DISSERTATION

MECHANISMS OF TIMING: AN INTEGRATIVE THEORETICAL APPROACH

Submitted by

Lara N. Pantlin

Department of Psychology

In partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2019

Doctoral Committee:

Advisor: Deana Davalos
Co-Advisor: Mark Prince

Matthew Malcolm
Don Rojas

ABSTRACT

MECHANISMS OF TIMING: AN INTEGRATIVE THEORETICAL APPROACH

Accurate timing allows individuals to perform essential tasks to meet societal demands, such as scheduling, responding to warning signals, and planning. Since timing impacts various functions, understanding the meaning of a timing deficit is necessary. Poor performance in neurophysiological measures of timing has been related to psychopathology but has not specifically been related to one's ability to plan or maintain a schedule. Inability to track elapsed time as done in behavioral tasks is often related to poor performance in academic settings, but the intricacies of how inaccurate timing in one task manifests in other timing tasks has not been examined. The present study proposes a comprehensive examination of timing by dividing the field into three sub-domains: neurophysiological, behavioral, and applied temporal processing. These sub-domains are organized based on the tasks traditionally used to assess timing. Neurophysiological timing (Level I) was assessed using a duration-based mismatch negativity paradigm (dMMN), which fundamentally requires minimal cognitive resources. Behavioral timing (Level II) introduces the role of attention and working memory to accurately determine the amount of elapsed time (verbal estimation) or the generation of a pause, which reflects a specified amount of time (interval production). These tasks do not require the higher-order cognitive functions such as decision making and planning which are needed to accurately perform applied temporal processing tasks (e.g., time management and scheduling) (Level III). Hypothesis I proposed a hierarchical relationship among the three subdomains in which each succeeding level in the mediation is informed by the previous one and is distinct from the others

based on the amount of cognition required to perform the task. Hypothesis II not only offered an extension of Hypothesis I, but also sought to examine the ways timing can be systematically improved through intervention methods. Across two time-points, participants were screened for select psychopathologies often associated with timing deficits (e.g., psychosis, traumatic brain injury, and substance use), underwent EEG recordings of dMMN to measure neurophysiology (Level I), performed two behavioral timing tasks (verbal estimation and interval production) (Level II), and completed three measures of applied temporal processing (letter-number sequencing and two time management surveys) (Level III). Hypothesis I was analyzed using a mediation model where neurophysiology (Level I) is expected to inform behavioral performance (Level II), which would subsequently influence accuracy on applied tasks (Level III). Hypothesis II was analyzed using repeated-measures ANOVAs to assess which intervention increases accuracy between time-points. Although Hypothesis I yielded nonsignificant results, interesting trends in the expected direction existed. Higher responses on the neurophysiological tasks were related to higher accuracy on behavioral and applied temporal processing measures. Hypothesis II yielded significant interactions between session and intervention and overall, suggested that using feedback to calibrate individuals to their abilities is the most appropriate intervention technique for increasing behavioral and applied accuracy. However, inclusion of tasks evaluating intermediate stages of timing is required if a full scale time continuum is to be modeled. Yet, this work provided the initial groundwork to further investigate the way time-related information is handled in the healthy brain.

ACKNOWLEDGMENTS

I would like to express thanks to committee members **Drs. Deana Davalos, Mark Prince, Don Rojas, and Matthew Malcolm** for their service on my dissertation committee and for dedicating time and effort towards my degree. I appreciate the many meetings, rounds of edits, and insightful feedback each of you provided throughout the project.

I would like to acknowledge the undergraduate research assistnats that volunteered countless hours and assisted in data collection. This list of undergraduate includes: Lisa Hirt and Adrienne Gaughan, who served as lab managers and trainers, Avery Scott, Zach Throvaldson, Hannah Crosby, Callie Harward, Jenna Klippenstein, Asher Augustinis, Lydia Akalu, Joyce LaBonte, Mark Carmen, Matthew Schiltz, Brianna Levine, Addison Babineau, Savannah Regensburger, Kym Schrock, Kaylynn Little, Morgan Stam, Natalie Miller, and Paige Cummings. I would specifically like to extend my gratitude to my dedicated and brilliant lab manager **Lisa Hirt**, who served on this project for over three years. During her time in the lab, she assisted with the pilot stages of this project and challenged me to learn new statistical modeling techniques to answer interesting theoretical questions that significantly informed the final hypotheses. She singlehandedly ensured that data collection and scheduling moved smoothly, allowing me to focus on other areas of the project. My timeline would have been compromised without her assistance and I am forever grateful for her service.

I thank **Dr. Mark Prince** for his dedication to this project long before the established title of co-advisor. Dr. Prince has supported my passion for statistical modeling since his arrival at Colorado State University. He has challenged me to explore new modeling techniques and given me the confidence to pursue a career in this field. I am forever in his debt for the countless hours

he spent with me on Friday afternoons dissecting my hypotheses and the appropriate modeling techniques. I cannot thank him enough for his dedication to this project and the many ways he has aided my journey to PhD candidacy.

I extend my gratitude to my advisor for the past 5 years, **Dr. Deana Davalos**. I thank you for choosing me as your graduate student and giving me the opportunity to pursue my dream of becoming a doctorate. I appreciate that ways you always kept me motivated throughout my academic career. I will greatly miss our meetings and your sense of humor, specifically around hard times, along with the academic and work-life guidance you have given me. Your ability to combine professionalism with sarcasm made graduate school significantly less stressful and you were all that I could have asked for in a mentor (and more).

Additionally, I would like to thank **Dr. Karla Gingerich**. Dr. Gingerich, I appreciate the many last minute meetings, the support, and guidance you have given me over the past two years. I cannot thank you enough for 'talking me off the cliff' and for teaching me to be a better instructor, mentor, and individual.

Meredith college faculty members**, Drs. Doreen Fairbank, Cynthia Edwards, and Steven Benko**, I thank each of you for preparing me for this journey and keeping up with me throughout it. **Dr. Doreen Fairbank**, my academic mother. Thank you for calling me and supporting me even though I had left your nest. I appreciate the ways in which you always checked up on me, reminding me of who I am, and where I came from. **Dr. Cynthia Edwards**, that bottle became quite the motivator. There was never a doubt in my mind that we would one day share that toast. Every time you called me with statistics or research questions for the lab I helped you build at Meredith, it reminded me why I began this journey. **Dr. Steven Benko**, my PhD shaman. Thank you for reminding me that done is done and enough is enough, and for

always knowing exactly what to say (aka providing me with lessons in running metaphors, since that is the only thing that truly makes sense to me). Thank you to each of you for mentoring me through my undergraduate career and life after, as each of you ultimately preparing me for success in graduate school.

I would like to extend gratitude to my dearest graduate friends, **Hannah Hausman, Alex Claxton, Dr. Nate Tenhundfeld**, and **Marielle Darwin**. Hannah, for always taking care of my interests outside of graduate school, for killing me in the gym to rid me of graduate school anxiety, and for taking that last minute guest lecture so I could finish my defense presentation. To Alex and Dr. Tenhundfeld, my graduate school brothers. Thank you for supporting me from day one until the end, and making fun of me every step of the way. Lastly, thank you Marielle Darwin. Your countless edits on this document, the number of last minute meetings where I was so frustrated I could not have created a sentence, for working through my ideas and helping me organize this project from Day 1. Academically, you are the best graduate lab partner. Thank you for seeing me through on this journey and I look forward to all of our other journeys we will have together. If graduate school did one great thing for me, it was bringing us together. Thank you to each of you for always celebrating my successes and being there for me for the failures. I would not have come out of graduate school in one piece without each of you.

I am grateful for the support and love that my family has provided me throughout this endeavor. My mom and dad, **Suzy and Tony Adams**, for supporting me and helping me build the confidence to go after my dream career. Mom, thank you for the many edits on this document as well and for making me write essays as a child when I was in trouble. I think that was the first stepping stone towards a PhD. To my dad, **Mark Pantlin**, and brothers, **Eric, Alexander, and**

Nicholas Pantlin: Thank you for celebrating the good times, listening to the complaints, and occasionally reading my papers.

Lastly, I would like to thank my soon-to-be-husband, **Austin Southard**. Austin's countless edits on my documents, read throughs and reread throughs, insightful commentary, and sacrifices amounted to a successful, timely dissertation defense. Whether he was working to keep us afloat, making sure I was well-fed, or listening to a presentation (or single sentence) for the hundredth time, Austin has been there for me through thick and thin. Thank you for teaching me to be kinder to myself, believe in my abilities, and trusting my process through graduate school. It was always clear to me that you never had a doubt I would successfully reach this goal and your confidence in me was a impetus to my successful defense. Also, proposing just before my defense was a great confidence boost and made the week leading up to my defense significantly less stressful. Although I moved to Colorado in search for academic and career advancement, we found each other and started something much more fulfilling than I could have ever imagined. You have learned with me and taught me more lessons than were available in graduate school. *The next adventure awaits us.*

DEDICATION

This work is dedicated to **Gloria Furman**

In loving memory of her lifelong pursuit for knowledge, quest for original experiences,

and her unyielding support of my educational journey.

TABLE OF CONTENTS

INTRODUCTION

Timing is central to a variety of general functions. Accurate timing allows individuals to perform the functions required to meet societal demands such as creating and maintaining schedules, understanding chronological sequences of events, and accurately executing other basic functions. Small inaccuracies, left unaccounted for, may lead to lateness or poor time management. The ability to create and maintain schedules is often required in employment and educational settings as well as other expectations in daily life. Maintaining the capacity to accurately manage time also allows people to perform precise motor functions, understand speech, respond to warning signals in the environment, sequence events, and appropriately carry out executive functions such as planning and decision making (Foster, Kisley, Davis, Diede, Campbell, & Davalos, 2013). Individuals would not be able to carry out these expected skills without the ability to process time in a linear fashion.

Timing fundamentally influences several areas of daily functioning; therefore, it is important to consider how timing is examined in research. Traditionally, timing is viewed as one, unified process, but the methods of evaluating this construct are contingent on distinct sub-disciplines within the field. For example, in neurophysiological research, timing is often assessed using event-related potentials (ERPs) to measure brain activity. Studies frequently related deficits in neurophysiological timing are to an overall ability to process and manage time units. Since the prevailing view of timing is as a single entity, abnormal neurophysiological activity is currently accepted as an indication of widespread inaccuracies in temporal processing. Studies using behavioral tasks to evaluate timing frequently operate under the same assumption. These conventional behavioral timing tasks, such as tracking the length of a minute without any cues,

require more cognitive resources than needed to complete a neurophysiological timing task. While neurophysiological and behavioral tasks seem to assess timing from different positions, the anticipation remains that a deficit in behavioral timing tasks could yield similar shortfalls in applied temporal processing or tasks requiring an initial understanding of time (i.e., meeting deadlines, time management, and planning). From here, it is a short leap to assume that timing deficits exposed in neurophysiological timing may also manifest in a behavioral task, but the relationship between these tasks has yet to be examined in experimental conditions. Thinking of timing in a holistic manner may be misleading to researchers in that the relationship between the abovementioned tasks and tasks directly measuring applied temporal processing (e.g. time management skills, planning, scheduling) has yet to be examined.

Assessing timing with distinct yet interdependent subconstructs permits a more conclusive understanding of a deficit in timing. This interpretation can better identify where these shortfalls reside, when deficits first manifest on an individual basis, and promote opportunities to investigate novel intervention methods seeking to increase time accuracy. Improving timing is an area of interest for researchers because deficits in timing are related to various disorders (Bartha-Doering, Deuster, Giordano, Zehnhoff-Dinnesen, & Dobel, 2015). The theory remains: improving accuracy may alleviate symptomology associated with these pathologies. Furthermore, research has shown the relationship between better accuracy on various timing tasks and higher intelligence, specifically in math, better working-memory, and mental health (Kramer, Bressan, & Grassi, 2011). Thus, examining timing as subdomains will not only allow for a better understanding of timing in general, but also provide a basis for intervention methods.

Hypothesis I

This study proposed a comprehensive understanding of timing by separating the field into subdomains based on the tasks traditionally used to evaluate timing. These subdomains are neurophysiological, behavioral, and applied temporal processing. If timing truly is the overarching process underlying each of the tasks in these subdomains, then a statistical relationship between subdomains should be evident. This proposed relationship is expected to be a mediation where each succeeding level in the mediation illustrates a more complex element of timing, requiring increasingly more cognitive resources to perform the tasks accurately. The subdomain of neurophysiology was considered as the first processing level of timing, then behavioral, followed by applied temporal processing. The following sections introduce the common tasks used to evaluate timing within each of the proposed subdomains. This review of the literature offers support and additional information regarding the tasks chosen in the proposed study. The last section of the introduction revisits the appeal to improving timing and introduces the second hypothesis.

Level I: neurophysiological timing. Neurological processes are thought to underlie cognitive abilities and therefore, neurophysiological timing is proposed as the first level in the mediation. Similar to other processes, timing is associated with an underlying neural circuitry primarily involving the auditory cortex and its connections with the prefrontal cortex (Buhusi & Meck, 2005; Ivry & Schlerf, 2008; Ivry & Spencer, 2004). Thus, neurophysiological timing is most often assessed using an event-related potential known as mismatch negativity (MMN) because increased brain activity in response to the MMN paradigm reflects more accurate temporal processing in the auditory cortex (Näätänen, Paavilainen, Rinne, & Aho, 2007). MMN also demands minimal cognitive resources, making this paradigm vastly distinct from the tasks

used to assess the proposed higher levels of timing (i.e. behavioral timing and applied temporal processing). Thus, MMN may be the purest assessment of timing at a neurophysiological level available.

