letters (B, P, R, C, G, E, F, O, Q) in Arial font with no repeats. The set of possible letters was determined based on the desirability to achieve low levels of chance performance when guessing and a combination of letters that have high similarity with at least one other letter in the set. The array of letters was displayed for 175 ms, after which all the letters were masked with a "#" for 2 seconds to allow the pupil response to return to baseline before responding. Following this, a response screen with a box for each letter was displayed and a mouse was used to select the target letter. After the response, there was a blank buffer display presented for 1.5 seconds, again to allow the pupil response to return to baseline before the start of the next trial.

The location of the cued location was semi-randomly determined such that there was an equal number of locations in the top-left, top-right, bottom-left, and bottom-right quadrants of the display. However, the location of the cue and target were more centrally constrained in this modification of the task in order to better control for the effect of peripheral eye gaze on pupil size (Hayes & Petrov, 2016). In the array of letters, the target letter always occurred at the center of the cued location. The non-target letters were semi-randomly dispersed within an 80×80 pixel square around the target letter with a minimum stimulus distance of 20 pixels to prevent any overlap.

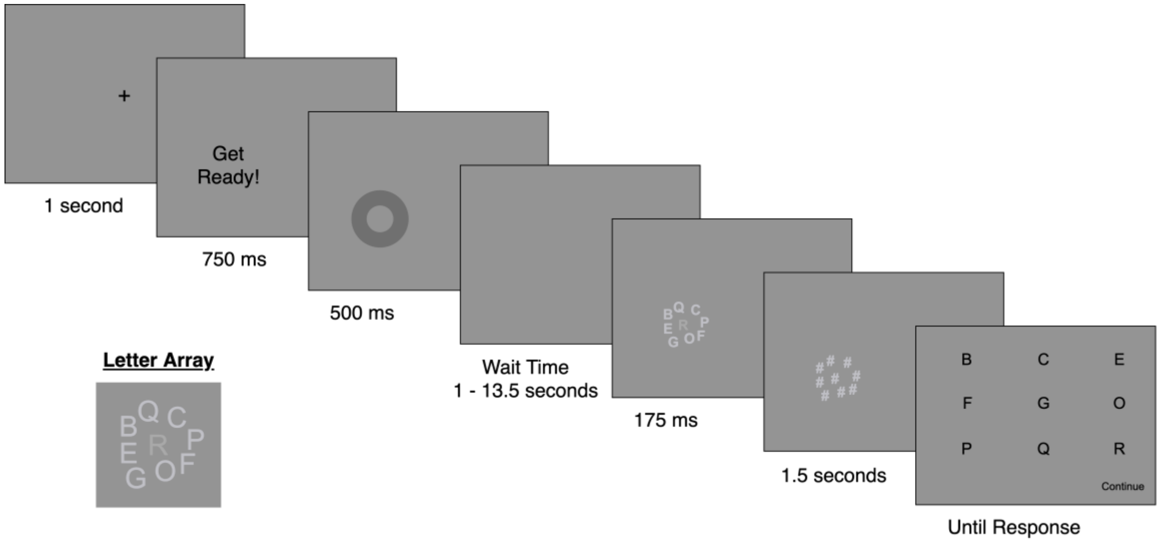

Figure 11 - Trial Procedure for SACT in Study 2

3.2.2.2 Self-Report Measures

Self-report thought and arousal probes occurred after 1/3 of the trials. At the end of the trial, once a response was made, a thought probe was displayed asking:

Which option best represents your state of mind during this latest trial?

You may have experienced one or more of these but try to respond to which one was more dominant and frequent.

1. I was focused on the cued location the entire time

2. I was focused on the task but not at the cued location

This can include trying to reorient back to the cued location, or thinking about your performance on the task overall

3. I was focused on / distracted by sights or sounds in the environment, or to physical sensations (like hunger, thirst, or temperature).

4. I was focused on / distracted by thoughts unrelated to the task (daydreaming or mind wandering)

5. I was not focused on anything in particular/I was spaced out

Four main dependent measures were obtained from the responses to this thought probe; 1) *Optimal Task Focus* represents responses to option 1, 2) *Task Engagement* represents responses to option 2, 3) *Mind Wandering* represents responses to either option 3 or option 4, and 4) *Inattentive* represents responses to option 5.

If they responded to options 3 or 4, a second thought probe was displayed to get at the intentionality of their task disengagement / distraction:

Which option best represents your state of mind during this latest trial?

1. At some point during the trial, I decided not to focus on the cued location or the task

2. I did not mean to get distracted, it just happened

Responses to this thought probe were coded as either *Intentional Mind Wandering* or *Unintentional Mind Wandering*. After the thought probes, a self-report arousal scale (Åkerstedt & Gillberg, 1990) was displayed asking: “Rate your level of energy using the scale below”. A scale was presented as a sequence of numbers, 1–9, from left to right. Some of the numbers had a descriptive label underneath; 1 = “Extremely sleepy – fighting sleep", 3 = "Sleepy – but no difficulty remaining awake", 5 = "Neither Alert nor Sleepy", 7 = "Alert", 9 = "Extremely Alert".

At the end of the task subjects were asked to self-report on their motivation, difficulty of the task, and the unpleasantness of the task. The motivation questions asked, "How motivated were you to perform well on the task?" with a labeled scale from 1 – 5; 1 = "Not at all motivated", 2 = "Slightly motivated", 3 = "Moderately motivated", 4 = "Highly motivated", and 5 = "Extremely motivated". The difficulty question asked, "How difficult did you find the task?" with a labeled scale from 1 – 5; 1 = "Not at all difficult", 2 = "Slightly difficult", 3 = "Moderately difficult", 4 = "Highly difficult", and 5 = "Extremely difficult". The unpleasantness questions asked, "How unpleasant did you find the task?" with a labeled scale from 1 – 5; 1 = "Not at all unpleasant", 2 = "Slightly unpleasant", 3 = "Moderately unpleasant", 4 = "Highly unpleasant", and 5 = "Extremely unpleasant".

3.2.2.3 Eye Tracking

A Tobii Pro Fusion eye tracker was used to continuously record eye gaze and pupil size binocularly at 250 Hz throughout the duration of the task. Subjects were seated 60 – 70 cm from the monitor and no head mobilization device was used. A 5-point eye gaze calibration procedure was conducted before task instructions. After every block, a track status screen appeared to make sure the subject was still positioned centrally, was 60 – 70 cm from the display, and both eyes could be detected. Additionally, after block 2 (half-way through) a 5-point re-calibration was performed.

Two main metrics of eye gaze were considered in this study. One was the proportion of eye gaze samples on the cued location at the end, last 500 ms, of the wait time delay.

Another was how long it took eye gaze to shift away from the cued location. The time it took eye gaze to shift away was calculated as the onset time of the first eye gaze shift away from the cued location. Eye gaze shifts away from the cued location were identified as epochs of gaze samples that were not on the cued location for more than 250 ms. Eye gaze samples that had x,y-coordinates within a 55 × 55 px square around the center of the cued location were considered as eye gaze samples on the cued location and any samples outside of those coordinates were considered as eye gaze samples not on the cued location. This region covered most, but not all, of the letter array display.

Pupil data were preprocessed using the pupillometry R package. Blinks were extended 100 ms in both directions and then pupil values at blinks were set to missing. Pupil values that exceeded a threshold of speed of change were detected and set to missing using the median absolute deviation (Kret & Sjak-Shie, 2019) with a threshold of n = 6. The pupil data were then smoothed using a moving-window average with n = 100 samples, followed by a cubic-spline interpolation but not interpolating over 1 second gaps of continuous missing data. The values to use for the median absolute deviation, smoothing, and interpolation was based on a visual inspection of the data from multiple trials and multiple subjects. Baseline correction was applied to two different epochs in the trial sequence by subtracting out the median of the preceding 200 ms. The first epoch was from the onset of the "Get Ready!" signal to the offset of the wait time delay. The second epoch was from the onset of the target stimulus to the end of the trial. After all preprocessing steps, the amount of missing data per trial was assessed. Trials with greater than 33% missing data were removed from further processing.

Several dependent measures were assessed. 1) Changes in pretrial pupil size across the reward groups and blocks. Pretrial pupil size was calculated as the average pupil size during the fixation period at the start of a trial. 2) Change in pupil size over the wait time was calculated as the average baseline-corrected pupil size during the last 500 ms of the wait time. 3) Pupil dilation to the onset of the get ready signal and 4) pupil dilation to the onset of the target stimulus. Pupil dilation was calculated as the average baseline-corrected pupil size from the onset of the event to 1500 ms after the onset time.

3.2.2.4 Heart Rate and Galvanic Skin Response

A GazePoint Biometrics finger-based system was used to continuously record heart rate and galvanic skin response at 150 Hz throughout the duration of the task. The finger-based biometric device was setup before administration of the task. The subject was told to place their finger with the device in a comfortable position on the table or chair arm rest and was reminded to do so at the start of every block. The primary measure that was assessed was changes in pretrial heart rate and galvanic skin response across the reward groups and blocks.

3.2.3 Overview of Study Design and Dependent Measures

The study is 2 (Reward Group) × 3 (Wait Time) × 4 (Block) mixed design with reward group as a between-subject independent variable and wait time and block as within-subject independent variables. The dependent variables are:

- Average accuracy
- Proportion of eye gaze samples on the cued location at the end of the wait time delay
- Average time to for eye gaze to shift away from the cued location
- Frequency of *Optimal Task Focus*, *Task Engagement*, and *Mind Wandering*
- Frequency of *Inattentiveness*
- Frequency of *Intentional* vs. *Unintentional Mind Wandering*
- Average arousal rating
- Pretrial pupil dilation, heart rate, and galvanic skin response
- Pupil size at the end of the wait time delay
- Pupil dilation to the get ready signal and target stimulus

3.2.4 Research Questions and Hypotheses

To better organize all the interacting pieces of this study, the hypotheses were formulated around three primary research questions.

1. **Are we more susceptible to distraction the longer we attempt to sustain our attention?**

Distraction and lapses of attention are expected to increase the longer attention must be sustained.

a. Accuracy will decrease with wait time and time-on-task. The effect of wait time on accuracy will be larger as time-on-task increases.

b. Lower probability of eye gaze at the cued location as wait time increases. Eye gaze will shift away from the cued location more quickly as time-on-task increases.

c. On-task reports will decrease with wait time and with time-on-task. The effect of wait time on mind wandering will be larger as time-on-task increases.

2. **Are failures of sustained attention due to fluctuations of arousal modulated by the locus coeruleus?**

2.1. Arousal and mindlessness models of the vigilance decrement predict that arousal should decrease as wait time and time-on-task increases.

a. Frequency of inattentiveness reports will increase with wait time and time-on-task. The effect of wait time on inattentiveness will increase as time-on-task increases.

b. Self-report arousal ratings will decrease with wait time and time-on-task. The effect of wait time on arousal rating will increase with time-on-task.

c. Pupil size over the wait time will decrease and it will decrease more at longer wait times. Pre-trial baseline pupil size and heart rate will decrease with time-on-task. Pupil dilation to the target will be smaller at longer wait times. Pupil dilation to the cue and target will decrease with time-on-task.

2.2. Task utility models predict that arousal will not necessarily decrease because the vigilance decrement is due to changes in utility not resources availability or arousal level. If changes in task utility *are* due to high locus coeruleus activity (indicative of a reward-explorative mode) then arousal is expected to increase as wait time and time-on-task increases.

These hypotheses can be distinguished based on the following dependent measures:

a. Frequency of inattentiveness reports will either stay the same or decrease with wait time and time-on-task.

b. Self-report arousal ratings will either stay the same or increase with time-on-task.

c. Pupil size over the wait time will either stay the same or increase at longer wait times. Pre-trial baseline pupil size and heart rate will either stay the same or increase with time-on-task. Pupil dilation to the target will be smaller at longer wait times. Pupil dilation to the cue and target will decrease with time-on-task.