While MMN can be elicited using various sensory stimuli, experiments examining temporal processing in the auditory cortex most often utilize an auditory stimulus. An individual who demonstrates accurate temporal processing also produces a larger MMN response, which is measured by a negative brain amplitude occurring 140-210 ms after an auditory stimulus (Näätänen, 1984). These auditory stimuli are categorized as standard tones or deviant tones. Standard tones share the same auditory elements such that all stimuli maintain the same duration, pitch, frequency, et cetera. Deviant tones, however, differ from standard tones in at least one of these auditory properties. When the standard set of stimuli are presented in a sequence, the individual listening quickly becomes familiar with the presented pattern. Information regarding the pattern is held in sensory memory, which allows the individual to predict the auditory properties of the next stimulus. Since all standard stimuli are the same, the expectation is that the next incoming stimulus resembles the previous stimuli. If a deviant tone is played in place of a standard tone, this violates the regularity of the pattern and the individual then produces a measurable brain response. Finally, to be considered MMN, an average response waveform is created from all standard tones. This average standard waveform is subtracted from the average response to all deviant tones and these responses must occur in 140 to 210 ms post-stimulus (Näätänen, 1984).

Since MMN processes are contingent upon sensory memory and the response is generated at an early sensory stage (140-210 ms), limited cognitive demand is required to produce an MMN response. When standard tones are presented in a sequence, an automatic

neural representation of the pattern becomes encoded in sensory memory and is referred to as the standard formation (Bartha-Doering et al., 2015). This information must be retained for the standard formation to be used as a baseline comparison for subsequent tones. Retention of the standard formation relies on an individual's sensory memory as opposed to the ability to attend to each incoming stimulus. If any attention is required, the type of attention related to sensory memory is akin to the *Pop out Phenomenon* of visual attention in which an individual's attention is grabbed, but not intentionally engaged. Other arguments stating that attention is a modulator of MMN are strictly lacking, especially once the standard formation has been encoded (Sussman, Winkler, Huotilainen, Ritter, & Näätänen, 2002). Studies have shown that the final MMN response waveform can be produced without specifically attending to the stimuli. Patients in comas, sleeping participants, and infants are all able to produce an MMN response (Garrido, Kilner, Stephan, & Friston, 2009; Sculthrope, Ouellet, & Campbell, 2009). Additionally, conscious, neurotypical individuals who were asked to specifically attend to stimuli did not demonstrate significant changes in MMN responses (Garrido et al., 2009). Therefore, this measurement is an appropriate tool for the first level of the proposed timing mediation as MMN demonstrates how temporal comparisons are automatically completed at an early, sensory level and without the involvement of higher-order cognitive functions (i.e., decision-making, memory, planning; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001).

Due to the low cognitive demand and passive nature of the paradigm, MMN is often used to assess timing abilities in a variety of populations (i.e., newborns, adults, and clinical patients). Additionally, MMN has been habitually used to isolate the ability to time in clinical patients who have known deficits in cognitive functions such as working memory, attention, and decision-making, which are required to complete more complex timing tasks (Ciullo et al., 2016; Shelley

et al., 1991). Those are considered healthy produce large, negative response waveforms when subjected to the MMN paradigm. This response is thought to be related to accurate or suitable timing abilities at the neurophysiological level. Studies comparing healthy individuals to clinical populations with certain types of deviant tones have found that attenuations in the MMN response waveform may be reflective of auditory disturbances and timing deficits (Umbricht & Krljes, 2005).

The difference in MMN responses between healthy and clinical populations is most evident when the standard and deviant tones differ in duration. For example, in duration MMN (dMMN) standard tones are each separated by 500 ms and the deviant tone is played sooner than 500 ms. Differences in dMMN are especially robust when comparing neurotypical individuals to patients with schizophrenia. This difference has been consistently found across studies and confirmed in a meta-analysis by Umbricht and Krljes (2005), which examined the relationship between MMN and schizophrenia across approximately 40 studies. Umbricht and Krljes (2005) determined that patients with chronic schizophrenia have decreased dMMN amplitudes when compared to controls (Cohen's d = .99). Due to this large effect, deficits in MMN, specifically dMMN, have been proposed as a cognitive endophenotype of schizophrenia (Ciullo, Spalletta, Caltagirone, Jorge, & Piras, 2015). While other types of MMN (e.g., frequency and pitch) also relate to differences in timing, the most robust effect is when dMMN is compared across populations. The proposed study screened for those who endorse symptoms of psychosis, utilize this information as a covariate (see Analysis Plan for Hypothesis I), and use dMMN to assess neurophysiological timing.

Level II: behavioral timing. While neurophysiological timing exists at an early sensory stage and needs minimal cognitive resources, behavioral tasks require the individual to attend to

external stimuli, hold information in working memory, and may recruit additional cognitive functions. These behavioral tasks often ask individuals to give an approximation of how much time has passed (verbal estimation) or construct a time interval (interval production) (Grondin, 2008). Traditional behavioral tasks often do not require individuals to engage in extensive decision-making, planning, or other higher-order cognitive functions (Wearden & Lejeune, 2008). Instead, those higher-order cognitive functions are needed to complete applied temporal processing tasks (i.e., general time management tasks or creating and maintaining a schedule) and are segregating factors between these two levels. Thus, behavioral timing is proposed as the second level in the mediation because to accurately complete a behavioral task a person must exercise more cognitive functions than are needed for neurophysiological timing, but the task requires less complex cognitive operations than are needed to accurately perform an applied temporal processing task.

In the past few decades, studies have most often used the following three tasks used to assess behavioral timing: verbal estimation, interval production, and temporal reproduction (Grondin, 2010; Poynter, 1983; Wearden, 2014; Zakay, 1990). In verbal estimation, an indication of the start and end of the time window or interval is produced, often using a tone or flash. Participants are asked to verbally specify the amount of time a target window lasted in the appropriate temporal unit (e.g., milliseconds, seconds, or minutes). For interval production, a participant is given a target window again, but then asked to produce this window by indicating the beginning and end of a specific interval using a stopwatch or other form of a start and stop button such as a finger tap, or a buzzer. In temporal reproduction, a participant first listens to a target window as done in verbal estimation; however, the participant must also then recreate the target window by indicating the start and stop of a specified time-period (Zeiler & Hoyert, 1989).

Accurate behavioral timing is demonstrated when a participant's response matches the target time window in all tasks. However, reports suggest that accurate temporal reproduction is more indicative of appropriate motor responses and reaction times as opposed to the ability to track the passage of time (Caldara et al., 2004; Mioni, Stablum, McClintock, & Grondin, 2014; Wearden, 2014). The present study therefore only utilized verbal estimation and interval production and did not examine temporal reproduction.

Mioni et al., (2014) criticized previous timing studies for choosing "the entire repertoire of tasks, but in most cases hav[ing] provided no rational for the selection of the specific task" and therefore aimed to investigate and compare the available behavioral timing tasks. Findings from this study deemed interval estimation and production tasks as appropriate measures for investigating variations amongst internal clocks by person (Mioni et al., 2014). Estimation tasks may be less accurate as participants are more likely to estimate a whole number as opposed to a fraction, but overall, these two tasks reflect "two-sides of the same coin" (Grondin, 2008, 2010; Zakay, 1990). Specifically, results of Mioni et al. (2014) found that participants were most accurate when instructed to use a key press to start and stop the production of an interval, as in the production task, but generally demonstrated a slight tendency to overproduce the intervals (Mioni et al., 2014).

Generally, individuals tend to overestimate brief intervals and underestimate long intervals (> 5 seconds; Zakay, 1990; see also Vierordt's Law, 1951). Current work has demonstrated that individual differences also explain inaccurate behavioral timing such as mood, stress, and emotions, but experimental factors also affect time perception (Allan, 1979; Eagleman & Holcombe, 2002; Eagleman, 2008). The type of stimulus used in an estimation or production task is crucial to how individuals perceive the duration and therefore is directly

related to if a participant over or under estimates an interval. When visual or auditory cues for the interval have high intensity (e.g., a bright light), this confounded responses, such that participants often perceived the intervals to be longer than if a more neutral stimulus was used (Poynter & Homa, 1983). Other confounding factors of the perception of time include mood, hunger, activities, or tasks during the interval (see meta-analysis by Block, Hancock, and Zakay, 2010); however, these factors were accounted for in the present study. Typically, studies examining these confounding variables found that individual variation is then most likely due to the person's cognitive performance. For example, memory and attention become increasingly more mandatory to accurately estimate longer intervals (Block et al., 2010; Eagleman & Holcombe, 2002; Eagleman, 2008; Fraisse, 1984).

To fully explain the relationship between cognition (i.e., attention and working memory) and behavioral timing tasks (i.e., verbal estimation and interval production), an explanation of the internal clock model must be addressed. The internal clock model proposed by Treisman (1963) includes an intrinsic pacemaker that emits a pulse, which is counted in the accumulator (see Figure 1). Time is ultimately tracked by counting these internal pulses. Suppose an individual is expected to verbally estimate an interval of 30 seconds. Once the start tone or flash is produced, the individual's pacemaker begins to emit pulses. These pulses are linearly related to the passing of time and are stored in the accumulator as subjective time units. Assuming there are no other distractions and an individual is lending all attentional resources to track time, a healthy individual should be able to accomplish this task without error. Thus, accurate time estimation is dependent upon the available attentional resources, which are limited and assigned based on the importance of the present task (Grondin, 2010). Accordingly, if the participant is supposed to complete an additional task during a specified interval, attentional resources are

reallocated to this task and away from tracking time. This shift may result in an imprecise estimation of time and is explained in the model by as a mechanism between the pacemaker and accumulator, referred to as the switch. The switch determines the flow of pulses when attending to time (Grondin, 2010). As additional attention-demanding tasks or stimuli are introduced, the switch closes or blocks the stream of pulses from the pacemaker. This adjustment ultimately explains inaccurate timing at the behavioral level in a healthy individual.

The internal clock model (Treisman, 1963) suggests that all individuals have an internal clock mechanism which keeps track of time, but this clock is dependent upon attention to time and is sensitive to distractions (e.g., arousal, emotions, motivation, and mood; Gros et al., 2015). However, errors in behavioral timing may also be due to shortfalls in working memory. To examine the relation between working memory and timing, Gibbon et al., (1984) built off Treisman's internal clock model to create the Scalar Expectancy Model of Timing (SET; see Figure 2). This model is divided into three stages: the clock stage (Treisman's the internal clock model), the memory stage, and the decision stage. Under the SET model, if an individual is tasked with producing a minute span of time (interval production), then the individual would indicate the 'start' and the pacemaker begins emitting pulses. As the accumulator is counting subjective time units, this information is held in working memory (memory stage) and continuously compared to the individual's memory of a minute (reference memory) until the individual decides whether or not the time elapsed is equal to a minute (the decision stage). Thus, if an individual has any limitations in working memory, then their ability to perform on both verbal estimation and interval production tasks would be compromised (Grondin, 2010; Wearden, 2014).

Interval length is another cause of differences among healthy individuals in behavioral timing tasks. Neurotypical participants tend to show positive relationships between accuracy and interval duration. Those who are inaccurate when estimating or producing short intervals will thereby become exponentially inaccurate as the intervals become longer. Therefore, the present study utilizes intervals that are considered short or less than one minute (Ciullo et al., 2015; Elvevåg, McCormack, Glibert, Brown, Weinberger, & Goldberg, 2003). The use of short intervals also adds to the proposed mediation. If there are deficits in short intervals as shown by behavioral timing, the expectation is that disturbances in applied timing tasks would be evident as well. Specifically, extrapolations of time for long range plans would be confounded by an attenuated ability to time in a short interval range.

Level III: applied temporal processing. Tasks testing intermediary levels of timing (i.e., behavioral timing tasks) may require participants to estimate the amount of elapsed time or produce an interval in the seconds to minutes range. Studies assessing behavioral timing often utilize intervals that are short (< 60 seconds) because even neurotypical participants become increasingly inaccurate with longer intervals (Ciullo et al., 2015). This result is partly due to the increasing amount of cognitive resources needed to complete these longer tasks. For instance, a participant engages less attentional resources when tracking a five second window compared to an hour. As time intervals lengthen, accuracy not only becomes increasingly dependent on the ability to engage attention and working memory but also the individual's ability to extrapolate from their fundamental understanding of time. Tasks measuring this skill, or applied temporal processing tasks, often require individuals to self-assess their proficiency in time management, use of time, and applying attention and working memory skills to timing. Therefore, to accurately complete applied temporal processing tasks, more cognitive power is required than in

behavioral or neurophysiological timing tasks. Applied temporal processing is hence proposed as the final level in the model (see Figure 3 for proposed model and Figure 4a-c for final models).

The variety of complex cognitive functions (e.g., decisions pertaining to task duration) needed to perform adequately on applied temporal processing tasks are extensions of an individual's ability to apply attentional resources and working memory capabilities towards timing. This proficiency is measured in the present study using the letter-number sequencing (LNS) task, which is a subsection of the Wechsler Adult Intelligence Scale (WAIS-IV; Wechsler, 2008). In this task, the participant listens to a series of numbers and letters in which one letter or number is spoken per second. The participant must hold this information in working memory, reorganize the series into ascending numerical order first, then letters alphabetically, and repeat this information back to the examiner. Therefore, the LNS tasks is thought of as a measure of auditory attention and working memory in which higher scores on the LNS task are related to better auditory attention, working memory, and overall intelligence (Vuoksimaa, 2004). In the present study, the LNS task and specific evaluations of time management ability are proposed as measures of applied temporal processing.

Time management is an inclusive way to measure a person's ability to apply temporal processing to daily living, as effective time management skills require an accumulation of accurate anticipation of task duration, creation of a plan pertaining to this estimation, and accounting for potential setbacks (Clasessens et al., 2007). This view of time management is often examined in two stages. The first stage requires the ability to accurately predict task duration and establish a schedule around this prediction (Francis-Smythe & Robertson, 1999). The second stage relates to maintenance of the planned schedule. This ability requires an extrapolation of the same type of processing used to monitor the passing of time (i.e. the internal

clock theory), but on a larger scale. Thus, to be a successful time manager, an individual must be able to accurately and efficiently allocate time.

Efficient time management requires an individual to correctly make decisions about task duration and allocate time accordingly. These two functions are informed by the individual's initial ability to estimate time intervals. This propensity is frequently measured using time management questionnaires, which are reliable indicators of an individual's proficiency in temporal processing. The Time Management Questionnaire (TMQ) and the Time Structure Questionnaire (TSQ) are two of the most conventionally used measures to self-assess time management abilities (Clasessens et al., 2007). The TMQ, which was specifically developed to examine time management practices in college students, measures time management behaviors and specifically, planning the allocation of time (Clasessens et al., 2007). Additionally, higher scores on the TMQ are related to work and academic success, which suggests that differences in the ability to allocate time may also be related to these outcomes (Razali, Rusiman, Gan, & Arbin, 2018). Conversely, the TSQ does not measure time management behaviors, but instead measures perceptions of how an individual makes use of allocated time. A high score on the TSQ indicates that an individual's use of time is perceived as structured and purposeful. Higher scores are also related to psychological well-being, optimism regarding future outcomes, academic success, and fewer physical and psychological symptoms related to depression (Bond & Feather, 1988; Macan, Shahani, Dipboye, & Phillips 1990). Perceived control of time has also been related to enhanced work and academic performance (Clasessens et al., 2007). Together, these measures offer a more comprehensive evaluation of an individual's applied temporal processing abilities than the LNS task alone.

The ability to apply one's understanding of timing to day-to-day functioning is fundamentally dependent upon lower levels of timing. Without a core understanding of timing, the ability to plan and make decisions around time is strictly diminished. To accurately allocate time or have successful applied temporal processing, one must be proficient in time estimation (i.e., behavioral timing tasks). A robust understanding of how time passes, measured with estimation and production tasks, determines how much time an individual believes should be dedicated to a task and ultimately inform time management skills. Furthermore, these applied temporal processing tasks require additional attentional resources and auditory working memory compared to behavioral and neurophysiological timing tasks. However, if timing is not examined in isolation from cognitive abilities, those who may be ineffective timers and overcompensate with strong cognitive abilities would not be addressed. Therefore, neurophysiological timing tasks are proposed to fundamentally inform upper levels of timing because these tasks best separate the ability to time from cognitive abilities. In sum, the present study introduced a statistical model which describes the interaction between levels of timing to ultimately explain differences among timing abilities (see Figure 3 for proposed model and Figure 4a-c for final models).

Hypothesis II

While Hypothesis I examined past literature that measures different types timing as possible subdomains, Hypothesis II (1) offered a possible extension of the previous hypothesis and (2) examines if timing can be systematically improved. To further investigate the nuances of Hypothesis I, interventions aimed to change timing are specifically introduced at each of the proposed subdomains. If these levels of timing in Hypothesis I are truly related in a hierarchical manner, then interventions which increase time accuracy at higher processing levels (i.e., applied

temporal processing) should not produce any changes in previous processing levels (i.e., neurophysiology or behavioral timing). Additionally, increased time accuracy in the first level of the mediation (i.e., neurophysiology) should produce increased accuracy at every level. If behavioral timing is justly placed as the second level of the mediation, then increasing behavioral accuracy should produce more accurate responses on the behavioral and applied temporal processing tasks while producing no effects on neurophysiological timing. The subsequent sections address the significance of improving temporal processing, regardless of the relationship between the proposed subdomains of timing.

As previously stated, timing affects a multitude of functions of daily living in all individuals; thus, the possibility of manipulating or improving time accuracy is an important area of focus. Differences in performance on behavioral and applied temporal processing tasks may suggest gradations in intelligence and/or cognitive functioning (Franssen, Vandierendonck, & Hiel, 2006; Kramer et al., 2011). Specifically, imprecisions in these tasks may manifest as problems with time management or procrastination but may be resultant of a deficit in other cognitive functions. For example, Franssen et al. (2006) determined that decreased precision in both verbal estimation and interval production was related to an individual's phonological working memory, suggesting that variations in working memory may account for some of the differences in behavioral timing. Research has also shown how the ability to accurately estimate time not only predicts better working memory, but also mathematical intelligence (Kramer et al., 2011). Those with lower intelligence, as measured by the intelligence quotient (IQ) test, performed worse on behavioral tasks compared to those with higher IQs (Elvevåg et al., 2003). This finding insinuates that while intelligence is not the driving mechanism behind accurate timing, it does play a partial role in the ability to perform on timing tasks. Additionally, Balci et

al. (2011) examined estimation abilities in participants during experimental risk-assessment conditions. Those who were accurate timers also had better decision-making performance in the risky conditions. As a precursor for overall functionality, Gallistel and Gibbon (2000) named accurate time estimation a key element of organized behavior, suggesting that behavioral differences in timing may also be related to differences in psychological health.

Like behavioral timing, those with psychopathology may demonstrate reduced responses during neurophysiological timing tasks (Davalos, Kisley, & Freedman, 2005; Light & Näätänen, 2013). These deficits in neurophysiological timing are often correlated with symptomology of psychosis and schizophrenia, which suggests that increasing time accuracy may reduce the associated symptomology (Bodatsch et al., 2011; Umbricht & Krljes, 2005). This deduction has been the headway for studies examining MMN amplitude changes following behavioral training. Schröger, Paavilaninen, Näätänen (1994) and Kraus, McGee, Carrel, King, Tremblay, and Nicol (1995) were among the first to compare changes in MMN amplitudes before and after a training task. The expectation is that, following behavioral training, participants would become increasingly more accurate in behavioral tasks (i.e., classifying, estimating, or discriminating intervals) and this ability would correlate with a larger MMN response. The six studies used an identification task demonstrated significant MMN amplitude changes across time-points (Heim, Choudhury, Benasich, 2016; Kharlamov, Campbell, Kazanina, 2011; Miller, Zhang, & Nelson, 2016; Tamminen, Peltola, Kujala, Näätänen, 2015; Tong, Melara, & Rao 2009; Ylinen et al., 2009). Twelve studies using MMN as a pre and post-test indicator of timing improvement found significant amplitude changes following a discrimination task (Atienza, Cantero, & Dominguez-Marin, 2002; Caruso, & Balaban, 2015; Gottselig, et al., 2004; Kharlamov et al., 2011; Kraus et al. 1995; Miller, Zhang, & Nelson, 2016; Petersen et al., 2015; Schröger et al.,1994; Tamminen

& Peltola 2015; Tamminen et al., 2015; Tremblay, Kraus, Carrel, & McGee, 1997). Of these studies, the longest time delay between pre- and post-tests with a significant amplitude change was 9-weeks, demonstrating that behavioral training could contribute to a long term neurophysiological effect (Russo, Nicol, Zecker, Hayes, & Kraus, 2004; Tong et al., 2009).

These studies found significant results using less complex behavioral tasks such as identification and discrimination of intervals. Limited studies have paired MMN amplitudes with the more common and slightly more cognitively demanding behavioral tasks (i.e., estimation and production) adding to the novelty of the present study. Although previous findings support that behavioral training can induce neurophysiological changes, none have systematically examined the social and cognitive manifestations of increasing neurophysiological responses.

Purpose and Hypothesis

Prior research has mostly examined one modality in isolation, providing a disjointed, myopic view of the mechanisms that comprise timing. Overall, the present study offers a new platform of timing by separating timing into three subdomains (i.e., neurophysiological, behavioral, and applied temporal processing). These were categorized based on the tasks commonly used to assess timing and the amount of cognitive resources required to complete each task. This study suggested that if these subdomains of timing are truly related to one another, then timing should be addressed as a one entity with secondary categories; however, if these levels were not related then timing should be examined as a single, fluid process.

The first hypothesis aimed to test the hierarchical relationship of timing using a mediation model (see Figure 3 for proposed model and Figure 4a-c for final models). Hypothesis I posited that better timing at the neurophysiological level is associated with accurate performance on behavioral tasks and better applied temporal processing. By statistically

supporting the hierarchical nature of the relationship between the three processing levels of timing, a more succinct understanding of timing can be obtained. This would presumably lead to better development of techniques aiming to increase time accuracy.

Improving timing could lead to symptom reduction in psychopathologies and increases in executive functioning such as intelligence and working memory. Therefore, Hypothesis II aimed to determine if time accuracy can be improved and sought to generate the optimal improvement method across dimensions of timing. Hypothesis II suggested that within a specified condition, time accuracy would increase at the level targeted by a specific intervention, such that those in Intervention I (MMN) would show improvements on the neurophysiology tasks across time-points. Those in Intervention II (Metronome) would demonstrate an increase in behavioral accuracy from Time 1 to Time 2 and those in Intervention III (Calibration) would display increased scores on the applied temporal processing measures across time. Controls would not show a significant change across time.

METHODS

Participants

A survey screening for various psychopathologies (e.g., depression, anxiety, autism spectrum disorder, prodromal) was administered to undergraduate, introductory level psychology students at a large university in the United States of America (N = 204; Final N = 145). This survey included demographic information, which ultimately provided information on variables that may increase or decrease MMN amplitudes. These factors included: use of nicotine, past or current traumatic brain injuries, and a score of 6 or higher on the 16-Item Prodromal Questionnaire (PQ; Ising et al., 2012) (Martin, Davalos, & Kisley 2009; Pantlin & Davalos, 2016; Tao, Sun, Li, & Chen, 2015). Other psychopathology such as anxiety, depression, or autism related disorders were not included in the final analysis due to either conflicting consensus on how these disorders relate to timing in the literature (e.g. autism) (Andersson, Posserud, & Lundervold, 2013; Gomot et al., 2011) or work suggesting that the disorder is not characterized by timing difficulties (Oberfeld, Thones, Palayoor & Hecht, 2014).

Those who indicated nicotine use or indices of brain injury were removed from the analyses as these variables introduce potential influence over the final effect (see Figure 5). Two and a half percent of participants indicated that they use nicotine and 18.1% of participants reported that they had been knocked unconscious before, suggesting a brain injury occurred. Thus, 5 participants were removed for nicotine use and 12 were removed for brain injury. One participant indicated both use of nicotine and indices of a brain injury, resulting in 187 participants. The remaining 42 participants were removed from the final analyses because these participant's MMN amplitude did not meet criteria for an MMN response or was considered an

outlier. This is further explained in Neurophysiology Preprocessing subsection under the Data Preprocessing for Analyses header. Information regarding cannabis use was asked in the form of 'Have you ever smoked cannabis?" "Did you smoke cannabis today?" "Approximately how often do you smoke cannabis?" There were 6 participants who indicated they smoked cannabis but had not on the day of the experiment after removing all other possible confounding variables. These individuals indicated they were not habitual users and were spread across interventions evenly.

Overall, the inclusion and exclusion criteria were as follows: 1. Participant with hearing impairment were excused from the study and no data were collected. 2. Participants who indicated use of nicotine were excluded from the final analyses. 3. The data associated with those who indicated previous traumatic brain injury or losing consciousness following forceful head contact was also excluded. 4. If the participant's amplitudes were not negative, then this response did not meet criteria for an MMN response and their data was excluded (Näätänen, 1984). 5. Participants with amplitudes that exceeded +/-3 standard deviations from the mean amplitude were counted as outliers, excluded, and this assisted with non-normal variables. 6. Those that had over 50% missing data were excluded from the analyses (see Figure 5).

Of the final 145 participants used in the analysis, 44.8% had no musical experience (n = 65), 13.1% had 1-3 years of experience (n = 19), 25.5% had 3-5 years of experience (n = 37), .7% reported 5-10 years of experience (n = 1), and 15.2% reported more than 10 years of experience (n = 22). Years of musical experience and higher scores on the 16-item PQ suggests endorsement of psychosis related symptoms both have influences over MMN amplitudes, as suggested by prior research (Pantlin & Davalos, 2016; Rammsayer & Altenmüller, 2006). There were 129 participants who did not endorse these symptoms or scored below a 6 on the 16-item

PQ and 15 participants who did endorse symptoms. All of these variables are included as covariates in both analyses to account for the potentially spurious effects each variable may introduce to the overall model (see Figure 3 for proposed model and Figure 4a-c for final models).

Of the final distribution of participants, 41 (28.3%) self-reported as male and 103 (71.0%) identified as female. The ages of participants ranged from 18-26, with a larger distribution in the 18-21 or traditional college age range. The highest percentage (46.2%) of participants were 18 and the lowest percentages (.7%) were 26. One hundred and thirty-three (91.7%) of participants were right handed. A majority of participants (n = 100, 69.0%) identified as Caucasian. Twenty-two (15.2%) of participants identified as Hispanic, 10 (6.9%) as Asian, 6 (4.1%) as African American, and ~3% as American Indian, Latin American, Middle Eastern, or Pacific Islander.

Before beginning, evaluators verbally explained the consent form which included the scope of the study and the procedure. Participants were required to sign and date consent forms and initial each page. Any individual who reported hearing impairments during the consent stage of the procedure was excused from the study. Participants were randomly assigned to one of three intervention groups or assigned to the control group and this information was used for the analysis of Hypothesis II only. There were 38 controls, 34 participants in Intervention I (Neurophysiology), 40 participants in Intervention II (Metronome), and 33 participants in Intervention III (Feedback). This study was approved by the Institutional Review Board. Participation was voluntary and anonymous, and participants were treated in accordance with the "Ethical Principles of Psychologists and Code of Conduct" (APA, 2017). Participants received extra credit or research credit for participation. Both verbal and written consent was included

prior to participation and debriefing upon conclusion of the study. At the conclusion of the study participants were given a debriefing form with contact information for the principal investigators as well as student health and psychological resources.

General Procedure

Participation took place across two time-points (see Figure 6). Two rooms without clocks were used. Baseline information was collected during the first time-point. Participants first completed the neurophysiological portion of the study, which included 12 minutes of dMMN data collection. During this task, participants were instructed to indicate when they heard the tones begin, which offered an additional inspection of hearing abilities. Participants were then tested behaviorally using a verbal estimation and an interval production task. The order of the behavioral tasks was randomized by time-point and across participants. Participants then completed the LNS task to measure executive functioning and answered two surveys about use of time and time management behaviors (Bond & Feather, 1988; Britton & Tesser, 1991).