3. **Are failures of sustained attention due to depletion of resources or due to diminishing task utility?**

3.1. Resource depletion models

a. The effect of wait time and time-on-task on accuracy (see hypothesis 1a) will either be the same or larger for the monetary incentive group.

b. The effect of wait time and time-on-task on the eye gaze metrics (see hypothesis 1b) will either be the same or larger for the monetary incentive group.

c. If there is an effect of wait time and time-on-task on mind wandering (see hypothesis 1c), the increase in mind wandering will be due to increases in unintentional mind wandering and not intentional mind wandering. The effect on mind wandering with wait time and time-on-task will either be the same or smaller—due primarily to unintentional mind wandering—for the monetary incentive group.

d. The change in pupil size over the wait time (see hypothesis 2.1c and 2.2c) will be the same for the monetary incentive and no monetary incentive groups. The effect of wait time and time-on-task (see hypothesis 2.1c and 2.2c) on pupil dilation to the cue and target will be the same for the monetary incentive and no monetary incentive groups.

3.2. Task utility models

a. The effect of wait time and time-on-task on accuracy (see hypothesis 1a) will smaller for the monetary incentive group.

b. The effect of wait time and time-on-task on the eye gaze metrics (see hypothesis 1b) will be smaller for the monetary incentive group.

c. If there is an effect of wait time and time-on-task on mind wandering (see hypothesis 1c), the increase in mind wandering will be due to increases in intentional mind wandering and not unintentional mind wandering. The effect on mind wandering with wait time and time-on-task will be smaller—due primarily to intentional mind wandering—for the monetary incentive group.

d. The change in pupil size over the wait time (see hypothesis 2.1c and 2.2c) will be smaller for the monetary incentive group. The effect of wait time and time-on-task (see hypothesis 2.1c and 2.2c) on pupil dilation to the cue and target will be smaller for the monetary incentive group.

3.2.5 Data Processing and Analysis

All data processing, cleaning, scoring, and analyses were conducted in R statistical software (R Core Team, 2020) and using the *tidyverse* R packages. Prior to conducting statistical models, outliers for each dependent variable were identified and removed per experimental condition. Outliers were identified as being +/- 3.3 SDs from the mean.

ANOVAs were conducted using *afex*, violations of sphericity were tested using Mauchly's test for Sphericity, Greehouse-Geisser correction was used across all within-subjects tests. For post-hoc main effect and interaction contrast comparisons, the *emmeans* package was used and Holm-Bonferroni p-value correction was used to control for multiple comparisons. All plots were made using *ggplot2*.

3.3 Results

3.3.1 Are we more susceptible to distraction the longer we attempt to sustain our attention?

The hypothesis for the first research question is that distraction and the waning of attention are expected to increase the longer attention must be sustained. For this hypothesis, analyses were conducted only on the no-reward group because this hypothesis is not concerned with the reward manipulation and the no-reward group served as the baseline.

Accuracy. If attention wanes over the wait time delay and as time-on-task increases, then accuracy on the SACT should decrease at longer wait times and at later blocks. There was a large decrease in accuracy as wait time got longer ($F(1.57, 81.45) = 56.73$, $p < .001$, $\eta_p^2 = .52$), see Figure 12a. Post-hoc comparisons showed that all pairwise comparisons between short, medium, and long wait times were significant (**short-medium**: diff = .09, Cohen's D = .45, $t(52) = 5.75$, $p < .001$), **short-long**: diff = .19, Cohen's D = .89, $t(52) = 8.69$, $p < .001$, **medium-long**: diff = .09, Cohen's D = .44, $t(52) = 6.70$, $p < .001$). There was no effect of time-on-task on accuracy ($F(1.79, 92.95) = 0.99$, $p = 0.367$, $\eta_p^2 = .02$), see Figure 13a. There was also no wait time × time-on-task interaction on accuracy ($F(4.96, 257.88) = 0.789$, $p = .557$, $\eta_p^2 = .02$).

Eye gaze. Two eye gaze metrics were analyzed, 1) *gaze sustained*: the proportion of eye gaze samples on the cued location at the end of the wait time delay, and 2) *gaze shift*: average time for eye gaze to shift away from the cued location.

If attention wanes over the wait time delay and as time-on-task increases, then eye gaze sustained on the cued location should become less likely at longer wait times and later blocks. Eye gaze was less likely to be at the cued location as wait time increased ($F(1.51, 76.89) = 39.98$, $p < .001$, $\eta_p^2 = .44$), see Figure 12b. Post-hoc comparisons showed that all pairwise comparisons between short, medium, and long wait times were significant (**short-medium**: diff = .07, Cohen's D = .25, $t(51) = 4.74$, $p < .001$, **short-long**: diff = .18, Cohen's D = . 63, $t(51) = 7.16$, $p < .001$, **medium-long**: diff = .11, Cohen's D = . 37, $t(51) = 5.63$,

$p < .001$). There was no effect of time-on-task on gaze sustained ($F(1.74, 88.58) = 0.65$, $p = 0.506$, $\eta_p^2 = .01$), see Figure 13b. There was also no wait time × time-on-task interaction on gaze sustained ($F(4.86, 247.99) = 0.79$, $p = 0.556$, $\eta_p^2 = .02$).

For the gaze shift measure, analyses were only conducted on the long wait times, the condition that provided the most time for eye gaze to shift away[2]. On average, eye gaze shifted away from the cue at about 7.4 seconds—longer than most of the medium wait times and shorter than all the long wait time conditions. If attention wanes as time-on-task increases, then eye gaze should shift away from the cued location more quickly as time-on-increases. However, there was no effect of time-on-task on gaze shift ($F(2.66, 149.09) = 2.14$, $p = 0.105$, $\eta_p^2 = .04$), see Figure 13c.

Optimal task focus. If distraction becomes more likely over the wait time delay and as time-on-task increases, then there should be a decrease in self-reports of optimal task focus. Optimal task focus decreased as wait time increased ($F(1.97, 110.09) = 43.57$, $p < .001$, $\eta_p^2 = .44$), see Figure 12c. Post-hoc comparisons showed that all pairwise comparisons between short, medium, and long wait times were significant (**short-medium**: diff = .10, Cohen's D = .25, $t(56) = 4.59$, $p < .001$, **short-long**: diff = .20, Cohen's D = .50, $t(56) = 8.86$, $p < .001$, **medium-long**: diff = .10, Cohen's D = . 26, $t(56) = 5.04$, $p < .001$). Optimal task focused decreased as time-on-task increased ($F(2.17, 121.35) = 5.73$, $p = .003$, $\eta_p^2 = .09$), see Figure 13d. Post-hoc comparisons showed that block 1 had the highest level of optimal task focus compared to the other blocks (**block 1v2**: diff = .09, Cohen's D = .23, $t(56) = 2.79$, $p = .029$; **block 1v3**: diff = .12, Cohen's D = .31, $t(56)$ =

[2] Analyses were also conducted on all wait times and trials and the results were the same.

3.42, p = .007; **block 1v4**: diff = .15, Cohen's D = .38, $t(56)$ = 2.96, p = .023). There was no wait time × block interaction on optimal task focus ($F(5.63, 315.37) = 0.91$, p = .482, η_p^2 = .02).

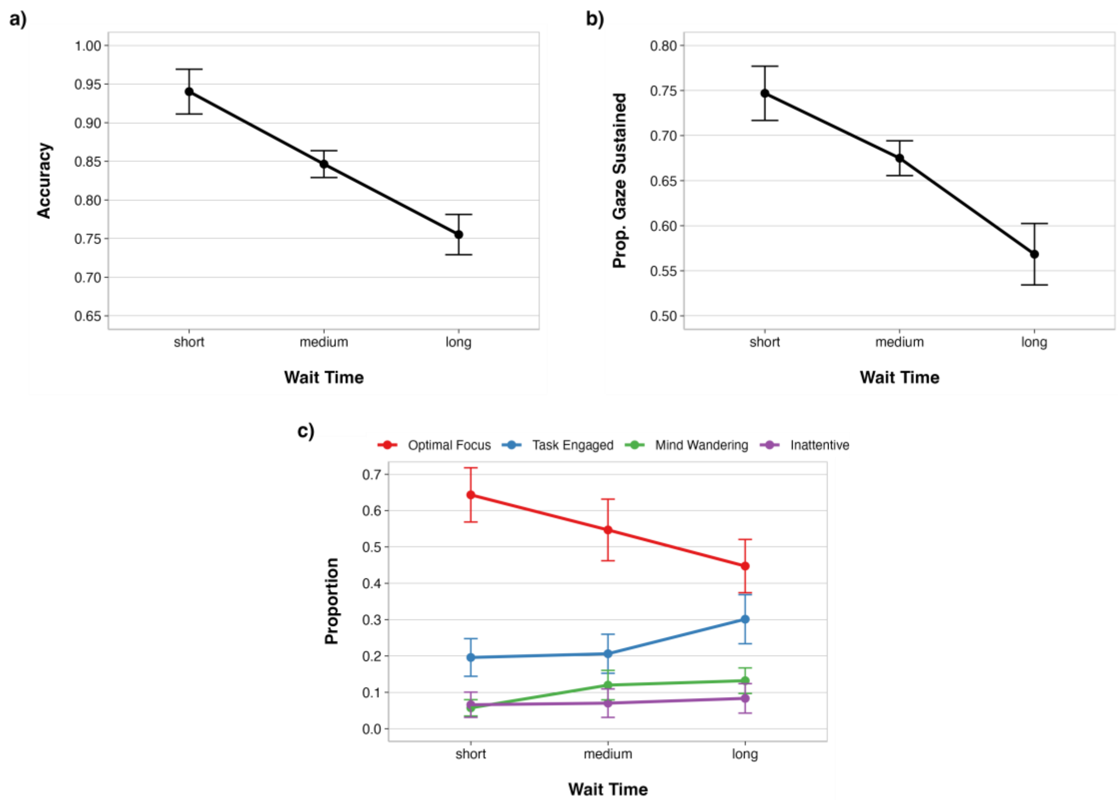

Figure 12 - Effect of Wait Time on the Waning of Attention

Note. Across four metrics a) accuracy, b) gaze sustained, c) gaze shift, and d) self-reported attentional state, the focus of attention waned as wait time increased. Error bars represent 95% confidence intervals.

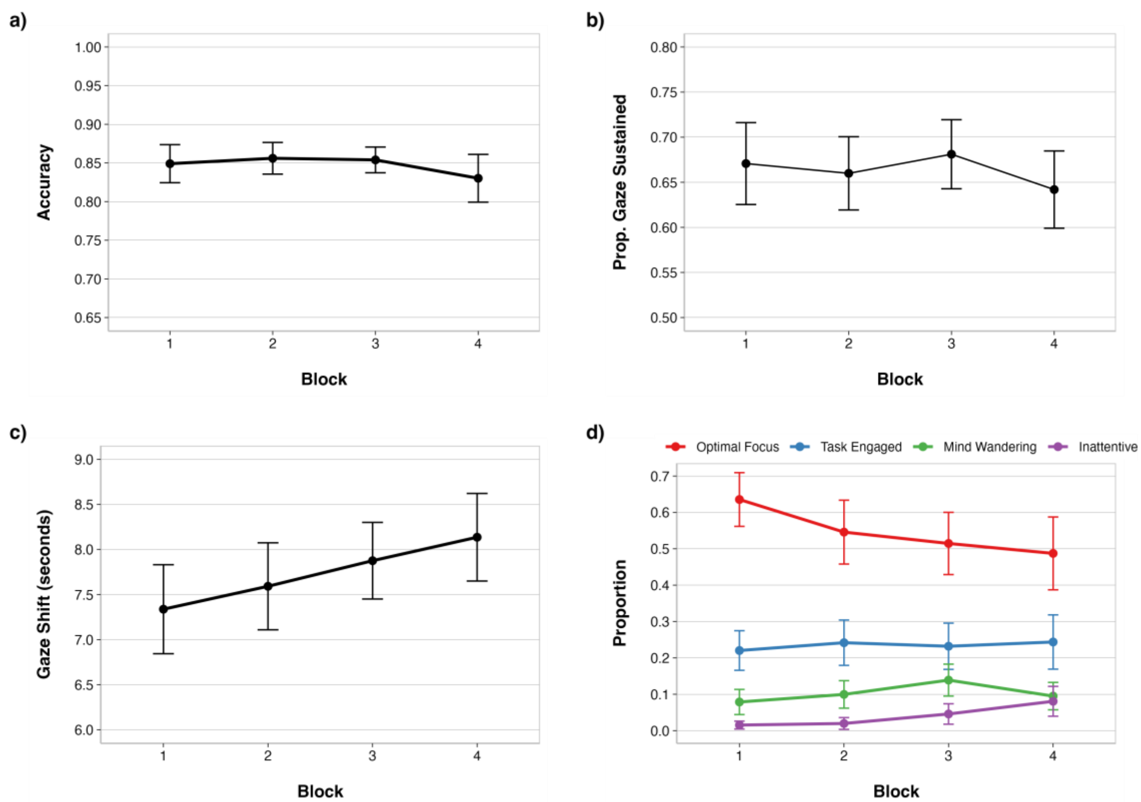

Figure 13 - Effect of Time-on-Task on the Waning of Attention

Note. Across four metrics a) accuracy, b) gaze sustained, c) gaze shift, and d) self-reported attentional state, there were no to small changes in the focus of attention as time-on-task increased. Error bars represent 95% confidence intervals.