Data collected in Time 2 took place exactly 1 week following the baseline testing (see Figure 6). Participants followed the same procedure as completed in Time 1 with the following exceptions: Participants completed an intervention based on group assignment (1, 2, 3, or Control) following the first MMN. To avoid practice effects, the behavioral tasks were presented in an order which was different from that participant's Time 1 order. The participants received an additional post-intervention MMN session, which was analyzed as the final neurophysiology recording in Time 2.

Procedure: Neurophysiology

Electroencephalography (EEG) was recorded at a rate of 1000hz using SynAmps^2 system (NeuroScan 4.0, United States of America) (acquisition bandwidth = 3000 Hz) and silver

electrodes plated with silver-chloride (Diameter = 122 cm; Grass Technologies Model: F-E5SHC). A .1 Hz high-pass filter and a 200 Hz low-pass filter were used during recordings. Electrode placements followed the 10-20 system and included the following locations: *Fz, Cz, Pz*, left and right mastoids, lateral and superior of the right eye for VEOG, a ground electrode from the forehead, and a reference electrode from the tip of the nose. This is a standard set-up procedure for Eprime and Neuroscan systems. Participants were presented with 2,880 samples (120 cycles of 24 samples) of randomized tones (standard = 500 ms; deviant = 350 ms; Sound Pressure Level (SLP) = 78 dbSPL) through binaurally headphones for 12 minutes (1000 Hz, 50 ms in duration). The deviant tone was programed to each occur 8% of the time and the standard tone was presented 92% of the time.

In the passive condition, participants were told to watch a silent, closed-captioned video and ignore the tones (Davalos et al., 2003). Deviant interval durations were coded in counterbalanced blocks with one deviant per block. The deviant tone duration was 70% of the standard tone duration. The intended purpose of the selected deviant duration is to make the difference between tones not so subtle that participants vary in their ability to produce an MMN in general, but not consequently so noticeable that there was a ceiling effect of MMN (Davalos et al., 2005). Since the purpose of the study is to demonstrate changes in MMN, the deviant was selected to avoid these possible ceiling or floor effects.

Intervention I: neurophysiology. This intervention involved training on MMN directly. Administrators read the intervention directions from a standardized script (see Appendix A). Participants listened to MMN tones and use a mouse click to indicate when the participant perceives a deviant tone (time = 6 minutes). The deviant tone and standard tones that were used in the intervention were the same set of tones used in the experiment. If timing was improved at

the neurophysiology level from this intervention, then this would be indicated by larger MMN amplitudes at Time 2 compared to Time 1. There was no trial period because participants had heard the MMN tones for 12 minutes immediately before this intervention. Participants then completed the two behavioral tasks and an additional MMN trial. This resulted in 30 minutes of MMN task total in Time 2 for those in this intervention. There were at least 10 minutes between the intervention MMN and the final MMN in Time 2 due to the behavioral and applied temporal processing tasks, which eliminates any chance the participant's MMN response would habituated (Weber, Hetzel, Fahnenstich, & Lütschg, 2006).

Procedure: Behavioral

The order of the behavioral tasks (i.e., verbal estimation, interval production) was randomized by participants and differ across time-points. Intervals of auditory tones were randomized into four lists and each participant received the same list for both tasks. Intervals in the list were chosen at random to be played twice for reliability. The participants did not receive any feedback information regarding performance for either task.

Interval lists. Participants were randomly assigned to 1 of the 4 lists (see Appendix B). The lists differed between time-points for each participant to avoid practice effects. All intervals were between 1 and 50 seconds. There were 4 interval lists consisting of 10 different intervals and 2 trial intervals. Although all presented intervals were considered short intervals (<60 seconds), each list consisted of the same number of 'short' (under 10 seconds), 'medium' (11-25 seconds) and 'long' (25-50 seconds) intervals.

Verbal estimation task. Tones were created in Audacity version 2.0.5 (2014). Each interval included a start and end beep, which lasted for .1 ms. The start tone began 2 seconds into the file for every interval. Intervals were randomized into separate playlists in iTunes. In the

early pilot study, each list initially consisted of 5 trials, each played twice. Participants noticed this aspect of the design, and therefore, only 3 of 10 intervals were played twice per playlist in the study. Participants were not able to view the screen while tones were played. Researchers read the directions from a standardized script (see Appendix C). At the end of each trial, participants verbally responded with the amount of time they believed had passed between the two beeps.

Interval production task. Stimuli were presented using a computer simulator of a stopwatch (see Appendix D). The participant was instructed to recreate specified intervals on the stopwatch, which had a sheet of paper taped across the screen to block the participant's view of the clock. Administrators read directions from a standardized script (see Appendix D). At the end of each trial, the researcher recorded the participant's production of the auditory intervals. The intervals were separated into lists and the presentation order was randomized within a specified list for each participant. The participants did not receive any feedback information regarding performance for the verbal estimation or interval production tasks.

Intervention II: metronome. An online metronome set at 60 beats per minute was used to regulate the rate of finger tapping. Researchers read the intervention directions from a standardized script (see Appendix E). Participants were then trained to tap by alternating the first 2 fingers of their dominant hand along with the metronome for 90 seconds. After a 60 second break, the participants continued their training for an additional 60 seconds. Participants were directed to use the finger tapping strategy during both behavioral tasks following the intervention.

The metronome paradigm used in this study is an example of the synchronization-continuation paradigm which includes finger tapping in synchrony with metronome beats

(Kuznetsov & Wallot, 2011; Okano, Shinya, & Kudo, 2017). This paradigm is often used to find a participant's local timing correction process. Additionally, studies have used self-paced finger tapping to probe the relationship between time perception and motor timing (Delevoye-Turrell, Wilquin, & Giersch, 2012; Vanneste, Pouthas, & Wearden, 2001). Findings have demonstrated that finger tapping can improve time accuracy and temporal perception.

The main goal of the metronome intervention is to improve accuracy on the estimation and production tasks (Delevoye-Turrell et al., 2012; Vanneste et al., 2001). Previous work, however, suggests that this intervention may also impact other proposed levels of timing. Although this has not been directly assessed by MMN paradigms, studies have demonstrated that when participants are trained to coordinate their finger tapping to an interval, this exercise increases their ability to detect deviant stimuli in ERP paradigms (Kamiyama & Okanoya, 2014). Additionally, synchronization-continuation paradigms are thought to be related to a fundamental ability to coordinate, suggesting that inability to finger-tap in synchrony with beats may allude to inability to participate in unintentional action coordination (i.e. interpersonal or social coordination; Okano et al., 2017).

Procedure: Applied Temporal Processing

Executive functioning. The LNS task was administered following the behavioral tasks in both sessions and as described in the script to evaluate executive functioning, specifically working memory and attention (see Appendix F; Vuoksimaa, 2004). This task is indirectly associated with timing as the LNS task measures the cognitive functions (e.g. attention and memory) required for successful timing in an applied environment. Recordings of each trial was played through headphones to each participant to ensure standardization across trials and participants. Administers of the test abided by the discontinuation rule, such that if a participant

misses all 3 trials in 1 block then the task was ended. There was a total of 21 possible trials. A ceiling effect is possible but unlikely, as the LNS task is often used as a subtest in the WAIS-IV or intelligence quotient testing in healthy populations (Wechsler, 2008; see Appendix F). Participants also completed the Time Management Questionnaire (TMQ; adopted from Britton & Tesser, 1991) and the Time Structure Questionnaire (TSQ; Bond & Feather, 1988) to evaluate executive functioning, but specifically, perceived time management skills (see Appendix G and Appendix H). This was completed at the beginning of Time 1 after the demographic surveys and again as the last task in Session 2 (see Figure 6).

Intervention III: calibration. Participants were trained on an analogous set of intervals, which included select intervals which may appear in the verbal estimation and interval production behavioral tasks. Participants went through the above described verbal estimation and interval production tasks using a new list of intervals (see Appendix I). The researcher read the intervention directions from a standardized script (see Appendix J). Intervals were presented in a sequential order and after each trial the examiner communicated to the participant by how much they over-estimated, under-estimated the interval, or if the participant correctly estimated the interval. After the intervention, participants proceeded with the general procedure: two behavioral tasks without feedback, the LNS task, the MMN paradigm, and the two surveys (see Figure 6). This intervention was developed based on prior literature, which shows a positive relationship between feedback and time estimation accuracy (Kuznetsov & Wallot, 2011). Specifically, when feedback is included on estimation trials, participants perform more accurately and have an easier time applying this newly constructed and reorganized sense of time to longer intervals (Kuznetsov & Wallot, 2011).

Data Preprocessing for Analyses

Neurophysiology preprocessing. The recorded signals were separated into epochs (400 ms) with a 100 ms pre-stimulus interval relative to time pulses. Linear detrend was then applied. For baseline correction, the average voltage of the 100 ms pre-stimulus interval was subtracted from each signal trial (tone-onset: time = 0). All channel's in which the voltage exceeds +/- 75 μV were removed to correct for artifacts and eye blinks. A low-pass zero phase shift filter with cut offs of 30 Hz and 96 Hz were then applied. For each participant, standard and deviant evoked responses were averaged and then the standard averages were subtracted from the deviant averages to determine the MMN amplitude (see Figure 7a-c for sample data). The required minimum was set to 10 clean trials for each participant. Group average waveforms pre and post intervention are presented in Figure 8a-d for the main electrodes of interest.

The MMN amplitude is defined as the peak negative amplitude occurring between 140 and 210 ms after the onset of the deviant tone (Näätänen, 1984). Peak detection was therefore set at the minimum amplitude occurring within this window. Any positive MMN amplitudes remaining were removed. The analytical programs used for analysis are capable of handling missing data; however, participants with amplitudes that did not meet criteria for MMN in all 3 electrodes during either session were removed from analysis entirely (Time 1 $n = 9$; Time 2 $n = 13$; Both $n = 1$; see Figure 5).

Behavioral preprocessing. Participant's behavioral accuracy was computed by comparing actual duration of the interval and participant's response. Actual duration of the interval and participant's responses were collected in both the verbal estimation and interval production tasks. The formula used to analyze participant's accuracy is as follows:

$$(1)\ b_{1...10} = \left(\frac{|\,[\text{a participant's estimate or interval production of interval} - \text{actual duration of interval}]\,|}{[\text{actual duration of interval}]} \right) 100$$

(adopted from Francis-Smythe & Robertson, 1999). The absolute value demonstrates the participant's level of accuracy by assigning equal error weight to over-estimating and under-estimating the target interval. During the behavioral tasks, intervals were randomized to avoid practice effects; therefore, each participant received slightly different intervals. This was analytically accounted for by z-scoring behavioral accuracy. Each participant's percent deviance was *z*-scored against their own mean and standard deviation, then averaged across all 10 trials. This was done for both the estimation and production tasks. For interpretation, a more positive *z*-score on either estimation or production tasks reflects greater inaccuracy, a score of 0 reflects the participant's mean, and a negative score reflects how the participant's response was often close to the true interval.

Applied temporal processing preprocessing. The LNS task was scored in accordance to the test's scoring manual (see Appendix F). For each trial, a correct response merits a score of 1 and an incorrect response receives a score of 0. The final score was a sum of all completed trials (Range = 0-21). The TMQ (Britton & Tesser, 1991) has 35 items, but 18 were used in this study (see Appendix G). The possible range of scores is from 18 to 90. TMQ has responses which are coded in five categories: *always, frequently, sometimes, infrequently,* and *never.* There is reverse coding on this survey and high scores on the TMQ refers to better time management abilities. The TSQ has 26 questions, scores range from 26 to 182, and 17 questions were reverse coded (see Appendix H). Responses were recorded on a 7-point Likert scale, where 1 represents '*yes,*

this always applies to me' and 7 represents *'no, this never applies to me*.' A higher score on the TSQ reflects better use of one's time (Bond & Feather, 1988). Since the survey scores range significantly higher than the other measures of timing, all applied temporal processing variables (LNS, TMQ, and, TSQ) were converted to *z*-scores.

GENERAL ANALYSIS PROCEDURE

All variables were scaled using a *z*-score conversion, continuous, and the assumption of normality was tested prior to running any analyses (see Table 1a). The *z*-score conversion was required for scaling since few all variables existed in the same metric. Normality was assessed using kurtosis and skew values of all scaled data (see Table 1b). Kurtosis and skew values which fell between -2 and +2 are considered acceptable (George & Mallory, 2010; Gravetter & Wallnau, 2014). The estimation and production *z*-scores and the three electrodes at both time-points required an outlier test, but all values on the executive functioning tasks were within a normal range. All variables attained normal distribution following outlier removal except for the average of neurophysiological responses at Time 2 (Kurtosis = 4.28; Skew = -1.82) and the *z*-score of estimation at Time 1 (Kurtosis = 3.67; Skew = 1.66). All skew values were within the acceptable range for continuous variables (see Table 1b). Outlier removal resulted in an additional 19 participants with missing values for all three electrodes at either Time 1 or Time 2 (Time 1 $n = 8$; Time 2 $n = 11$; see Figure 5) and the final number of participants used in the analysis was 145. The kurtosis values of Time 1 estimation *z*-score and Time 2's *Cz* and *Pz* remained slightly positive, but the three electrodes were averaged which should remove the influence of nonnormality (see Table 1a and Table 1b).

The three electrodes *Fz, Cz, Pz,* were averaged to represent a measure of overall neurophysiology of timing. For the behavioral variable, each individual proportion for verbal estimation and interval production was analyzed as a percent deviance *z*-score. The variables for the applied temporal processing tasks were analyzed differently for Hypothesis I and Hypothesis II (see description of variables).