Task engagement. If distraction becomes more likely over the wait time delay and as time-on-task increases, then there should be a decrease in self-reports of task engagement. Task engagement actually increased as wait time increased ($F(1.88, 104.98) = 14.25$, $p < .001$, $\eta_p^2 = .20$), see Figure 12c. Post-hoc comparisons showed that task engagement levels were higher at the long wait time compared to the short and medium wait times (**short-medium**: diff = -.01, Cohen's D = -.03, $t(56) = -0.52$, $p = .603$, **short-long**: diff = -.11, Cohen's D = -.35, $t(56) = -4.34$, $p < .001$, **medium-long**: diff = -.10, Cohen's D = -.31,

$t(56) = -4.501$, $p < .001$). This was likely due to a shift from a more optimal task focus to a less optimal focus but still engaged with the task. There was no main effect of time-on-task on task engagement ($F(2.30, 128.56) = 0.25$, $p = .808$, $\eta_p^2 < .01$), see Figure 13d. There was no wait time × time-on-task interaction on task engagement ($F(5.32, 298.10) = 0.67$, $p = .653$, $\eta_p^2 = .01$).

Mind wandering. If distraction becomes more likely over the wait time delay and as time-on-task increases, then there should be an increase in self-reports of mind wandering. Mind wandering increased as wait time increased ($F(1.96, 92.33) = 10.96$, $p < .001$, $\eta_p^2 = .19$), see Figure 12c. Post-hoc comparisons showed that mind wandering was less frequently reported on short wait times compared to medium (**short-medium**: diff = -.06, Cohen's D = -.33, $t(47) = -3.51$, $p = .002$) and long (**short-long**: diff = -.08, Cohen's D = -.39, $t(47) = -4.69$, $p < .001$), but that there was no difference between medium and long wait times (**medium-long**: diff = -.01, Cohen's D = -.06, $t(47) = -0.69$, $p = .492$). There was no main effect of time-on-task on mind wandering ($F(2.86, 134.54) = 2.61$, $p = .057$, $\eta_p^2 = .05$), see Figure 13d. There was only a marginally significant wait time × time-on-task interaction on mind wandering ($F(4.67, 219.41) = 2.31$, $p = .049$, $\eta_p^2 = .05$).

In summary, attention primarily fluctuated on a shorter timescale that led to a decrease in accuracy, sustained eye gaze at the cued location, less self-reported levels of optimal focus, and more self-reported mind wandering as wait time increased. Interestingly, reports of task engagement increased at the longest wait time compared to the short and medium wait times. This was likely due to a shift of attention from a more optimal task focus to a less optimal but still task engaged focus. The only change in attention with time-on-task was that fewer reports of optimal task focus were made on blocks 2-4 compared to block 1.

3.3.2 Are failures of sustained attention due to fluctuations of arousal modulated by the locus coeruleus?

Arousal and mindlessness models of sustained attention predict that arousal should decrease as wait time and time-on-task increases. Task utility models do not make any specific predictions; however, if changes in task utility are due to high tonic locus coeruleus activity (indicative of a reward-explorative mode) then, according to adaptive gain theory, arousal is expected to increase as wait time and time-on-task increases. For this hypothesis, analyses were conducted only on the no-reward group because this hypothesis is not concerned with the reward manipulation and the no-reward group served as the baseline.

Inattentiveness. Overall reports of inattentiveness were low, only 56% of subjects made at least 1 inattentiveness report. Nevertheless, there was an increase in inattentiveness reports as time-on-task increased ($F(2.511, 140.614) = 6.820$, $p = .001$, $\eta_p^2 = .11$), see Figure 13d. Post-hoc comparisons showed that inattentiveness reports were higher on block 4 than block 1 (**block 1v4**: diff = -.10, Cohen's D = -.52, $t(56) = -3.61$, $p = .004$). There was no difference in inattentiveness reports between wait times ($F(1.674, 93.723) = 1.03$, $p = .351$, $\eta_p^2 = .02$) and no wait time × time-on-task interaction ($F(4.490, 251.415) = 0.55$, $p = .722$, $\eta_p^2 = .01$).

Self-reported arousal. Self-reported arousal ratings decreased as wait time got longer ($F(1.970, 108.333) = 13.67$, $p < .001$, $\eta_p^2 = .20$), see Figure 14a. Post-hoc comparisons showed that all pairwise comparison were significantly different (**short-medium**: diff = .12, Cohen's D = .06, $t(55) = 2.10$, $p = .041$, **short-long**: diff = .27, Cohen's D = .13, $t(55) = 5.48$, $p < .001$, **medium-long**: diff = .16, Cohen's D = .08, $t(55) = 3.03$, $p = .008$). Self-reported arousal rating decreased with time-on-task ($F(1.766, 97.143) = 37.86$, $p < .001$, $\eta_p^2 = .41$), see Figure 14b. Post-hoc comparisons showed that all pairwise comparisons between blocks were significant (**block 1v2**: diff = 0.83, Cohen's D = .40, $t(55) = 4.48$, $p < .001$, **block 1v3**; diff = 1.58, Cohen's D = .77, $t(55) = 6.72$, $p < .001$, **block 1v4**; diff = 2.02, Cohen's D = .98, $t(55) = 7.35$, $p < .001$, **block 2v3**: diff = 0.75, Cohen's D = .37, $t(55) = 4.88$, $p < .001$, **block 2v4**: diff = 1.19, Cohen's D = .58, $t(55) = 5.43$, $p < .001$, **block 3v4**: diff = .44, Cohen's D = .22, $t(55) = 3.94$, $p < .001$). There was no wait time × time-on-task interaction ($F(5.309, 292.004) = 0.264$, $p = .940$, $\eta_p^2 = .01$).

Pre-trial baseline arousal. Pretrial pupil size, heart rate, and galvanic skin response provided physiological measures of changes in arousal with time-on-task. Consistent with the self-reported arousal ratings, pre-trial pupil size ($F(1.857, 94.73) = 5.66$, $p = .006$, $\eta_p^2 = .10$) and pre-trial heart rate ($F(2.041, 93.869) = 5.76$, $p = .004$, $\eta_p^2 = .11$) both decreased with time-on-task, see Figure 14c and Figure 14d. There were no changes in pre-trial galvanic skin response ($F(1.263, 58.104) = 1.26$, $p = .277$, $\eta_p^2 = .03$), see Figure 14e. Therefore, self-reported arousal ratings, pre-trial pupil size, and pre-trial heart rate, were more consistent with arousal and mindlessness models of sustained attention.

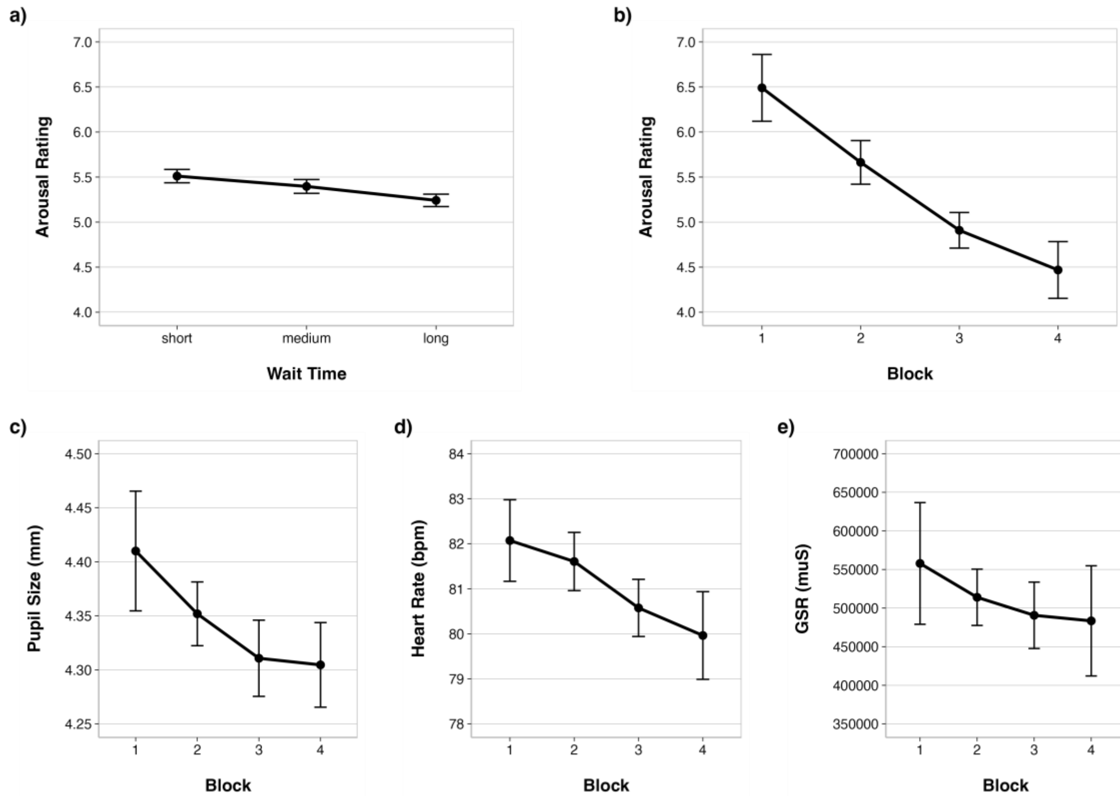

Figure 14 - Changes in Arousal with Wait Time and Time-on-Task

Note. Arousal decreased as wait time and time-on-task increased. a) self-reported arousal ratings decreased with wait time and time-on-task. b) pre-trial pupil size and c) pre-trial heart rate decreased with time-on-task. There were no changes in d) pre-trial galvanic skin response (GSR) with time-on-task. Error bars represent 95% confidence intervals.

Pupil size over the wait time. Overall, pupil size decreased over the wait time delay. The decrease in pupil size was larger as wait time got longer ($F(1.789, 87.664) = 5.112$, $p = .010$, $\eta_p^2 = .09$) and as time-on-task increased ($F(2.551, 124.978) = 3.601$, $p = .021$, $\eta_p^2 = .07$). However, there was also a wait time × time-on-task interaction ($F(4.494, 220.196) = 2.943$, $p = .017$, $\eta_p^2 = .06$), see Figure 15a. Interaction contrasts showed that there was no effect of wait time on the decrease in pupil size on block 1 but there was an effect on block 4 (**short-medium block 1v4**: $t(49) = -3.39$, $p = .001$; **short-long block 1v4**: $t(49) = -3.28$, $p = .002$). That is, early in the task attention was able to be sustained over long wait times—as measured by pupil size—but as time-on-task increased, attention decreased the longer attention had to be sustained over the wait time.