HYPOTHESIS I: ANALYSIS PROCEDURE

The direction of the relationship between neurophysiology, behavioral accuracy and applied temporal processing was tested using a mediation model in MPlus 8.1 (Muthén & Muthén, 1998-2012), with applied temporal processing as a latent variable (see Figure 3). This format tested behavioral timing as a mediator of the relationship between neurophysiological timing and applied temporal processing. Covariates included sex, years of musical training, and score on the 16-item PQ (Ising et al., 2012). To detect excellent model fit at 80% power, 113 participants were required, and 145 participants were obtained for the entire model (see Figure 9; Schoemann, Preacher, & Coffman, 2010). This is a well-cited and appropriate approach to a power analysis when using structural equation modeling (SEM).

Description of the Latent Variable

Latent variables use multiple indicators to represent a construct which would otherwise only be accessible indirectly (Muthén, 2002). In order to capture executive functioning, performance on three tasks were used as direct measurements of higher-order timing. The LNS task and the two surveys were *z*-scored and loaded onto the latent variable of applied temporal processing. All other variables were treated as described in the general analysis plan (see Figures 10a-f).

Overall Model Fit

The criteria suggested by Hu and Bentler (1999) was used to evaluate overall model fit. Model fit is considered excellent when comparative fit index (CFI) and Tucker-Lewis Index (TLI) are close to 1, the standardized root mean square residual (SRMR) is below .08, and the chi-squared test of model fit is nonsignificant. SRMR was used instead of RMSEA because

SRMR is a better measure of model fit for smaller samples (Hu & Bentler, 1999). RMSEA is frequently used for large sample sizes (~500) and at lower sample sizes, this measure becomes less accurate. A nonsignificant chi-square test of model fit indicates the model has excellent fit to the data.

Effect Sizes

Standardized betas with bias corrected bootstrapped confidence intervals were used to assess the relationship between variables. The known cutoffs for small, medium, and large effects (.1, .3, and .5, respectively) is used to assess the strength of each relationship (Cohen, 1988). Unstandardized regression coefficients are also reported in the units of the dependent variable.

Indirect Effects

To assess the strength of an indirect effect in mediation models, the regression slopes of the "a" and "b" paths are multiplied; however, this computation results in a non-normal distribution and thereby violates normality, an underlying assumption of mediation models. When assumptions are violated, the ability to accurately detect significance diminishes. Thus, asymmetrical confidence intervals (ACIs) should be used to detect significance as these types of confidence intervals best represent the underlying distribution of the indirect path. A type of ACI known as bias-corrected bootstrap intervals (Efron & Tibshirani, 1993) were used to assess measures of effects based on 1,000 bootstrapped samples (Fritz & MacKinnon, 2007). ACIs that do not contain zero at the 95% level indicate statistical significance.

According to Hayes (2009) an indirect effect of a predictor on an outcome variable through a mediating variable can occur without a significant direct effect between the predictor and outcome variables. This argument is supported by the idea that the total effect is the

summation of many paths, not all of which may be included in the current model. Therefore, indirect effects were reported in the final results regardless of the relation between direct effects and the .05 alpha level.

Covariates

Covariates were interpreted in the units of their respective dependent variable. Individuals with schizophrenia demonstrate reduced MMN responses and this effect is most easily detected using dMMN paradigms (Umbrict & Krjles, 2005; see Introduction). Those who endorse items on the PQ also have attenuated MMN amplitudes (Pantlin & Davalos, 2016). Therefore, PQ scores were treated as a covariate of MMN amplitude. The PQ score is a binary variable coded as 1 = control and 2 = endorsing symptoms of psychosis. Any participant who scores above the established cut off score of 6 on the PQ was given a value of 2 (Ising et al., 2012). There was an expected positive relationship between PQ score and MMN responses (see Figure 10d & e).

Those who have musical training tend to perform better on timing tasks which are considered automatic, such as MMN (Rammsayer & Altenmüller, 2006). Musical training was also analyzed as a covariate of neurophysiological timing abilities. Years of musical experience was coded as an ordinal variable in which 1 represents no experience (n = 73, 44.5%), 2 represents 1-3 years of experience (n = 21, 12.8%), 3 represents 3-5 years of experience (n = 41, 25.0%), 4 represents 5-10 years of experience (n = 2, 1.2%), and 5 represents more than 10 years of experience (n = 26, 15.9%).

Sex differences have been found in the LNS task and in both time management surveys (Macan et al., 1990; Vuoksimaa, 2004). Sex was coded as a binary variable with a 1 representing *male* and a 2 representing *female*. Women tend to show better time management

skills whereas men tend to score higher on the LNS task; therefore, sex was treated as a covariate of the latent variable applied temporal processing (see Figure 3; Macan et al., 1990; Vuoksimaa, 2004).

HYPOTHESIS I: RESULTS

Model 1: Proposed Model

Overall Model Fit. Model fit indices overall were poor (see Table 2). CFI and TLI values were not close to 1, SRMR was greater than .08, but the chi-squared test of model fit was significant due to the small sample size ($\chi^2(16) = 35.26$, $p = .00$). Together, this information indicated poor overall fit and thus, results should not be interpreted. The factor loadings of the latent variable were nonsignificant and uninterpretable due to poor model fit (TMQ: b = .02, SE = .04, $p = .54$; TSQ: b= -.07, SE = .05, $p = .11$) (see Table 2). Nonsignificant factor loadings for the LNS task and both of the time surveys revealed that change in TMQ was not correlated with change in TSQ or the LNS task. Additionally, model results indicated that averaging behavioral tasks diluted the overall effect of behavioral accuracy on the final outcomes. Model fit was improved when the estimation and production tasks were analyzed in separate models.

Figures 4a-c represent the new model where each of the three applied temporal processing tasks are analyzed in three separate mediation models. The reliability of the task was loaded as a single factor for each of the three new models (see Figures 4a-c). By analyzing the model in this fashion, measurement error is accounted for in the mediation model and the latent variable is considered to be normally distributed. Estimation and production tasks were analyzed in two separate codes to avoid saturation of the final model.

Single Indicator Mediation Models

Overall model fit. Model fit was poor when single indicator models were run with estimation and production tasks in the same model (see Table 3); therefore, models were run as shown in Figures 4a-c. All models retained nonsignificant chi-squared tests for both estimation

and production tasks, all CFI and TLI values were equal to 1, and all SRMR values were less than .08 (see Table 3). Therefore, based on the established criteria for model fit, all models were in the excellent range. A scored response on the PQ and years of musical experience were evaluated as covariates to neurophysiological response. PQ responses was also a covariate of average amplitude. Sex was removed as a covariate in the final models due to large, negative residual variances. When this covariate was removed, model fit improved and there were only positive residual variances; thus, sex was not an appropriate covariate for this model (Muthén & Muthén, 2002).

LNS. The LNS task was evaluated as a latent variable with reliability as a single indicator (Fisher *z* calculated with transformation = .82; Silva, 2008). The two behavioral tasks, estimation and production, were run as the mediator of the relationship between neurophysiological responses and LNS scores in two separate models due to model fit outcomes (see Table 3).

Direct effects. There were no significant relationships in the single indicator models analyzing the LNS task, but models using estimation as the mediator followed the expected directionality. A 1-unit change in LNS scores was related to a -.09 unit change in average MMN response, but this effect was not significant (b = -.09, SE = .23, [-.57, .36]. The standardized betas revealed that this trend had a minor effect (β = .07). Scores on the LNS task were also negatively, but not significantly, related to behavioral accuracy (b = -.10, SE = .59, [-1.40, 1.00]). The standardized betas for this relationship revealed that this trend had a trivial effect (β = .03). The effect of neurophysiological responses and accuracy on the estimation task also had a minor effect (β = .03), but retained a positive, nonsignificant relationship (b = .01, SE = .04, [-.07, .10]) (see Table 4). The LNS single indicator model with the production task did not maintain the overall expected directionality or any significant effects but is reported in Table 4.

Covariates. As expected, those with higher scores on the PQ performed worse on the LNS task (b = -.05, SE = .08, [-.19, .00]) and had smaller neurophysiological responses (b = .002, SE = .03, [-.05, .06]), but these relationships only retained the expected directionality and did not meet criteria for significance. The relationship between PQ responses and neurophysiological responses had a trivial effect (β = .01). Years of musical experience had a small effect (β = .1) but did not significantly relate to neurophysiological responses (b = .13, SE = .14, [-.16, .41]) (see Table 4).

Indirect effects. Indirect effects demonstrated a nonsignificant relationship using bias-corrected bootstrapped intervals at the .05 level (b = -.001, SE = .29, [-.07, .04]). The directionality of the indirect effect was as expected, but this was a trivial effect (β = .001). The indirect effect of the production model was nonsignificant and did not follow the expected directions (see Table 4).

TSQ. The TSQ was evaluated as a latent variable with reliability as a single indicator (Test re-test reliability = .76; Bond & Feather, 1988). Estimation and production tasks were each analyzed as the mediator of neurophysiological responses and TSQ scores in two separate models (see Table 3 for overall model fit).

Direct effects. There were no significant relationships and only the models using estimation retained the expected directionality. All information for TSQ single indicator models and the production task are reported in Table 5. Standardized estimates of effect size revealed that all relationships were trivial (a-path: β = .03; b-path: β = .002; c': β = .06). Unstandardized estimates showed that those who retained high scores on the TSQ had high accuracy on the estimation task (b = -.004, SE = .29, [-.66, .47]) and larger neurophysiological responses (b = -.06, SE = .10, [-.24, .14]),but this effect was not significant. Trends also showed that those with

larger MMN responses had accurate estimation of time intervals (b = .01, SE = .04, [.-07, .10]) (see Table 5).

Covariates. None of the covariates had significant relationships with neurophysiology or TSQ. PQ scores did not affect the neurophysiological response waveforms (b = .01, SE = .02, [-.04, .05]; β = .04). Musical experience also did not significantly relate to strength of neurophysiological responses (b = .13, SE = .14, [-.16, .03]; β = .1). All covariate effects had small effect sizes (see Table 5).

Indirect effects. Bias-corrected confidence intervals at the .05 level revealed a nonsignificant relationship between the indirect effects and total effect for both the estimation (b = 0, SE = .01, [-.04, .03] and production (b = -.01, SE = .03, [-.08, .05]) tasks (see Table 5). The standardized betas were very small (Estimation: β = .00; Production: β = .01).

TMQ. The TMQ was not analyzed in the same fashion as the previous applied temporal processing tasks. There were no reports of test-retest reliability in the literature, but only a value for Cronbach's alpha (Cronbach's alpha = .87; Alay & Koçak, 2002). Since use of internal consistency measures is inappropriate for measurement error models, the model was explored in two ways: first, using a value of .1, which assumes good reliability (see Table 6a) and second, using the observed score of TMQ (see Table 6b). The differences among the results of the observed score and assumed reliability model were miniscule. Therefore, the model using the observed score for TMQ was evaluated in the results.

Direct effects. The results of the single indicator models using TMQ followed the same findings as the previous measures: there were no significant findings and only the models using the estimation task displayed directionality which aligned with the hypothesis. All standardized and unstandardized betas for the production task are reported in Table 6b. Trends suggested that

participants with larger neurophysiological responses also scored higher on the TMQ (b = -.02, SE = .07, [-.16, .11]; β = .03) and had higher accuracy on the estimation task (b = .01, SE = .04, [-.07, .10]; β = .03). There was also a nonsignificant trend, suggesting that accurate responses on the estimation task related to higher scores on the TMQ (b = -.15, SE = .20, [-.60, .18]; β = .01) (see Table 6b).

Covariates. All covariates had small, standardized, and nonsignificant effects on neurophysiology. The covariate PQ had the expected directionality in relation to neurophysiological responses (b = .06, SE = .02, [-.05, .05]; β = .03). Years of musical experience did not portray the expected directionality with neurophysiological responses (b = .13, SE = .14, [-.16, .14]; β = .01) (see Table 6b).

Indirect effects. The indirect effect of estimation and neurophysiology demonstrated the expected directionality but did not have a significant effect (b = .00, SE = .01, [-.03, .012]; β = .00). The model using production as a potential mediator did not retain significance or the expected directionality (see Table 6b).

HYPOTHESIS I: DISCUSSION

The goal of the first hypothesis was to develop a relationship between the tasks most often used to measure the various domains of timing. Tasks were selected to evaluate three proposed levels of timing (neurophysiology, behavioral, and applied). Level I, neurophysiology, was evaluated by measuring brain responses to stimuli that differ. Behavioral timing represented the second level of the proposed mediation and was measured using two tasks: interval estimation and interval production. Lastly, two surveys (TMQ and TSQ) and the LNS task were used as measures of applied temporal processing (Level III). Based on the proposed mediation, the expectation was that neurophysiological timing (Level I) would inform behavioral timing abilities (Level II). These levels would subsequently relate to timing in an applied setting (Level III); however, the results did not support these relationships.

Neurophysiology and Behavioral Timing

In Hypothesis I, it was anticipated that higher accuracy on the estimation and production tasks (Level II) would relate to larger responses on the neurophysiological timing task (Level I). Although neither behavioral task was significantly related to neurophysiology, estimation demonstrated trends in the expected direction. These trends suggested that those who have a larger neurophysiological response would also be able to accurately track elapsed time between two stimuli. Neurophysiology responses to the MMN paradigm are thought of as the earliest indication of detecting changes in our temporal environment (Näätänen et al., 2007). This response represents a fundamental ability to time in milliseconds. Thus, it is plausible to assume that our ability to time on a longer timescale, such as in the behavioral tasks, is dictated by our ability to time on a smaller timescale, such as in neurophysiological responses (Level I). In other words, early indices of timing may be built upon to produce an accurate reading on elapsed time.

The nonsignificant relationship between neurophysiology (Level I) and behavioral timing (Level II) may have been caused by large gap between the timescales. Between Level I and Level II, the timescale under evaluation rose from milliseconds to seconds. Both behavioral tasks (estimation and production) measure timing ability in a range that exceeds a second but does not exceed a minute while neurophysiology measures timing in the milliseconds range. Thus, Level I neurophysiology is thought of as more of an automatic response which contrasts with the behavioral tasks, where higher-order cognition (e.g. attention, working memory, and decision-making) is required to successfully and accurately perform the behavioral tasks (Gibbon et al., 1984; Mioni, 2018; Näätänen et al., 2007).

Previous work has supported that cognitive resources, such as attention, working memory, and decision-making, are required for accurate behavioral timing unlike the requirements for accurate neurophysiological timing (Mioni, 2018). To complete the estimation task, an individual must attend to the interval in order to track its duration, then hold that duration in working memory, retrieve a memory about previous experiences of that duration for comparison, then lastly decide about the duration that has passed. Thus, estimation and production not only measure the ability to time in the seconds range, but inadvertently evaluate a person's working memory capacity, attention, and decision making. This idea is supported by the Gibbon et al. (1984) Scalar Expectancy Model, representing the ways that timing and cognitive resources are intertwined. The timescale difference from milliseconds (Level I), which does not have a high cognitive demand, to seconds (Level II), which requires additional working memory and attentional resources, may have been too distant.

The present study laid down the initial outline of the time continuum, but since the first two levels were not significantly related, careful evaluation of intermediate stages are required to

fully understand timing (see Figure 11). These intermediate stages could be measured using psychophysical tasks (Bausenhart, Luca, & Ulrich, 2018). These tasks would allow the researcher to evaluate a participant's timing ability between the timescales analyzed in Level I (neurophysiology) and Level II (behavioral timing). Two examples of psychophysical tasks that could be used in future studies are the discrimination and identification tasks. The discrimination task requires participants to distinguish between two stimuli and has been successfully used with intervals ranging from 100-1480 ms (Matthews & Meck, 2014). The identification task, in which participants are required to classify durations, also measures the psychophysical intermediate stages of timing, but is limited by the range of stimuli that can be used. However, foundational work by Miller (1956) found that overall participants can distinguish (discrimination task) between more stimuli than they are able to recognize (identification task). This suggests that, like the estimation and production tasks, these two psychophysical tasks are two sides of the same coin in that both employ similar temporal mechanisms, such as the internal clock (Block, 1990; Treisman, 1963). In other words, the data collected from these psychophysical tasks would offer a window into an individual's ability between neurophysiological and behavioral timing by isolating their internal clock (Treisman, 1963, see Figure 1).

The internal clock as theorized by Treisman (1963) can be evaluated using psychophysical tasks, such as the identification and discrimination tasks. These tasks are thought to capture if a person's internal clock or subjective experience of time is initially accelerated or decelerated (Mioni, 2018). The psychophysical or intermediate stage between Level I (neurophysiology) and Level 2 (behavioral timing) provides the opportunity to evaluate timing before additional cognitive resources are required for accurate timing. Adding these additional

psychophysical measures to analyze an intermediate stage may provide a way to further deconstruct the time continuum and offers an explanation for the null results.

Behavioral Timing and Applied Temporal Processing

The expectation was that higher scores on the two surveys, TMQ and TSQ, and the LNS task (Level III) would be related to better accuracy on the estimation and production tasks (Level II). A high score on the LNS task reflects good working memory and attention, two cognitive abilities that are indirectly related to timing (Gibbon et al., 1984, see Figure 2). The ability to accurately respond on the estimation and production tasks require the same cognitive skills that are measured by the LNS task. Although neither the estimation nor the production task was significantly related to LNS score, estimation demonstrated a trend. To successfully complete the estimation task, a participant must attend and engage their working memory. Therefore, it is plausible that individuals who have better auditory working memory and attention, as indicated by a higher score on the LNS task, would also perform more accurately on the estimation task.

Although the estimation task revealed an interesting trend, neither behavioral task was significantly related to the LNS task. This is most likely due to the limited variation that was observed in the sample population. Participants in the current study were strictly university students, who are likely to retain better executive functioning capacity compared to the general population. Students are constantly receiving feedback on their working memory and attention abilities (e.g. exams, essays) and are challenged to implement strategies that improve cognition to meet the demands of university life. Additionally, there are entire fields of research devoted to supporting students and improving their working memory because of the intense demand of attention and working memory in school environments (Alloway, Gathercole, Kirkwood, & Elliot, 2008; Gropper & Tannock, 2009; Rogers, Hwang, Toplak, Weiss, & Tannock, 2011)

Thus, it can be deduced that a student, who receives a LNS score that is comparatively lower than his or her peers, is potentially still operating at a level that is on par with the average member of the general population. The average standard deviation of LNS scores in the current study supports this claim as participants did not waiver from one another by more than a point. This limited variation is largely responsible for the nonsignificant effects found between the behavioral tasks and the LNS task.

While the LNS task offers an indirect measurement of timing, the two surveys, TSQ and TMQ, were used to isolate timing ability at the applied level. These surveys did not have a significant relationship with accuracy on the behavioral tasks. The estimation task presented a trend which suggested that those who were effective time managers, as indicated by high survey scores, had a propensity to accurately estimate interval duration. This trend was expected as estimation is associated with the first stage of time management: determining task duration and creating a plan to allocate time to each task accordingly (Francis-Smythe & Robertson, 1999). To complete this stage, a participant must retain a fundamental understanding of the duration. Participants often are capable of estimating elapsed time, as outlined by their performance on the estimation task. Thus, inaccuracy at this stage is less often due to misunderstanding the duration and more often due to overestimating their own abilities, supporting why the estimation task presented a trend with the surveys and the production task did not.

Neither of the behavioral tasks (Level II) were significantly related to the TSQ or TMQ surveys (Level III), yet unlike the estimation task, the production task did not demonstrate a trend. The estimation task relates most closely to creating an agenda based on estimations of task duration, whereas the production task relates to the carrying out of said tasks. Often, self-efficacy and accuracy in estimating the duration of a task is misaligned (Francis-Smythe & Robertson,

1999). Past work has demonstrated how participants are often over-ambitious in their self-assessments of their own abilities (e.g. the planning fallacy, overconfidence bias, etc.) (Sanna, Parks, Chang, Carter, 2005). Thus, the production task did not present a trend because it is primarily related to the second stage of time management: maintenance of a planned schedule (Francis-Smythe & Robertson, 1999). Inability to maintain a planned schedule or inaccurate applied timing can be caused by either an overestimation or underestimation of one's own abilities. For example, an individual may anticipate a task taking five minutes, but miss his or her self-imposed deadline when the task takes longer. On the contrary, an individual who anticipates a task taking him or her an hour but can complete the task in less time demonstrates inaccurate timing as well. To empirically disentangle differences in applied timing, future studies should aim to develop a battery that includes various time management tasks along with self-efficacy scales, such as the surveys used in the present study, to capture time perception at an applied level (Level III).

Covariates

The PQ and years of musical experience were both analyzed as covariates of neurophysiological responses, but neither produced a significant relationship. Although the effect was trivial, those who had high scores on the PQ also tended to have worse neurophysiological responses. Previous work suggests that those who endorse symptoms of psychosis should have attenuated neurophysiological amplitudes compared to controls (Pantlin & Davalos, 2016). The other covariate, musical experience, had no relationship with neurophysiological responses. Previous work suggests that retaining experience with music (e.g. dance, song, and/or instruments) is related to differences in neurophysiological responses (Petersen et al., 2015; Rammsayer & Altenmüller, 2006). Specifically, one study noted that only highly trained

musicians would show changes in the temporal environment as measured by neurophysiological amplitudes (Rammsayer & Altenmüller, 2006). The present study only solicited information about the type and the number of years of experience to explore this under-researched relationship, which may explain the nonsignificant relationships. Individuals in the present study may have differing ideas about the defining elements that constitute as a year of musical experience. These differences would not have been reflected in their survey responses and therefore, were not able to be addressed in ad-hoc analyses.

Summary

This was the first study that attempted to compare performance on neurophysiological measures (Level I) with behavioral tasks (Level II), and then relate this to performance in an applied settings (Level III). While the current study did not show a clear relationship between the proposed levels of time processing, trends paved the road for a better understanding of how time is handled in healthy individuals. The boundaries of the first two levels (neurophysiological and behavioral) may have been inadvertently overstretched. More attention to potential intermediate stages (e.g. psychophysical tasks) between these two subdomains is required if a continuum is to be outlined. The nonsignificant findings were in part accounted for by the minimal variation in the tasks across participants. Yet, this was a powered study which insinuates that these effects would still exist even with additional participants. Findings of the present study have also set the stage for the development of a time management battery, which should include both indirect and direct measures of timing. While none of the relationships reached significance, this study provided the initial groundwork for unpacking how time is handled in healthy individuals and trends provide a promising trajectory for future studies to follow and further deconstruct the time continuum.

HYPOTHESIS II: ANALYSIS PROCEDURE

Hypothesis II is as follows: within a specified condition, timing would increase at the level targeted by the specific intervention across time-points. Those in Intervention I (MMN) would demonstrate an increase in neurophysiology across time-points. Those in Intervention II (Metronome) would demonstrate an increase in behavioral accuracy from Time 1 to Time 2 and those in Intervention III (Calibration) would demonstrate an increase in applied temporal processing measures across time. Controls should not demonstrate a significant change across time compared to intervention groups (see Figures 4a-c and Figure 12).

Hypothesis II was examined using 4x2 repeated-measures analyses of variance (RM-ANOVA) in IBM SPSS Statistics 25 (SPSS Inc., 2017). Each RM-ANOVA assessed change in each subdomain, neurophysiology, behavior, applied temporal processing, across time, separated by intervention. Main effects for change across time, intervention and the interaction of time and intervention was assessed in each RM-ANOVA using significance values, partial eta-squared, and estimated marginal means.

Partial eta-squared values are interpreted as the statistical representation of the amount of variance explained by each variable in the model. Additionally, estimated marginal means are produced by the RM-ANOVA and represent new values that are adjusted for the variance attributed to the covariate; therefore, estimated marginal means were also interpreted. To retain 80% power, 116 participants were needed for a small-medium effect to be detected (~.15) and 145 were used in the final analysis.

Assumptions and Variables

All assumptions of RM-ANOVA were met after required adjustments. The distribution of each variable was tested, and variables were continuous. Time was assessed as an independent variable, which was statistically tested for independence of observations and errors. All significant outliers were removed from the dataset in the final model. Unequal variances across dependent variables was assessed using the Levene's Test of Equality of Error Variances as opposed to Mauchly's test of sphericity as there were only two levels of repeated measures (Time 1, Time 2). All RM-ANOVAs maintained a nonsignificant Levene's Test (see Tables 7a-c) except for the estimation task and the TMQ at Time 2 only. Thus, the Greenhouse-Geisser statistic was used for interpreting these models as this offers a more stringent assessment of significance to avoid a Type I error.

Box's Test of Equality of Covariance Matrices was used. In this test, a significant statistic suggests that the variable has violated the assumption of equal covariances, is heavily affected by group size, and increases the likelihood of a Type 1 error (Box, 1949; Cohen, 2008). All RM-ANOVAs maintained a nonsignificant Box's Test of Equality of Covariance Matrices (Neurophysiology: $F(9, 207377) = 17.03$, $p = .056$; Applied Temporal Processing: $F(63, 24082.66) = 69.05$, $p = .52$) except for the two behavioral tasks ($F(30, 43825.02) = 52.96$, $p = .013$).

HYPOTHESIS II: RESULTS

Neurophysiology

Main effects. All means and standard deviations are reported in Table 8. Across sessions, participants' neurophysiology was significantly affected ($F(1,138) = 4.88$, $p = .03$), suggesting that overall participants' neurophysiological responses were increased at Time 1 compared to Time 2. The main effect of intervention was significant ($F(1,138) = .29$, $p = .006$), suggesting there was a difference in neurophysiology across the three intervention groups and controls. The covariates also maintained a significant main effect (PQ: $F(1,138) = .92$, $p = .007$; Musical experience: $F(1,138) = .27$, $p = .002$) (see Table 9).

Interactions. Although the main effects for all variables retained significance, the interactions of each variable across sessions did not. The differences in amplitude between Time 1 and Time 2 was not due to the intervention ($F(3, 138) = 1.69$, $p = .17$; see Table 9) nor the potential covariates (PQ: $F(1,138) = 3.12$, $p = .08$; Musical Experience: $F(1,138) = .291$, $p = .59$) (see Table 10).

Effect size. Table 10 reports partial eta-squared values and shows the amount of unexplained variance accounted for by each covariate and variable. The interaction of time and intervention explained 3.5% of the variance in the model, but this was not significantly more than time alone (partial $\eta^2 = 3.4\%$). Out of the covariates, the interaction of time and responses on the PQ (partial $\eta^2 = 2.2\%$) explained more variance than time and years of musical experience (partial $\eta^2 = .2\%$).

Comparison of interventions. The effects of time and intervention were evaluated using estimated marginal means to control for covariates (see Tables 11a-c). Pairwise tests were used

to compare estimated marginal means across time and intervention separately (see Tables 12a-b). Pairwise tests were not examined for time and intervention because this interaction was not significant.

Main effect: Time. Estimated marginal means revealed that all participants had a lower neurophysiological response at Time 1 (Mean = -1.49, SE = 0.09) compared to Time 2 (Mean = -1.31, SE = 0.08), after controlling for PQ scores and years of musical experience (see Table 11a). However, this attenuation was not significant (Mean difference = -.18, SE = .12, p = .09, [-.39, .03]; see Table 12a).

Main effect: interventions. After controlling for covariates, those in the neurophysiology intervention had the largest mean response amplitudes (Mean = -1.46, SE = 0.13) where as those in the metronome intervention had the weakest responses (Mean = -1.31, SE = 0.13; see Table 11b). Pairwise comparison of estimated marginal means revealed that there were no significant differences of neurophysiological responses between interventions (see Table 12b).