Pupil dilation. Pupil dilation to the onset of the get ready signal and to the onset of the target stimulus was analyzed. Because the get ready signal precedes the wait time, only the effect of time-on-task was analyzed for pupil dilation to the get ready signal. There was no effect of time-on-task on pupil dilation to the get ready signal ($F(2.377, 30.905) = 0.54$, $p = .617$, $\eta_p^2 = .04$), see Figure 15b. There was a main effect of time-on-task on pupil dilation to the target stimulus ($F(2.775, 135.991) = 5.07$, $p = .003$, $\eta_p^2 = .09$), no main effect of wait time ($F(1.493, 73.141) = 2.23$, $p = .128$, $\eta_p^2 = .04$) but there was a wait time $\times$ block interaction ($F(4.873, 238.756) = 2.99$, $p = .013$, $\eta_p^2 = .06$), see Figure 15c. Interaction contrasts showed that pupil dilation to the target decreased over blocks, but only for the short wait time (**block 1v4 short-medium**: $t(49) = 2.08$, $p = .043$; **block 1v4 short-long**: $t(49) = 3.30$, $p = .002$).

In summary, there was a decrease in arousal ratings, baseline pupil size, and heart rate—suggesting that arousal decreases the longer attention must be sustained. Pupil size also decreased over the wait time delay and this effect was more prominent as time-on-task increased suggesting it was harder to sustain a certain level of focus over the duration of the task. There was also a decrease in pupil dilations to the target in the short wait time condition as time-on-task increased. If the locus coeruleus-norepinephrine system modulates arousal and attention on sustained attention tasks, then this pattern of findings suggests that the waning of attention over time is due to a shift toward low tonic and low phasic locus coeruleus activity (as would be predicted by arousal and mindlessness models) and not a shift toward high tonic and low phasic activity (as would be predicted by adaptive gain theory of locus coeruleus function).

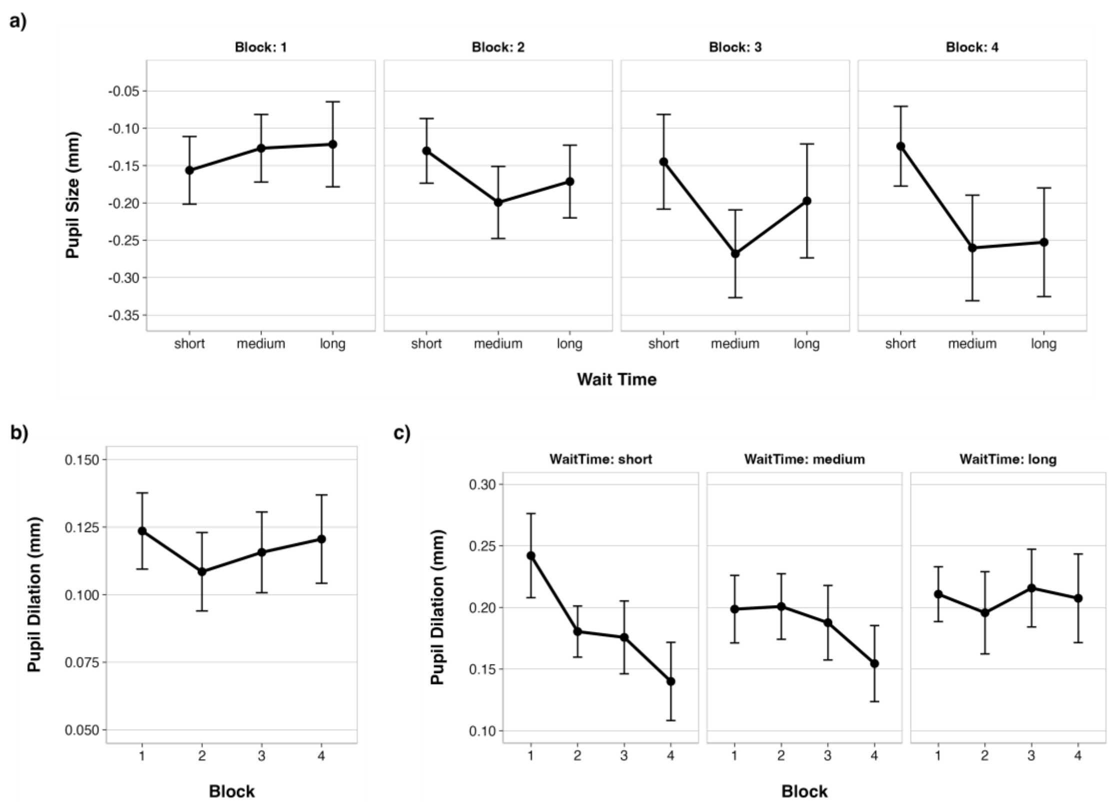

Figure 15 - Changes in Pupillary Response with Wait Time and Time-on-Task

Note. a) Pupil size decreased over the wait time delay. The decrease in pupil size did not change as wait time increased on block 1, but on block 4 the decrease in pupil size was larger at medium and long wait times—suggesting it was harder to sustain a certain level of focus over the wait time as time-on-task increased. b) There was no change in pupil dilation to the get ready signal. c) Pupil dilation to the target stimulus decreased with time-on-task but only for the short wait time. Error bars represent 95% confidence intervals.

3.3.3 Are failures of sustained attention due to depletion of resources or due to diminishing task utility?

Resource depletion and task utility models make opposite predictions as to the effect of reward on sustained attention. Resource depletion model predict that reward, and thereby greater effort and resource consumption, should lead to greater declines in sustained attention. Task utility models, on the other hand, predict that reward should lead to smaller declines in sustained attention. Reward × time-on-task and reward × wait time interactions were analyzed to test the effect of reward on sustained attention at longer and shorter timescales[3].

Accuracy. Overall, the reward group performed better than the no-reward group (F(1.000, 99.000) = 8.324, p = .005, η_p^2 = .08). There was no reward × wait time interaction on accuracy (F(1.562, 154.617) = 2.623, p = .089, η_p^2 = .03) and the presence of a main effect of wait time demonstrated that both groups showed the same decline in performance as wait time increased (F(1.562, 154.617) = 90.38, p < .001, η_p^2 = .48), see Figure 16a. There was a reward × time-on-task interaction (F(1.985, 196.494) = 3.280, p = .040, η_p^2 = .03), see Figure 16b. Interaction contrasts showed that the reward group had a relative increase in performance as time-on-task increased compared to the no-reward group (**no-reward vs reward at block 1v2**: t(99) = -2.02, p = .047; **no-reward vs reward at block 1v3**: t(99) = -2.06, p = 042; **no-reward vs reward at block 1v4**: t(99) = -2.28, p = .025).

[3] Note that the 3-way interaction of reward × wait time × time-on-task was also analyzed but was statistically significant for only one dependent measure, mind wandering, and no specific predictions were made as to the presence of the 3-way interaction.

Eye gaze. Looking at the effects on gaze sustained, there was no main effect of reward ($F(1.000, 95.000) = 2.81$, $p = .097$, $\eta_p^2 = .03$) no reward × wait time interaction ($F(1.593, 151.362) = 1.177$, $p = .303$, $\eta_p^2 = .01$), see Figure 16c, nor a reward × time-on-task interaction ($F(1.813, 172.250) = 0.994$, $p = .365$, $\eta_p^2 = .01$), see Figure 16d. The presence of a main effect of wait time means that both reward groups showed the same decline in sustaining eye gaze at the cued location as wait time increased ($F(1.593, 151.362) = 63.05$, $p < .001$, $\eta_p^2 = .40$), see Figure 16c. Looking at the effects on time for gaze to shift, there was no main effect reward ($F(1.000, 109.000) = 1.47$, $p = .228$, $\eta_p^2 = .01$) and no reward × time-on-task interaction ($F(2.42, 263.779) = 0.58$, $p = .595$, $\eta_p^2 = .01$), see Figure 16e.

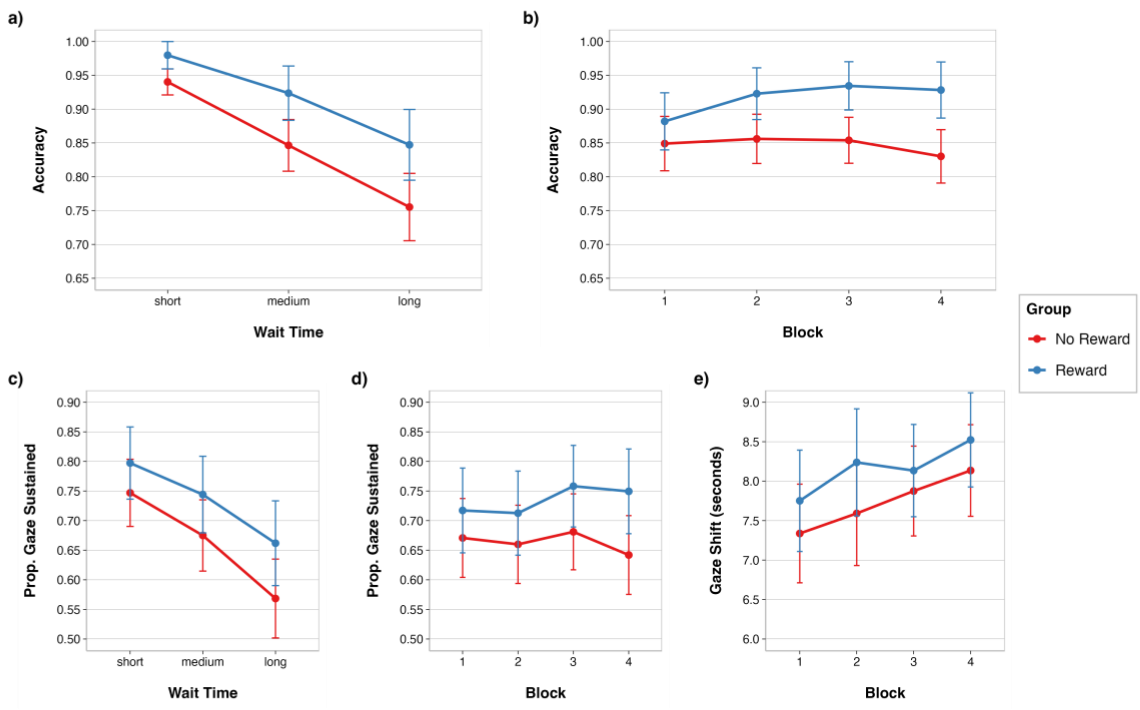

Figure 16 - Effect of Reward on the Focus of Attention

Note. Reward did not lead to improvements in accuracy with a) wait time but it did with b) time-on-task. Reward let to overall improvements on c-d) sustaining eye gaze over the wait time but not in e) how long it took for eye gaze to shift away from the cued location. Error bars represent 95% confidence intervals.

Self-report thought probes. Looking at the effects on optimal task focus, there was no main effect of reward ($F(1.000, 109.000) = 2.59$, $p = .111$, $\eta_p^2 = .02$), no reward × wait time interaction ($F(1.968, 214.475) = 0.63$, $p = .534$, $\eta_p^2 = .01$), see Figure 17a, nor a reward × time-on-task interaction ($F(2.392, 260.743) = 1.43$, $p = .240$, $\eta_p^2 = .01$), see Figure 18a. The presence of a main effect of wait time ($F(1.968, 214.475) = 90.19$, $p < .001$, $\eta_p^2 = .45$) and time-on-task ($F(2.392, 260.743) = 8.18$, $p < .001$, $\eta_p^2 = .07$) showed that both groups had fewer reports of optimal task focus as wait time and time-on-task increased, see Figure 17a and Figure 18a.

For task engagement, there was no main effect of reward (F(1.000, 109.000) = 0.66, p = .418, η_p^2 = .01). There was a reward × wait time interaction (F(1.916, 208.871) = 5.18, p = .007, η_p^2 = .05), see Figure 17b. Interaction contrasts showed that the reward group had a larger increase in task engagement with wait time compared to the no-reward group (**no-reward vs reward at short-medium**: t(109) = -3.04, p = .003; **no-reward vs reward at short-long**: t(109) = -2.48, p = .015). The increase in task engagement over wait times reflects a shift from an optimal task focus to a less focused but still task engaged state. There was no reward × time-on-task interaction on task engagement (F(2.769, 301.834) = 2.05, p = .112, η_p^2 = .02), see Figure 18b.

For mind wandering, those in the reward group reported less mind wandering than the no-reward group (F(1.000, 92.000) = 13.36, p < .001, η_p^2 = .13). There was a reward × wait time interaction on mind wandering (F(1.959, 180.242) = 4.578, p = .012, η_p^2 = .05), see Figure 17c. Interaction contrasts showed that the no-reward group had a larger increase in mind wandering as wait time increased (**no-reward vs reward at short-medium**: t(92) = 2.49, p = .015; **no-reward vs reward at short-long**: t(92) = 2.92, p = .004). There was no reward × time-on-task interaction on mind wandering (F(2.912, 267.932) = 1.74, p = .161, η_p^2 = .02), see Figure 18c.

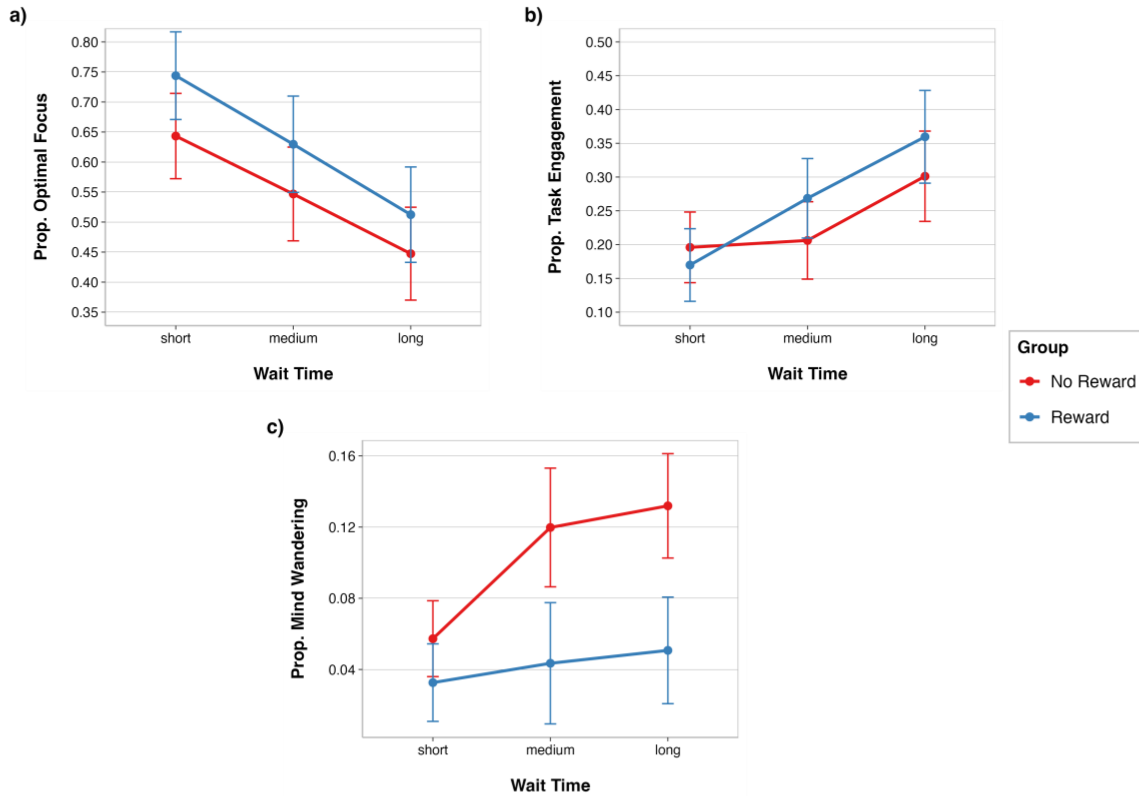

Figure 17 - Effect of Reward and Wait Time on Self-Reported Attentional State

Note. a) There were no differences between reward groups on the proportion of optimal task focus. b) The reward group showed a larger increase in task engagement as wait time increased. c) The no-reward group showed a larger increase in mind wandering as wait time increased. Error bars represent 95% confidence intervals.

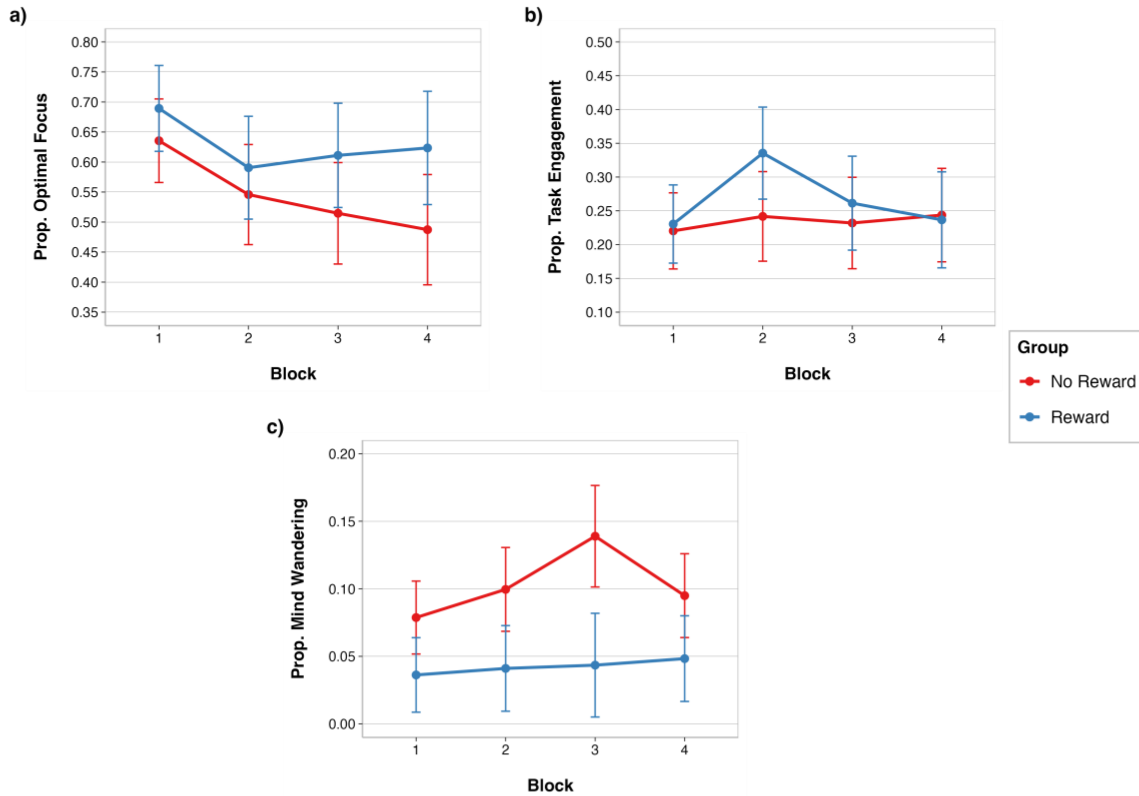

Figure 18 - Effect of Reward and Time-on-Task on Self-Reported Attentional State

Note. There were no differences due to reward on the proportion of a) optimal task focus, b) task engagement, and c) mind wandering across blocks. Error bars represent 95% confidence intervals.

There were very few reports of intentional mind wandering. Only 13.5% of subjects in the reward group reported at least one instance of intentional mind wandering. Therefore, instead of conducting an ANOVA on intentional mind wandering rates, I conducted a chi-square test on the number of subjects that did and did-not report at least one instance of intentional mind wandering in the reward and no-reward groups. The chi-square test was significant ($\chi(1)^2$ = 6.221, p = .013), there were more subjects in the no-reward group

(35%) that reported at least one instance of intentional mind wandering compared to the reward group (13.5%).

For inattentiveness, there were also very few inattentiveness reports, particularly in the reward group. Only 24% of subjects in the reward group reported at least one instance of inattentiveness. Therefore, instead of conducting an ANOVA on inattentiveness rates, I conducted a chi-square test on the number of subjects that did and did-not report at least on instance of inattentiveness in the reward and no-reward groups. The chi-square test was significant (($\chi(1)^2 = 10.536$, $p = .001$), there were more subjects in the no-reward group (56%) that reported at least on instance of inattentiveness compared to the reward group (24%).

In summary, reward led to an overall improvement in performance, but both the reward groups showed a similar decrease in accuracy as wait time increased. There was a small difference of performance as time-on-task increased with the reward group showing a relative superiority in performance compared to the no-reward group. Reward did not improve sustaining eye gaze at the cued location over the wait time delay. Both reward groups showed the same decline in reported optimal task focus as wait time increased. However, that decline seemed to be for different reasons. The reward group showed a greater shift towards a less optimal but still task engaged focused while the no-reward group showed a greater shift towards mind wandering.

Self-reported arousal. Those in the reward group reported higher levels of arousal overall ($F(1.000, 108.000) = 8.59$, $p = .004$, $\eta_p^2 = .07$). There was no reward × wait time interaction on self-reported arousal ($F(1.922, 207.597) = 0.140$, $p = ..862$, $\eta_p^2 < .01$), see

Figure 19a. There was a reward × time-on-task interaction (F(1.755, 189.590) = 9.46, p < .001, η_p^2 = .08), see Figure 19b. Interaction contrasts showed that the reward group had less of a decline in arousal as time-on-task increased (**no-reward vs reward at block 1v2**: t(108) = -1.29, p = .201; **no-reward vs reward at block 1v3**: t(108) = -3.07, p = .003; **no-reward vs reward at block 1v4**: t(108) = -3.48 p = .001).

Pre-trial baseline arousal. There was no main effect of reward on pre-trial pupil size (F(1.000, 95.000) = 0.24, p = .627, η_p^2 < .01) pre-trial heart rate (F(1.000, 86.000) = 0.02, p = .904, η_p^2 < .01) or pre-trial galvanic skin response (F(1.000, 86.000) = 2.21, p = .141, η_p^2 = .03). There was no reward × time-on-task interaction on pre-trial pupil size (F(1.938, 184.147) = 0.27, p = .757, η_p^2 < .01) pre-trial heart rate (F(2.19, 188.375) = 0.62, p = .551, η_p^2 = .01) or pre-trial galvanic skin response (F(1.959, 168.458) = 0.81, p = .446, η_p^2 = .01), see Figure 19c-e.