Interaction: time and intervention. After controlling for covariates, all interventions demonstrated an attenuation in responses between Time 1 and Time 2, except those in the neurophysiology intervention (Time 1: Mean = -1.40, SE = .19; Time 2: Mean = -1.52, SE = .15; see Table 11c and Figure 13). The metronome intervention demonstrated the largest decrease in neurophysiological response across time (Time 1: Mean = -1.56, SE = .18; Time 2: Mean = -1.06, SE = .14). Only descriptive statistics are reported because the interaction of time and intervention did not reach significant. Therefore, there were no pairwise comparisons to analyze.

Behavioral Timing

Main effects. Descriptive statistics for the estimation and production tasks are presented in Table 13. The main effect of time on the two behavioral tasks were both significant (Estimation: $F(1, 128) = 17.52$, $p = .00$; Production: $F(1, 128) = 14.16$, $p = .00$) (see Table 14).

Interactions. The interaction of intervention and time each had a significant effect on production: $F(3, 128) = 2.73$, $p = .046$), but not on estimation ($F(3, 128) = 1.99$, $p = .12$) (see Table 14).

Effect sizes. For production, the variable time explained 10% of the variance in the model and the interaction of time and intervention explained 6% (see Table 14). For estimation, 12% of the variance was explained by time and 4.4% of the variance was explained by the interaction of time and intervention (see Table 14).

Comparison of interventions. Estimated marginal means and pairwise tests were evaluated to compare the effect of each time point and intervention. Pairwise comparisons were evaluated only for the production task since the interaction of time and intervention was significant. This interaction was not significant for the estimation task, thus, pairwise tests were not examined for time and intervention interaction.

Main effect: time. Overall accuracy on the estimation task increased across sessions (Time 1: Mean = -0.06, SE = 0.04; Time 2: Mean = -0.22, SE = 0.03) as did production (Time 1: Mean = 0.36, SE = 0.13; Time 2: Mean = -0.17, SE = 0.09; see Table 15a). Pairwise comparisons of estimated marginal means indicated that participants significantly became more accurate on behavioral tasks across time, regardless of intervention (Estimation: Mean difference = .53, SE = 0.1, $p = 0.0$, [0.252, 0.81]; Production: Mean difference = .15, SE = 0.04, $p = 0.0$, [0.08, 0.23]; see Table 16a).

Main effect: intervention. An analysis of pairwise comparisons also showed that those in the calibration intervention (EMM = -.181, SE = .18) retained a significantly higher overall behavioral accuracy on the production task than those in the metronome intervention (EMM = .34, SE = .17; Mean difference = .52, SE = 0.25, p = 0.04, [-1.02, -0.02]; see Table 15b & Table 16b). There were no other significant differences between interventions; however, estimated marginal means revealed that those in the calibration intervention retained the highest mean accuracy on estimation (EMM = -.19, SE = .05) and production (EMM = -.18, SE = .18) tasks compared to all other interventions. Those in the metronome task had the lowest accuracy on both tasks (Estimation = .338; Production = .09), followed by controls (EMM = .18, SE = .17; see Table 15b).

Interaction: time and intervention. All intervention groups demonstrated an increase in accuracy for both behavioral tasks (see Table 15c). Controls did not demonstrate an increase in accuracy on the production task across time (Time 1: Mean = 0.211, SE = 0.26; Time 2: Mean = 0.14, SE = 0.18; see Table 15c), but did for estimation (Time 1: Mean = -0.09, SE = 0.07; Time 2: Mean = -.16, SE = 0.05; see Table 15c). The calibration and neurophysiological interventions had the highest estimated marginal means accuracy on both behavioral tasks at Time 2, compared to all other interventions.

To compare performance on the production task across time and intervention, the data was pulled into a regression framework and simple slopes were examined (see Figure 14). Production had a significant interaction (time*intervention) and was therefore, the only behavioral task examined in this way. Based on the negative simple slopes presented in Figure 14, all participants improved their mean accuracy on the production task across sessions. As expected, the control group demonstrated the least change across sessions, indicated by the

smallest slope (y = -.35x + .84). Controls had higher accuracy on production at Time 1 compared to all groups except the calibration intervention, but at Time 2 retained the lowest mean accuracy score. The calibration intervention retained the highest accuracy and the largest improvement on the production task across both sessions (y = -.98x + 1.37). Those in the metronome (y = -.8x + 1.55) and neurophysiological (y = -.84x + 1.58) interventions followed similar trends, but those in the neurophysiology intervention showed slightly more improvement on the production task during the second session, but this difference was trivial (see Figure 14).

Applied Temporal Processing

Main effects. All descriptive statistics for the applied temporal processing tasks are presented in Table 17. The main effect for time (TMQ: $F(1, 104) = .33$, $p = .57$; TSQ: $F(1, 104) = 2.73$ $p = .10$; LNS: $F(1, 104) = .26$, $p = .61$), sex (TMQ: $F(1, 104) = .90$, $p = .34$; TSQ: $F(1, 104) = .01$, $p = .03$; LNS: $F(1, 104) = .64$, $p = .43$), and intervention (TMQ: $F(3, 104) = .77$, $p = .51$ TSQ: $F(3, 104) = 2.02$, $p = .12$; LNS: $F(3, 104) = .39$, $p = .76$;) were not significant (see Tables 18 and 19). The covariate PQ responses demonstrated a significant main effect on TSQ ($F(1, 104) = 5.01$, $p = .03$), but not on TMQ ($F(1, 104) = 1.34$, $p = .25$) or LNS ($F(1, 104) = .64$, $p = .43$) (see Table 18).

Interactions. The interaction of time and sex had nonsignificant relationships with all three applied temporal processing tasks (TMQ: $F(1, 104) = .07$, $p = .79$; TSQ: $F(1, 104) = .12$, $p = .73$; LNS: $F(1, 104) = .00$, $p = .99$). The PQ and time interaction also had no significant relationships (TMQ: $F(1, 104) = .11$, $p = .74$; TSQ: $F(1, 104) = 2.01$, $p = .16$; LNS: $F(1, 104) = .00$, $p = .95$). The interaction of time and intervention had a significant interaction with TSQ ($F(3, 104) = 3.66$, $p = .02$), but not with TMQ ($F(3, 104) = .34$, $p = .80$) or the LNS task ($F(3, 104) = .95$, $p = .95$) (see Table 19).

Effect sizes. The interaction of time and intervention with TSQ explained the most variance in the model (partial $\eta^2 = 9.6\%$) followed by the LNS task (partial $\eta^2 = 2.7\%$). The interaction of the covariate sex and time explained the least amount of variance for TMQ (partial $\eta^2 = .1\%$), TSQ (partial $\eta^2 = .1\%$), and scores on the LNS task (partial $\eta^2 = 0\%$). The interaction of time and PQ also had a minimal amount of variance explained for TMQ (partial $\eta^2 = .1\%$) and the LNS task (partial $\eta^2 = 0\%$) (see Table 19).

Comparison of interventions. Each timepoint and intervention was evaluated and compared using estimated marginal means pairwise comparison tests. The model examining TSQ retained a significant interaction. Therefore, the TSQ model was the only reported pairwise comparison of the interaction (time and intervention).

Main effect: time. The LNS task had lower scores at Time 1 (Mean = 6.29, SE = 0.24) compared to Time 2 (Mean = 6.86, SE = 0.23; see Table 20a). The difference across sessions, controlling for PQ scores and sex, was significant (Mean difference = -.57, SE = 0.22, $p = 0.01$, [-1.01, -0.13]; see Table 21a). Both time management surveys also had demonstrated an increase between sessions (see Table 20a), but this effect was not significant (TMQ: Mean difference = -0.34, SE = 0.47, [-1.28, 0.60]; TSQ: Mean difference = -1.29, SE = 0.70, [-2.68, 0.10]; see Table 21a).

Main effect: intervention. Overall those in the metronome intervention had the highest LNS estimated marginal mean score of 6.84 (SE = .41) but score the lowest on the TMQ (Mean = 49.58, SE = 1.49; see Table 20b). The calibration intervention had the lowest LNS mean score (Mean = 6.20, SE = .45), the highest TMQ mean score (Mean = 52.84, SE = 1.62), and the lowest TSQ mean score (Mean = 88.82, SE = 2.06), controlling for the covariates (see Table 20b). The neurophysiology intervention group had the highest TSQ estimated marginal means

(Mean = 94.78, SE = 1.79). However, none of the differences were significant except for the mean difference between the neurophysiology and calibration interventions on the TSQ (Mean difference = 5.96, SE = 2.74, p = 0.03, [0.52, 11.40]). This demonstrated that those in the neurophysiology intervention did significantly better on the TSQ than those in the calibration intervention (see Table 21b).

Interaction: time and intervention. Table 20c displays all estimated marginal means for each applied temporal processing task by intervention. All interventions increased their respective LNS scores across time, except for those in the calibration intervention (Time 1: Mean = 6.26, SE = 0.52; Time 2: Mean = 6.15; SE = .50; see Table 20c). This was also the case for the TMQ, in which those in the calibration intervention again did not increase their scores across time (Time 1: Mean = 53.03, SE = .67; Time 2: Mean = 52.66; SE = 1.74; see Table 20c). Those in the neurophysiology intervention were the only ones to not increase their scores across sessions (Time 1: Mean = 95.76, SE = 1.95; Time 2: Mean = 93.79, SE = 1.87; see Table 20c).

The main effects and interactions of time and intervention were not significant when the TMQ or the LNS task was analyzed. Since the TSQ retained a significant interaction (time and intervention), a comparison of simple slopes is reported below and shown in Figure 15. To increase interpretability, the predicted raw scores were used in this figure. Participants in the calibration and control groups reported they were more efficient users of time at Time 2, indicated by the negative simple slopes. Those in the metronome intervention retained the highest TSQ scores during both sessions but did not show much change in scores across sessions ($y = -0.94x + 95.56$). The control group demonstrated the most change across time on the TSQ survey ($y = 2.25x + 88.3$) and the neurophysiology group showed the largest decrease in scores ($y = -1.77x + 95.76$); however, the change in scores did not fluctuate by more than 2 points on

average. The calibration intervention reported the lowest mean scores on the TSQ at Time 1 but showed approximately a 2-point increase on average by Time 2 ($y = 2.15x + 86.56$; see Figure 15).

HYPOTHESIS II: DISCUSSION

Hypothesis II examined if time accuracy, as measured by the various tasks, could be systematically improved through use of interventions. These interventions were developed to specifically target each of the proposed levels of timing. First, participants underwent all tasks at Time I. Performance at Time 1 was used as a baseline measure and was ultimately compared to performance on the same tasks after completing an intervention in a one-week follow-up session.

There were three interventions aimed to improve accuracy at each level: Neurophysiology (Level I), behavioral timing (Level II), and applied temporal processing (Level III). Intervention I (Neurophysiology) required participants to attend to and indicate when they heard a tone that differed from the pattern. This intervention was intended to directly target timing at Level I (Neurophysiology). The second intervention (Metronome) targeted Level II (Behavioral Timing) and had participants tap along with a metronome. Participants in Intervention II (Metronome) were then instructed to use this finger tapping strategy when completing the behavioral tasks during the second session. The last intervention was the calibration intervention, which aimed to improve timing at the applied level (Level III). In the calibration intervention (Level III), participants received feedback on their performance during the behavioral tasks. RM-ANOVAs were conducted to test the effectiveness of the interventions by comparing pre- and post-intervention performance on timing tasks. The interaction of time and intervention was not significant in any of the models, suggesting that the interventions did not support the hypothesis. However, the trends in each of these analyses are reported below and yield important considerations for the field.

Neurophysiology

The expectation was that individuals in the neurophysiological intervention would show the largest improvement in the associated task across time while the control group would show comparatively less change in accuracy. Larger negative responses at Time 2 compared to Time 1 would have supported Hypothesis II, but this increase in responses was not observed in the data even after controlling of the covariates. Overall, neurophysiological amplitudes were closer to zero or worse at Session 2 compared to Session 1. However, when this information was separated by intervention and session, estimated marginal means revealed that those in the neurophysiological intervention were the only participants to demonstrate an improved response across time. Although the anticipated directionality was observed, this effect was trivial and nonsignificant.

Those who were directly trained on MMN paradigm (Intervention I: Neurophysiology) were expected to demonstrate the largest change in amplitude; however, no change was evident, providing further support for MMN as a pre-attentive biomarker of pathology (Light & Näätänen, 2013; Shelley et al., 1991). Previous work suggesting use of the MMN paradigm in clinical populations and as a biomarker rests upon the idea that MMN can be produced without the influence of attention (Garrido et al., 2009; Sculthrope et al., 2009). These arguments suggest that neurophysiological responses can be induced without the individual specifically paying attention to the stimuli in the paradigm (MMN). The participants in the present study who attended to the deviant stimuli, or those in the neurophysiological intervention, did not retain significantly better neurophysiological responses compared to other participants. Thus, these findings support the claim that attention is not a modulator of the MMN response (Campbell & Davalos, 2015; Garrido et al., 2009; Näätänen et al., 2001; Sussman et al., 2002).

Behavioral Timing

The two behavioral tasks, verbal estimation and interval production, were analyzed as separate measures within one RM-ANOVA. The expectation was that those in the metronome intervention would show the highest accuracy at Time 2 in both behavioral tasks. This was not supported by the final analysis. Based on the means, those in the calibration intervention retained the highest accuracy on the behavioral measures while those in the metronome intervention had the lowest accuracy. The calibration intervention improved the production task significantly better than the metronome intervention.