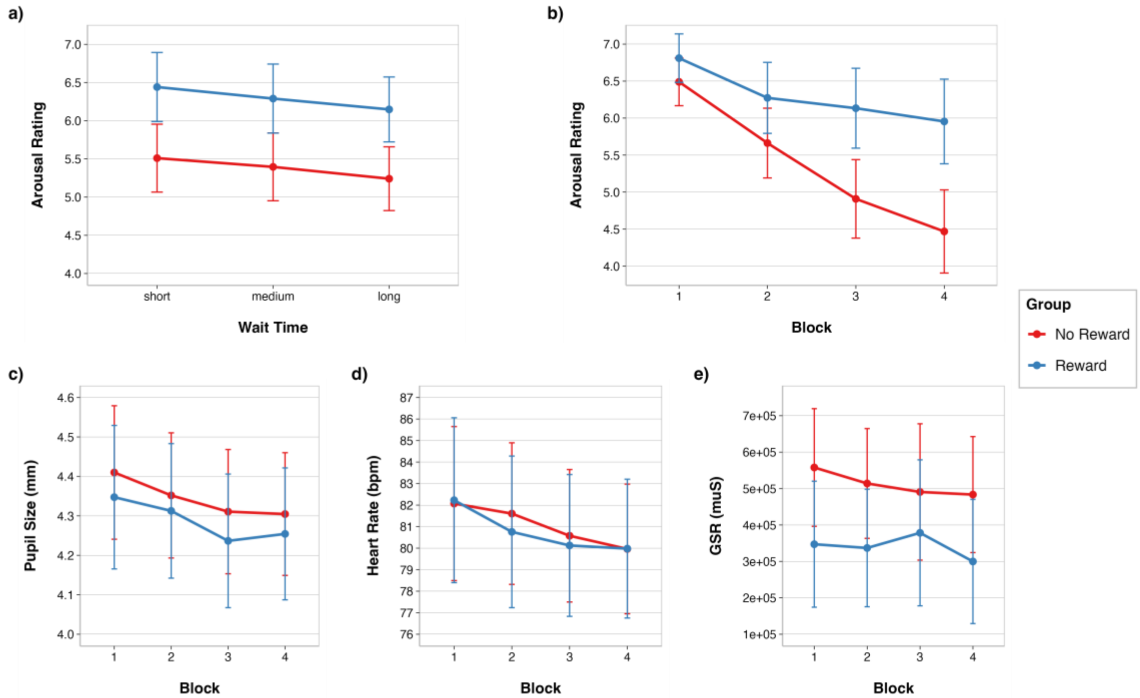

Figure 19 - Effect of Reward on Arousal

Note. Reward increased overall levels of self-reported arousal ratings. There were no difference in the change in arousal as a) wait time increased for the reward and no-reward groups, but b) reward did lead to less of a decrease in arousal as time-on-task increased. There were no differences due to reward on c) pre-trial baseline pupil size, d) pre-trial heart rate, and e) pre-trial galvanic skin response (GSR). Error bars represent 95% confidence intervals.

Pupil size over the wait time. In the analyses on the no-reward group reported above, there was a wait time × time-on-task interaction such that there was no effect of wait time on the decrease in pupil size on block 1 but there was an effect on block 4 (Figure 15a). There were no two-way (reward × wait time: $F(1.674, 152.352) = 1.00$, $p = .358$, $\eta_p^2 = .01$; reward × time-on-task: $F(2.578, 234.583) = 0.39$, $p = .727$, $\eta_p^2 < .01$) or three-way (reward × wait time × time-on-task: $F(5.013, 456.159) = 0.59$, $p = .710$, $\eta_p^2 = .01$) interactions with reward, therefore both groups showed the same wait time × time-on-task interaction, see Figure 20a. There was, however, a main effect of reward such that those in the reward group showed less of a decline in pupil size over the wait time overall (F(1.000, 91.000) = 4.73, p = .032, η_p^2 = .05).

Pupil dilation. There was no main effect of reward on pupil dilation to the get ready signal (F(1.000, 89.000) = 0.15, p = .700, η_p^2 < .01), nor to the target stimulus (F(1.000, 91.000) = 3.01, p = .086, η_p^2 = .03). There was no reward × time-on-task interaction on pupil dilation to the get ready signal (F(2.72, 242.065) = 0.14, p = .923, η_p^2 < .01), see Figure 20b, nor to the target stimulus (F(2.747, 250.006) = 1.13, p = .335, η_p^2 = .01), see Figure 20d. There was no reward × wait time interaction on pupil dilation to the target stimulus (F(1.574, 143.233) = 0.363, p = .645, η_p^2 < .01), see Figure 20c. In the analyses reported on the no-reward group above, there was a wait time × time-on-task interaction with pupil dilation to the target stimulus Figure 15d. The lack of an interaction with reward means that both the reward and no-reward group showed the same wait time × time-on-task interaction—such that pupil dilation to the target decreased over blocks, but only for the short wait time.

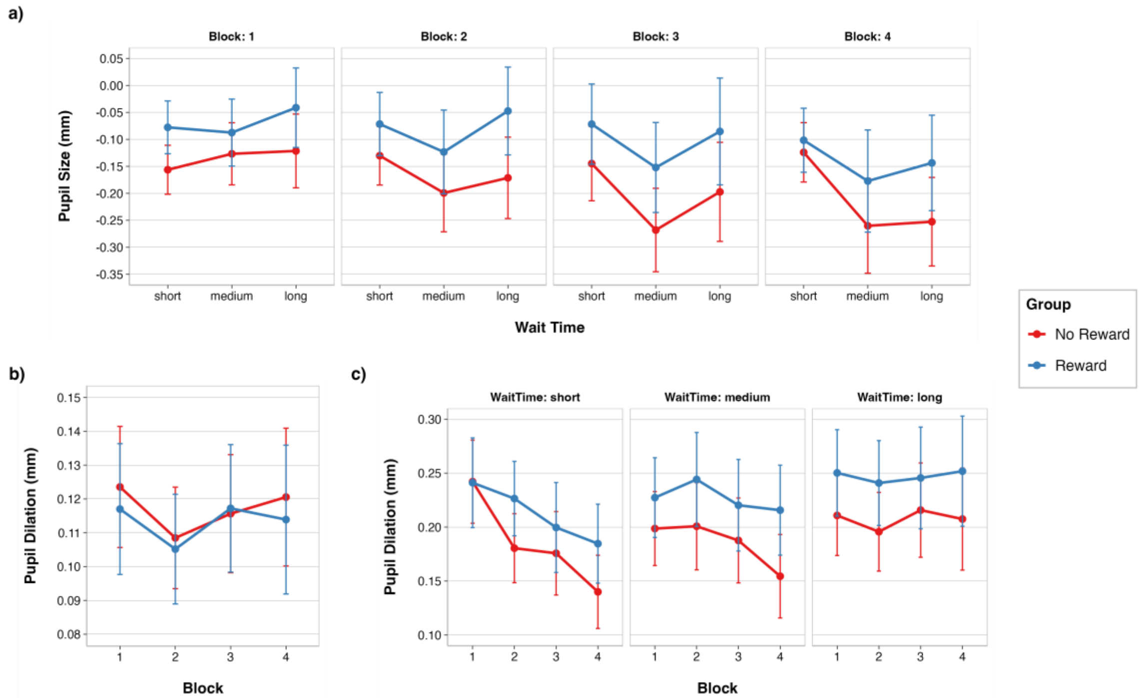

Figure 20 - Effect of Reward on the Pupillary Response

Note. a) The reward group showed less of a decline in pupil size over the wait time delay overall. There were no differences due to reward on b) pupil dilation to the get ready signal, nor to d) pupil dilation to the target stimulus. Error bars represent 95% confidence intervals.

In summary, based on self-report arousal ratings the reward group was able to maintain their level of arousal as time-on-task increased. However, there were no differences in physiological measures of arousal between the reward groups. The reward group showed an overall smaller decrease in pupil size over the wait time delay suggesting they were better able to sustain their attention. There was no difference due to reward on pupil dilation to the target stimulus.

3.3.4 Post-Task Questionnaires

Subjects rated their level of motivation, task difficulty, and task unpleasantness after completing the task. Those in the reward group (M = 4.02, SD = .86) reported higher levels of motivation than those in the no-reward group (M = 3.39, SD = .84; diff = .63, $t(109) = 3.93$, $p < .001$). There were no differences between the reward group (M = 2.39, SD = .81) and the no-reward group (M = 2.56, SD = .87) on levels of difficulty (diff = .17, $t(109) = 1.08$, $p = .282$). There were no differences between the reward group (M = 2.57, SD = .1.14) and the no-reward group (M = 2.91, SD = 1.03) on levels of unpleasantness (diff = .34, $t(108) = 1.62$, $p = .108$).

Motivation. The purpose of the reward manipulation was to increase the motivation for those in the reward group to perform well and engage more effort and attention to the task. However, even though the reward group reported higher levels of motivation the difference is not large and there is considerable overlap between the two groups on motivation ratings (Table 3). On average, the no-reward group reported being somewhere between moderately and highly motivated, whereas the no-reward group on average reported being highly motivated—this is not a large difference in motivation. This was largely due to the no-reward group reporting high levels of motivation—45.6% of subjects in the no-reward group reported being highly or extremely motivated. There was also a subset of those in the reward group reporting moderate to low levels of motivation—27.8% of subjects in the reward group reported being slightly or moderately motivated. Therefore, I conducted the same analyses above except replaced reward group with self-reported motivation to test the effect of motivation on sustained attention.

Table 3 - Motivation Ratings for the No-Reward and Reward Groups

Motivation Rating	No-Reward (*N* = 57)	Reward (*N* = 54)
5 – Extremely motivated	7.0%	33.3%
4 – Highly motivated	38.6%	38.9%
3 – Moderately motivated	42.1%	24.1%
2 – Slightly motivated	10.5%	3.7%
1 – Not motivated at all	1.8%	0.0%

Motivation was grouped into high motivation (highly and extremely motivated; n = 65) and low motivation (not at all, slightly, and moderately motivated; n = 46). A 2 (Motivation) × 4 (Block) × 3 (Wait Time) mixed ANOVA was conducted on the various dependent measures. For the most part, the only differences were on accuracy and gaze sustained.

Accuracy. The main effect of motivation on accuracy ($F(1.000, 98.000) = 33.07$, $p < .001$, $\eta_p^2 = .25$) was considerably larger than when reward group was used ($\eta_p^2 = .08$). Unlike when using reward group, there was no motivation × time-on-task interaction ($F(1.991, 195.077) = 0.146$, $p = .863$, $\eta_p^2 < .01$). There was a motivation × wait time interaction on accuracy ($F(1.632, 159.914) = 13.09$, $p < .001$, $\eta_p^2 = .19$), see Figure 21a, this interaction was not significant when using reward group as the moderator. Those that reported higher levels of motivation showed less of a decline in accuracy as wait time got longer (**low vs high motivation at short-medium**: $t(98) = -4.45$, $p < .001$; **low vs high motivation at short-long**: $t(98) = -4.04$, $p < .001$).

Eye gaze. Unlike when using reward group, there was a motivation × wait time interaction on gaze sustained ($F(1.613, 153.240) = 4.55$, $p = .018$, $\eta_p^2 = .05$), see Figure 21b. Interaction contrasts showed that higher motivation was associated with less of a decline in sustaining eye gaze on the cued location as wait time increased, with the biggest difference being at long wait times (**low vs high motivation**: $t(95) = -2.44$, $p = .017$).

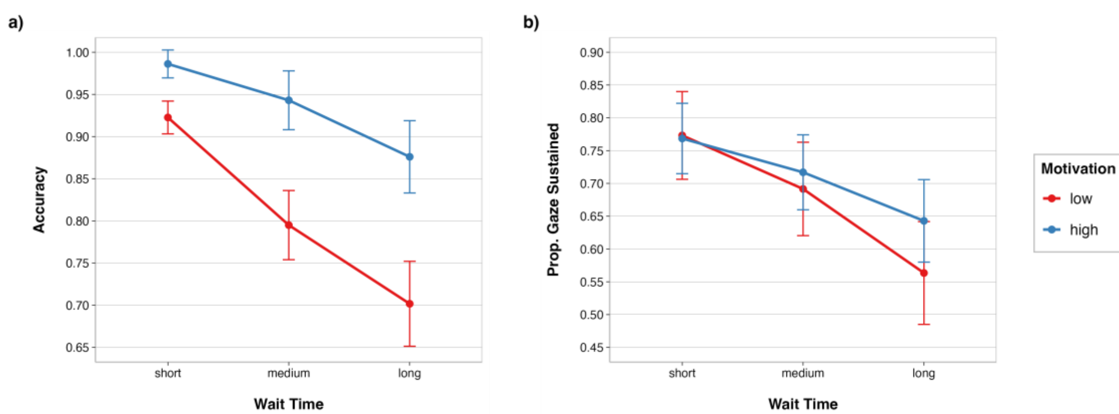

Figure 21 - Effect of Motivation on Sustained Attention

Note. Those that reported a higher motivation to perform well on the task showed a) less of a decline in accuracy as wait time increased and b) less of a decline in sustaining eye gaze at the cued location as wait time increased. High motivation: highly motivated or extremely motivated. Low motivation: moderately motivated, slightly motivated, or not motivated at all. Error bars represent 95% confidence intervals.

3.4 Discussion

This study investigated how various factors are related to the fluctuation and waning of attention over shorter and longer timescales. Multiple indices of attention and arousal were measured as subjects performed the sustained attention-to-cue task. On a shorter timescale, the duration that attention had to be sustained varied from short (0 – 3.5 seconds), to medium (6 – 8.5 seconds), and to long (11 – 13.5 seconds) durations. On a longer timescale, performance was assessed over about 42 minutes of task performance broken up into 4 blocks of trials. In addition, some subjects did the task with a performance-based reward while others did it with no-reward. The purpose of the reward manipulation was to test how changes in task utility and motivation impact sustained attention performance.

Based on behavioral performance, eye gaze, and self-report thought probes, attention primarily waned over a shorter timescale—getting worse from shorter to longer durations (Figure 12). There were no significant changes in attention across the entire duration of the task (Figure 13). Self-report thought probes revealed that as the duration that attention had to be sustained increased, attention tended to shift from a more optimal task focused state to a less optimal but still task engaged state and to some extent off-task states of mind wandering. These findings suggest that attention can fluctuate and wane on a relatively short timescale of around 10 seconds or less. The lack of a waning of attention over the duration of the entire task suggests that the focus of attention can be restored on each trial, but then temporarily wanes the longer attention must be sustained on that trial.

Overall, there was strong evidence that arousal decreased the longer attention had to be sustained at both shorter and longer timescales (Figure 14). Self-reported arousal ratings decreased both as wait time durations increased and as time-on-task increased. Arousal rating went from being close to “7 - Alert” (average of 6.5 on block 1) to somewhere between “3 – Sleepy, but no difficulty remaining awake” to “5 – Neither Alert nor Sleepy” (average of 4.5 on block 4). Additionally, baseline pre-trial pupil size and heart rate decreased over the duration of the task.

It is interesting to note that although self-reported arousal decreased with time-on-task, the indices of attention—accuracy, eye gaze, and self-reported mind wandering—did not. This can suggest several different things. One possibility is that arousal fluctuates on a larger timescale and cannot be restored as quickly as the focus of attention can be. If that is the case, this could also suggest that even as arousal decreases, one can still restore or re-initiate the focus of attention temporarily. There was some evidence that this was the case as pupil dilations to the ‘get ready’ signal did not decline over the course of the task, indicating a similar level of focusing attention at the start of the wait time delay. However, there was some evidence that temporarily sustaining that focus of attention was less successful over the duration of the entire task. For instance, the decrease in pupil size over the wait time delay on medium and long wait times was larger by the third and fourth block of trials compared to the first block. That is, initially pupil size over the wait time delay was maintained at a certain level, even at long wait times, but was not able to be maintained on medium or long wait times towards the end of the task.

Overall, reward had a modest impact on sustained attention. Those in the reward group performed better overall on the task but both groups showed a similar decline in performance as wait time got longer (Figure 16). However, this was partially because nearly half of the subjects in the no-reward group were reporting high to extremely high motivation on the task (Table 3). When separating groups based on motivation levels, rather than reward group, there was an interaction between motivation and wait time such that those with high motivation showed less of a decline in performance as wait time got longer—that is, high motivation improved sustained attention (Figure 21). This finding suggests that the waning of attention over time is partly due to a lack of motivation to stay engaged with the task. As such, there was also an interaction between reward group (also motivation) and wait time on mind wandering; those in the no-reward group/low motivation showed an increase in mind wandering reports as wait time got longer (Figure 17). Therefore, this provides evidence that the waning of attention over time is related to a diminishing task utility and a resulting shift of attention to task-unrelated concerns (e.g., mind wandering)—supporting task utility models of sustained attention. The other major impact that reward/motivation had was on self-report arousal ratings; those in the reward group showed less of a decline in arousal ratings over the duration of the task Figure 19. This suggests the possibility that motivation moderated levels of arousal to keep task utility high and improve sustained attention.

3.4.1 Resource Depletion or Diminishing Task Utility?

A major motivating factor for the reward manipulation in the current study was to distinguish between resource depletion (Grier et al., 2003; Warm et al., 1996) and task utility models (Esterman et al., 2014, 2016; Kurzban et al., 2013) of sustained attention (Pattyn et al., 2008). In general, the results do not make a strong case in favor of one model over the other. One major reason for this was that there were no time-on-task effects on sustained attention performance in the study. The stronger prediction made by the two models is on the nature of the vigilance decrement, that is, on time-on-task effects. Therefore, the lack of time-on-task effects in the present study is a major limitation for drawing any strong conclusions in regard to these models.

With that said, the results with motivation as a moderator of accuracy and mind wandering on the sustained attention-to-cue task is more consistent with task utility models. When a task has greater utility sustained attention performance, at least on a shorter timescale, is improved and there are fewer shifts of attention away from the task and towards other concerns.

Another difficulty in reconciling differences between these models is the vagueness in resource depletion models as to the nature of the resources that are being depleted. What are those resources and how do they relate to arousal, attention, and executive control? Some attempts have been made to be more specific about these relationships; for instance, the resource-control model of mind wandering (Thomson et al., 2015) specifies that it is not resources per-se that is depleted over time but the capacity to engage in executive control processes. This depleted executive control capacity, then, results in more frequent lapses of attention and mind wandering. Therefore, the findings in the current study can be consistent with either task utility models or a resource-control model. These models are not necessarily incompatible, but they do provide different explanations as to *why* attention wanes over time.

Models of sustained attention are often characterized more broadly as either overload or underload accounts. Overload accounts assume that sustained attention tasks are cognitively over-demanding such that they lead to a depletion of resources and performance decrements over time (Caggiano & Parasuraman, 2004; Grier et al., 2003). Underload accounts assume that sustained attention tasks are cognitively under-demanding such that they lead to boredom, under-arousal, lack of motivation, and mindlessness (Manly et al., 1999; Robertson et al., 1997). The relationship between arousal and motivation/utility in underload accounts, however, are not well specified.

The adaptive gain theory of locus coeruleus function does make a direct association between task utility, arousal, and locus coeruleus activity (Aston-Jones & Cohen, 2005). When task utility is high, arousal and tonic locus coeruleus activity are moderate and phasic locus coeruleus activity is high. This mode of locus coeruleus activity, termed exploitative mode, is optimal for task performance to filter and enhance responding to only task-relevant events—that is, it is optimal for exploiting a current source of reward. When task utility is low, arousal and tonic locus coeruleus activity are high and phasic locus coeruleus activity is low. This mode of activity, termed explorative mode, is optimal for exploring alternative sources of reward and is associated with states of distractibility and mind wandering (Unsworth & Robison, 2016). In contrast to arousal and mindlessness accounts of sustained attention, this would suggest that decrements in sustained attention performance over time should be due to a shift towards an explorative mode of locus coeruleus activity and heightened arousal, not under-arousal. However, the findings from the current study suggest that arousal decreases and distractibility, by mind wandering, increases over time. Additionally, the decrease in pre-trial pupil size, an indicator of locus coeruleus activity, suggests a decrease in tonic locus coeruleus activity, not an increase. Therefore, these patterns of findings suggest that the waning of attention over time was due to a shift towards a low tonic mode of locus coeruleus activity, while at the same time an increased distractibility by mind wandering.

In general, the findings from the current study do suggest that less optimal states of arousal and attention are associated with declines in sustained attention performance. Additionally, reward/motivation led to a better regulation of more optimal levels of arousal—based on self-report arousal ratings—and attention—based on accuracy, eye gaze, and thought probes.

3.4.2 Limitations

In Study 1, there was a decline in performance across the duration of the sustained attention-to-cue task (at least on medium and long wait times). That was not the case in the current study, even though the duration of the task was more than twice as long. One possibility is that the task in Study 1 was administered towards the end of a 2-hour and 30-minute session of other cognitive demanding tasks, whereas this was the only task administered in the current study. Therefore, there were likely differences in attention and/or motivation at the start of the task between the two studies. This makes the comparison of time-on-task effects hard to compare.

The subjects in the current study were Georgia Tech students, who tend to be higher in cognitive ability, and probably abiding motivation, relative to our non-college sample in Study 1. It is possible that a different pattern of findings would be found with a different or more diverse sample.

The overall proportion of inattentiveness and intentional mind wandering reports were very low, with the majority of subjects making no such reports. This made it difficult to analyze differences on these measures. However, based on chi-square equivalence testing, it did seem that those in the reward group were less likely to make reports of being inattentive or engaging in intentional mind wandering.

CHAPTER 4. GENERAL DISCUSSION

The faculty of sustaining attention to on one thing over other things for a period of time is critical for carrying out almost any task as we manage distractions from the environment and internally generated thoughts, desires, and emotions. Amidst all this external and internal activity, it takes **effort** to *voluntarily sustain attention* on any one thing over the course of minutes or even seconds. In two studies, I investigated how various factors are related to sustaining attention over shorter and longer timescales.

In Study 1, I looked at the measurement of individual differences in sustained attention ability and how it relates to other cognitive abilities. I had previously developed a novel task, the sustained attention-to-cue task, to improve on the reliability and validity of measuring individual differences in attention control (Draheim et al., 2021). In the sustained attention-to-cue task, the critical element is to focus attention at a cued location on the display and sustain attention there for a variable amount of time. In the first version of the task, the cue remained on the screen the entire time and therefore it was possible for attention to move away and then re-orient back to the location (Draheim et al., 2021). As such, there were no differences in performance if attention had to be sustained for a short (2 seconds) or long (12 seconds) period of time. Therefore, even though the task was a reliable and valid measure of attention control, it was not capturing differences in sustaining attention over time. Study 1 was a follow-up to the first version of the sustained attention-to-cue task. A second version was developed in which the cue was removed from the screen during the wait time delay. This modification led to significant improvements in this task to capture differences in the ability to sustain attention over time. Performance declined as the amount of time attention had to be sustained increased. Additionally, that decline in performance was related to individual differences in attention control—those high on attention control ability showed less of a decline in performance, that is they showed a greater ability to sustain attention. Sustained attention, however, was not related to individual differences in working memory capacity or fluid intelligence. This finding may have important implications as to the relationship between sustained attention and higher-order cognitive abilities like working memory capacity. Additionally, the sustained attention-to-cue task represents a successful attempt to develop more reliable and valid measures of attention control and one that does not rely on reaction time, nor reaction time difference scores.

In Study 2, I looked at how susceptibility to internal distraction, changes in arousal, depletion of resources, and changes in task utility are related to the fluctuation and waning of attention over time. I measured indicators of attention and arousal using several metrics based on behavioral performance, eye gaze, pupil size, heart rate, and self-report mind wandering and arousal measures. A critical manipulation was providing a performance-based reward to one group of subjects and no-reward to another group of subjects. The reward manipulation provided a way to dissociate resource depletion and task utility models while also looking at how differences in motivation influence sustained attention. The sustained attention-to-cue task was used to assess sustained attention over shorter (0 – 13.5 seconds) and longer (~45 minutes) timescales. Overall, there was a large decline in attention on a shorter timescale based on performance, eye gaze, and mind wandering measures. There were no changes in attention at a longer timescale, however there was strong evidence that arousal declined over the course of the task. Reward and motivation lead to improvements in attention overall and motivation led to improvements in sustained attention at a shorter timescale.

In summary, I found that 1) the ability to control attention was related to better sustained attention, 2) working memory capacity and fluid intelligence were not related to sustained attention, 3) attention can wane at a relatively short timescale (seconds), 4) mind wandering increased as attention waned, 5) arousal decreased at a larger timescale (minutes), and 6) motivation led to improvements in sustained attention. These findings have important implications for how we understand individual differences in cognitive ability and whether sustained attention is an important factor for some abilities and not others. The results also point to a need to better specify the complex interactions of sustained attention, distractibility, arousal, and motivation at different timescales.

REFERENCES

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*, 29–37.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research Reviews*, *42*, 33–84. https://doi.org/10.1016/S0165-0173(03)00143-7

Burgoyne, A. P., Mashburn, C. A., Tsukahara, J. S., & Engle, R. W. (2022). Attention control and process overlap theory: Searching for cognitive processes underpinning the positive manifold. *Intelligence*, *91*, 101629. https://doi.org/10.1016/j.intell.2022.101629

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. *Psychonomic Bulletin & Review*, *11*(5), 932–937. https://doi.org/10.3758/BF03196724

Christoff, K., Gordon, A., & Smith, R. (2011). The role of spontaneous thought in human cognition. In *Neuroscience of decision making* (pp. 259–284). Psychology Press.

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Cohen, R. A. (2014). Focused and sustained attention. In R. A. Cohen, *The Neuropsychology of Attention* (pp. 89–112). Springer US. https://doi.org/10.1007/978-0-387-72639-7_6

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, Z. D., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786.

Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*, 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652–655. https://doi.org/10.3758/BF03200977

Doran, S. M., Van Dongen, H. P., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes de Biologie*, *139*(3), 253–267.

Dorrian, J., Rogers, N. L., & Dinges, D. F. (2004). Psychomotor vigilance performance: Neurocognitive assay sensitive to sleep loss. In C. A. Kushida (Ed.), *Sleep Deprivation* (1st ed., pp. 39–70). CRC Press. https://doi.org/10.3109/9780203998007-4

Draheim, C., Harrison, T. L., Embretson, S. E., & Engle, R. W. (2018). What item response theory can tell us about the complex span tasks. *Psychological Assessment*, *30*(1), 116–129. https://doi.org/10.1037/pas0000444

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*, 508–535.

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*, *150*(2), 242–275. https://doi.org/10.1037/xge0000783

Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. In *Princeton NJ Educational Testing Service* (Vol. 102). https://doi.org/10.1073/pnas.0506897102

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*(2), 190–193. https://doi.org/10.1177/1745691617720478

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 44, pp. 145–199). Elsevier Science.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and function of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102–134). Cambridge University Press.

Esterman, M., Grosso, M., Liu, G., Mitko, A., Morris, R., & DeGutis, J. (2016). Anticipation of monetary reward can attenuate the vigilance decrement. *PLOS ONE*, *11*(7), e0159741. https://doi.org/10.1371/journal.pone.0159741

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. https://doi.org/10.1093/cercor/bhs261

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), 2287–2295. https://doi.org/10.1037/xge0000019

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, *29*, 174–180. https://doi.org/10.1016/j.copsyc.2019.03.005

Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research: Sustained attention. *Annals of the New York Academy of Sciences*, *1396*(1), 70–91. https://doi.org/10.1111/nyas.13318

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, *102*(27), 9673–9678. https://doi.org/10.1073/pnas.0504136102

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

Gonthier, C. (2022). An easy way to improve scoring of memory span tasks: The edit distance, beyond "correct recall in the correct serial position." *Behavior Research Methods*, 16. https://doi.org/10.3758/s13428-022-01908-2

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: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 349–359. https://doi.org/10.1518/hfes.45.3.349.27253

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research*, *18*, 1279–1296.

Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*. https://doi.org/10.3758/s13428-015-0588-x

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*(3), 429–437. https://doi.org/10.1007/s00221-011-2749-1

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*(3), 429–436. https://doi.org/10.1007/s00221-012-3322-2

Hitchcock, E. M., Dember, W. N., Warm, J. S., Moroney, B. W., & See, J. E. (1999). Effects of cueing and knowledge of results on workload and boredom in sustained attention. *Human Factors*, *41*(3), 365–372.

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(4), 645–662. https://doi.org/10.1037/0278-7393.33.4.645

Joshi, S., & Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. *Trends in Cognitive Sciences*, *24*(6), 466–480. https://doi.org/10.1016/j.tics.2020.03.005

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*, 221–234. https://doi.org/10.1016/j.neuron.2015.11.028

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall. https://doi.org/10.2307/1421603

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348–354. https://doi.org/10.1177/0963721412454875

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932. https://doi.org/10.1126/science.1192439

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods*, *51*(3), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

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. https://doi.org/10.1017/S0140525X12003196

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. https://doi.org/10.1177/1745691611427305

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), 870–900. https://doi.org/10.1037/a0030694

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–284. https://doi.org/10.1038/36846

Manly, T., Owen, A. M., McAvinue, L., Datta, A., Lewis, G. H., Scott, S. K., Rorden, C., Pickard, J., & Robertson, I. H. (2003). Enhancing the sensitivity of a sustained attention task to frontal damage: Convergent clinical and functional imaging evidence. *Neurocase*, *9*(4), 340–349. https://doi.org/10.1076/neur.9.4.340.15553

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

Martin, J. D., Tsukahara, J. S., Draheim, C., Shipstead, Z., Mashburn, C. A., Vogel, E. K., & Engle, R. W. (2021). The visual arrays task: Visual storage capacity or attention control? *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001048

Matthews, R. W., Ferguson, S. A., Sargent, C., Zhou, X., Kosmadopoulos, A., & Roach, G. D. (2017). Using interstimulus interval to maximise sensitivity of the Psychomotor Vigilance Test to fatigue. *Accident Analysis & Prevention*, *99*, 406–410. https://doi.org/10.1016/j.aap.2015.10.013

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196–204. https://doi.org/10.1037/a0014104

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to “D’oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(3), 525–549. https://doi.org/10.1037/a0025896

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*(2), 109–145. https://doi.org/10.1007/BF01109051

Murphy, P. R., O’Connell, R. G., O’Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*, 4140–4154. https://doi.org/10.1002/hbm.22466

Norman, D. A., & Shallice, T. (1986). Attention to action. Willed and automatic control of behavior. In *Consciousness and Self-Regulation* (pp. 1–18). Springer.

Parasuraman, R. (1984). Sustained attention in detection and discrimination. In R. Parasuraman & D. Davies (Eds.), *Varieties of attention* (pp. 243–271). Academic Press.

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. https://doi.org/10.1016/j.physbeh.2007.09.016

Poffenberger, A. T. (1928). The effects of continuous work upon output and feelings. *Journal of Applied Psychology*, *12*(5), 459.

Psychology Software Tools, [E-Prime 2.0]. (2016). *Psychology Software Tools, Inc. [E-Prime 2.0]*. https://www.pstnet.com

Psychology Software Tools, Inc. [E-P. 3. 0]. (2016). *Psychology Software Tools, Inc. [E-Prime 3.0]*. https://www.pstnet.com

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society for Neuroscience Abstracts*, *19*, 974.

Raven, J., Raven, J. C., & Court, J. (1998). Manual for Raven's progressive matrices and vocabulary scales. In *Raven manual*.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(4), 501–526. https://doi.org/10.1037/xlm0000450

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. https://doi.org/10.1016/S0028-3932(97)00015-8

Robertson, I. H., & O'Connell, R. (2010). Vigilant attention. In A. C. Nobre & J. T. Coull (Eds.), *Attention and Time* (pp. 79–88). Oxford University Press.

Robison, M. K., & Unsworth, N. (2015). Working memory capacity offers resistance to mind-wandering and external distraction in a context-specific manner. *Applied Cognitive Psychology*, *29*, 680–690. https://doi.org/10.1002/acp.3150

Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 85–98. https://doi.org/10.1037/xlm0000444

Robison, M. K., & Unsworth, N. (2019). Pupillometry tracks fluctuations in working memory performance. *Attention, Perception, and Psychophysics*, *81*(2), 407–419. https://doi.org/10.3758/s13414-018-1618-4

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223. https://doi.org/10.1038/nrn2573

Schmiedek, F., Oberauer, K., Wilhelm, O., Süss, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology. General*, *136*(3), 414–429. https://doi.org/10.1037/0096-3445.136.3.414

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Seli, P., Cheyne, J. A., Xu, M., Purdon, C., & Smilek, D. (2015). Motivation, intentionality, and mind wandering: Implications for assessments of task-unrelated thought. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1417–1425. https://doi.org/10.1037/xlm0000116

Seli, P., Jonker, T. R., Cheyne, J. A., & Smilek, D. (2013). Enhancing SART validity by statistically controlling speed-accuracy trade-offs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00265

Seli, P., Jonker, T. R., Solman, G. J. F., 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. https://doi.org/10.3758/s13428-012-0266-1

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences*, *22*(6), 479–490. https://doi.org/10.1016/j.tics.2018.03.010

Seli, P., Risko, E. F., Purdon, C., & Smilek, D. (2017). Intrusive thoughts: Linking spontaneous mind wandering and OCD symptomatology. *Psychological Research*, *81*(2), 392–398. https://doi.org/10.1007/s00426-016-0756-3

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. https://doi.org/10.1177/0956797616634068

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*(8), 605–617. https://doi.org/10.1016/j.tics.2016.05.010

Seli, P., Smallwood, J., Cheyne, J. A., & Smilek, D. (2015). On the relation of mind wandering and ADHD symptomatology. *Psychonomic Bulletin & Review*, *22*(3), 629–636. https://doi.org/10.3758/s13423-014-0793-0

Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2016). On the relation between motivation and retention in educational contexts: The role of intentional and unintentional mind wandering. *Psychonomic Bulletin & Review*, *23*(4), 1280–1287. https://doi.org/10.3758/s13423-015-0979-0

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116–141. http://dx.doi.org/10.1016/j.jml.2014.01.004

Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process-occurrence framework for self-generated mental activity. *Psychological Bulletin*, *139*(3), 519–535. https://doi.org/10.1037/a0030010

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Smit, A. S., Eling, P. A. T. M., & Coenen, A. M. L. (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, *115*(1), 35–42. https://doi.org/10.1016/j.actpsy.2003.11.001

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A Multidisciplinary Approach to Anterior Attentional Functions. *Annals of the New York Academy of Sciences*, *769*(1 Structure and), 191–212. https://doi.org/10.1111/j.1749-6632.1995.tb38140.x

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(2), 183–194. https://doi.org/10.1518/001872000779656480

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. https://doi.org/10.1177/1745691614556681

Thurstone, L. L. (1938). *Primary mental abilities.*

Tsukahara, J. S., & Engle, R. W. (2021). Fluid intelligence and the locus coeruleus–norepinephrine system. *Proceedings of the National Academy of Sciences*, *118*(46), e2110630118. https://doi.org/10.1073/pnas.2110630118

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. https://doi.org/10.3758/BF03192720

Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychologica*, *150*, 14–25. https://doi.org/10.1016/j.actpsy.2014.04.001

Unsworth, N., & Miller, A. L. (2021). Individual differences in the intensity and consistency of attention. *Current Directions in Psychological Science*, *30*(5), 391–400. https://doi.org/10.1177/09637214211030266

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review*, *22*(3), 757–765. https://doi.org/10.3758/s13423-014-0747-6

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience*, *16*, 601–615. https://doi.org/10.3758/s13415-016-0417-4

Unsworth, N., & Robison, M. K. (2017). A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin & Review*, *24*, 1281–1311. https://doi.org/10.3758/s13423-016-1220-5

Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. *Attention, Perception, and Psychophysics*, *80*(2), 461–484. https://doi.org/10.3758/s13414-017-1455-x

Unsworth, N., & Robison, M. K. (2020). Working memory capacity and sustained attention: A cognitive-energetic perspective. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(1), 77–103. https://doi.org/10.1037/xlm0000712

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and application* (pp. 183–200). CRC Press.