When means were separated and compared across all interventions and sessions, results suggested that those in the calibration intervention were significantly more accurate on the production task than those in the metronome intervention. Previous work suggests that the metronome intervention should have successfully improved behavioral timing, which is supported in the present study (Kuznetsov & Wallot, 2011; Okano, Shinya, & Kudo, 2017). While those in the metronome intervention improved across sessions, they did not improve as much as those in the calibration intervention. The present work is the first known study to compare interventions (e.g. calibration vs. metronome) aimed at improving time accuracy, thus this divergence from findings in previous work is feasible. This significant difference between interventions may be explained by the fact that those in the metronome task were trained on a second by finger-tapping along with a metronome, then instructed to carry out this function during the succeeding behavioral tasks. Perhaps the additional motor task and tracking time were competing cognitive functions. This concept can be explained by the internal clock theory (Treisman, 1963), which suggests that as other tasks are introduced, such as finger-tapping, attention is diverted away from tracking time and errors can occur, accounting for less accurate

recordings from individuals in the metronome intervention compared to the calibration intervention. This comparison demonstrates that the metronome task may not be the most appropriate intervention method and instead the calibration intervention may provide a more comprehensive method for improving behavioral timing.

Those in the calibration intervention retained the highest accuracy on the production task compared to all other interventions and performed significantly better on this task than those in the metronome intervention. In the calibration intervention, participants receive feedback regarding their timing abilities after each response. The instantaneous feedback allows participants to calibrate their beliefs about their own timing ability and perform an online adjustment to improve their accuracy. Successful implementation of this intervention requires the participant to introspect and adjust based on the feedback received. The calibration intervention does not demand additional cognitive resources while completing the estimation and production measures, unlike the metronome task (Treisman, 1963). Therefore, according to the internal clock theory (Treisman, 1963), all attentional resources are allocated to the behavioral tasks and any improvements in timing can be attributed to the previous calibration training. However, published work has only validated that feedback improves accuracy on the estimation task (Wearden & Farrar, 2007). Since the estimation and production tasks are considered measurements of the same underlying temporal mechanisms, it is expected that feedback would also impact the production task, but this effect has not been confirmed (Block, 1990). Thus, these findings complement the current body of literature by supporting the idea that feedback can also improve accuracy on the production task.

Previous work indicates that both behavioral tasks measure the same temporal mechanisms; even so other models of timing conversely suggest that production may be a more

demanding task that requires more attentional and working memory resources from the participant compared to estimation (Block, 1990). This latter explanation of timing may explain why participants showed significant differences in the production task across interventions and time, but not on the estimation task. The estimation task is considered passive, wherein participants may incorporate personal strategic methods to track elapsed time (e.g., counting seconds in their heads). On the contrary, the production task is a more active task. Here, the participant is given an interval, retrieves his or her memory of that interval, clicks the start button, activates personal methods for tracking durations, attends, and clicks the stop button to end the trial. Based on prior models of timing, there are theoretically more opportunities for errors during the production task due to the slightly higher cognitive demand (Treisman, 1963). This difference in cognitive demand has been deduced from studies that examined the differences in neurophysiological responses when participants were asked to attend to stimuli. Individuals who were active during the task, or attended to the neurophysiological stimuli, demonstrated increased P2 amplitudes, an ERP believed to relate to attention. This increase in the P2 amplitude supports the idea that more attention is required in these active tasks (i.e., the production task) compared to less active tasks (i.e., the estimation task). (Campbell & Davalos, 2015; Gebuis & Reynvoet, 2013). Therefore, those who are more susceptible to attentional problems may find it slightly more difficult to accurately respond on the production task compared to the estimation task, even if timing ability is consistent, which accounts for the variability observed between the two tasks.

Applied Temporal Processing

Each of the tasks used to measure applied temporal processing were analyzed as separate measures, as done in the behavioral RM-ANOVA. Participants who are in the calibration

intervention were expected to show the greatest change in the applied temporal processing tasks. However, the hypothesis was not supported by the results of the present study.

LNS. Those in the metronome intervention had the highest estimated marginal means on the LNS task whereas those in the calibration intervention had the lowest. All interventions except the calibration intervention showed improvements in these scores across sessions, but there was generally low variation across participants in this task. The sample was drawn from a university, which is an environment that requires high levels of intelligence for success. Since the LNS task is a subtest of the WAIS-IQ test, it is expected that similar, high scores would be observed in this sample. The largest difference between timepoints was a 1-point mean increase seen from those in the metronome intervention. The reduced variation observed in the LNS task may be due to the sample used in this study. The low observed change across sessions supports previous work stating that the LNS task does not fall suit to practice effects but does not support the present study's hypothesis (Beglinger et al., 2005).

Time management surveys. There are limitations to the sample used in the study; however, using a student-based sample may also offer an added benefit. Due to constant feedback on their cognitive abilities, students are potentially more calibrated than typical participants sourced from the community. Thus, the minuscule, nonsignificant changes in survey scores observed across time and intervention may be meaningful in this respect. To examine these small trends, estimated marginal means were compared and revealed that the calibration intervention maintained the highest means on the TMQ, but lowest on the TSQ. Notably, these individuals demonstrated the largest increase in scores on TSQ across time, but still retained the lowest average compared to all other groups.

This low average on the TSQ amounted to the only significant difference: those in the metronome intervention self-reported significantly better TSQ scores compared to the calibration intervention across time, but there were no other significant differences between interventions. Since participants in the calibration intervention were the only ones to systematically receive feedback on their timing abilities, scores were expected to improve on these surveys across time. The calibration intervention demonstrated an observable improvement across time but did not retain mean survey scores as high as those in the metronome intervention. This finding may be explained by the overconfidence effect: Without proper exposure to one's own abilities, individuals tend to be overconfident in their assessments of their skills. This effect is specifically robust when participants self-assess their aptitude on time-related tasks (e.g., the planning fallacy, meeting deadlines) (Sanna et al., 2005).

Since participants in the calibration intervention were the only ones to systematically receive feedback on their timing abilities, scores were expected to improve on these surveys across time. However, improved scores may not be necessarily related to accuracy. Frances-Smyth and Robertson (1999) suggested a similar trend in stating that those who think about time often may be more critical in their self-assessments. For example, those who are hypervigilant about timing may be more likely to inaccurately report that they are late whereas an individual who does not reflect on their timing abilities may seldom report lateness. This alternative explanation for the relationship between survey scores and accurate assessments of time management may explain the overall low survey scores reported by those in the calibration intervention. The calibration intervention required participants to be introspective about their timing abilities, resulting in the most conservative assessments of time managing abilities, indicated by lower survey scores, compared to all other groups. However, this explanation of

survey scores and overconfidence is preliminary since no qualitative data was specifically collected on this phenomenon. The inconsistencies among the data and gaps in literature both allude to the complex relationship between awareness of our own abilities and objective time perception, which has fueled the current, chaotic understanding of timing.

Summary

The field of timing has long sought to explore the ways in which time accuracy can be improved. The interest in improved accuracy is born from studies examining timing in patients with psychopathology, where deficits in neurophysiological and behavioral timing have been related to disorders such as schizophrenia and psychosis (Pantlin & Davalos, 2016; Umbrict & Krjles, 2005). However, differences in performance on timing also exist in healthy populations where gradations in accuracy on behavioral and applied temporal processing tasks may relate to intelligence and/or general cognitive functioning (Franssen, Vandierendonck, & Hiel, 2006; Kramer, Bressan, & Grassi, 2011). Thus, the need to develop tactics to improve timing is important for both clinical and healthy populations. The present work was the first study to validate ways to improve timing by comparing existing interventions.

GENERAL DISCUSSION

Accurate timing affects many important functions for day to day living. Neurotypical or healthy individuals show differences in their timing perception and abilities, which can be evaluated by performance on various timing tasks (i.e. verbal estimation, time management surveys). Previous work has found that shortfalls in these tasks may relate to performance on intelligence tests and cognitive tasks (Franssen, Vandierendonck, & Hiel, 2006; Kramer, Bressan, & Grassi, 2011). Not only are there differences in abilities amongst neurotypical populations, but deficits in time perception may also relate to existing or underlying psychopathology (e.g. schizophrenia, psychosis; Pantlin & Davalos, 2016; Umbricht & Krjles, 2005). Thus, unlocking how time is handled, processed, and applied in healthy individuals may provide answers to important questions about intelligence, psychopathology, and general cognitive functioning.

The present work aimed to unravel and outline the time continuum through two hypotheses. The first hypothesis proposed three levels of timing: neurophysiology (Level I), behavioral timing (Level II), and applied temporal processing (Level III) and used classic timing tasks to represent and evaluate each level. The second hypothesis sought to validate existing interventions aimed to improve timing by comparing task accuracy before and after the implementation of these interventions. Each intervention was selected to specifically target a level of timing. To target Level 1 (Neurophysiology), participants trained on the MMN paradigm during Intervention I (Neurophysiology). The metronome intervention was designated to improve abilities at Level II (Behavioral Timing) and had participants train to finger tap at a rate of one tap per second. Lastly, Level III (Applied Temporal Processing) was targeted by the

calibration intervention which used feedback to calibrate individuals on their timing abilities. Each intervention was specifically chosen to affect one proposed level from Hypothesis I; however, performance on all timing tasks was compared across interventions to determine the most appropriate way to improve timing.

Although the connections between the proposed levels of Hypothesis I were nonsignificant, the results displayed trends in the anticipated direction which have important considerations for the field of timing. Findings of the first hypothesis suggest that the relationship between a millisecond response (Level I: Neurophysiology), the ability to track seconds in a minute (Level II: Behavioral Timing), and plan a day (Level III: Applied Temporal Processing) was not strictly linear. Each level evaluated a different timescale (i.e., milliseconds, seconds, hours, etc.) and perhaps the nonsignificant relationships were due to the vast change in timescales between levels. However, the results displayed trends in the anticipated direction, thus setting the stage to explore intermediate timescales. Tasks such as interval identification and discrimination can be used to address this gap by offering a stepping stone between the timescales measured in the first two levels. These tasks often require comparisons or judgements of intervals which allows researchers to observe timing between milliseconds (Level I: Neurophysiology) and a second (Level II: Behavioral timing).

Systematically measuring decisions that participants make about time in a short interval, such as with tasks that evaluate intermediate stages of timing (>500 ms and < 1 second), may provide additional and necessary information to connect neurophysiological (Level I) to behavioral timing (Level II). It is not feasible to have participants estimate or produce a time interval under a second because the intervals become increasingly difficult to distinguish. For example, asking a participant to estimate or produce 750 ms versus 1 second may not yield

results that truly capture the accuracy of their internal clock (Mioni et al., 2014). To overcome this issue, tasks that employ comparisons or judgements of intervals (e.g., identification and discrimination), allow researchers to evaluate timing in the milliseconds range. These tasks evaluate intermediate stages by requiring the participant to compare, instead of estimate or produce, an interval of 750 ms to 1 second and determine which is the longer interval. Thus, these tasks produce data on the participant's internal clock, and therefore provide an overlap between the firsts two levels.

Accurately comparing intervals, as done in the identification and/or discrimination tasks, directly involves a memory and decision stage that is akin to the one described in the Gibbon (1984) model of timing (see Figure 2). When two intervals are played for comparison, the participant tracks the length of both intervals (clock stage; Gibbon, 1984), retains both of these in working memory (memory stage; Gibbon, 1984), then makes a decision (comparator, decision stage; Gibbon, 1984) about which interval was longer. Thus, these tasks still provide a direct measurement of an individual's internal clock, but also provide information about cognitive functions that indirectly affect timing (i.e., attention and working memory). These cognitive functions become increasingly more impactful on timing as the interval increases, such as in verbal estimation and interval production (Level II: Behavioral timing). With longer intervals, more attentional resources are required to track elapsed time during the clock stage and this information must be held in working memory for a longer period. Therefore, future studies should aim to relate neurophysiological responses to behavioral timing through use of these tasks.

Comparable to the first two levels, the relationship between Levels II (Behavioral Timing) and III (Applied Temporal Processing) was not significant; however, these levels can be

bridged through the development of a time management battery. This battery should include tasks that evaluate how an individual determines the time needed to complete a task, creates a plan for finishing said task, then executes that plan. The ability to perform the first step in time management, determining how long a task will take to complete, is underscored by the ability to track time in the minutes range (e.g., measured by the behavioral tasks). If a battery was to be developed that broke down and evaluated each of the steps of time management and was implemented into a study that also recorded participants' performance on behavioral tasks, then a more concrete understanding of how participants shift between Level II (Behavioral Timing) and III (Applied Temporal Processing) would be captured.

The need for a time management battery is not only essential to connect Levels II (Behavioral Timing) and III (Applied Temporal Processing) in Hypothesis I, but also revealed necessary for developing the most applicable intervention method in Hypothesis II. There are large inconsistencies between perceived timing ability and execution of behaviors that was not addressed in the present study. Although untangling this relationship was beyond the scope of the study, elaborating on the relationship between perception and actual timing ability would have provided a clearer understanding of the findings in Hypothesis II. Those in the metronome intervention had low behavioral accuracy (Level II) but outperformed all other participants on the applied measures. Meanwhile, those in the calibration intervention, who were expected to perform the best on the applied measures, displayed the opposite relationship. These individuals were accurate during the production task (Level II: Behavioral Timing) and retained the lowest mean score on the TSQ, indicating that they did not perceive themselves to be efficient users of time. The development of a time management battery should still include the established surveys (TMQ, TSQ) as these surveys provide foundational information about the participant's perceived

abilities. The TMQ and TSQ combined with a time management task would provide the required information to tease apart the discrepancy between perceived and actual ability at the applied level. Overall, the development of a time management battery would provide a clearer understanding of the inconsistencies found in Hypothesis II and timing in general.

Future Research

Future studies should aim to establish a battery that measures applied temporal processing (i.e., a time management battery) in totality. The measures used in the present study offer a suitable basis, but to fully grasp applied temporal processing, other measures should be included in this battery. As mentioned, the battery could provide a link between Levels II (Behavioral Timing) and III (Applied Temporal Processing) in Hypothesis I by including tasks that would evaluate overlapping timescales (e.g., minutes to hours range). Including a combination of surveys and time management tasks would provide a clearer explanation of the inconsistencies between perceived and tangible timing abilities found in Hypothesis II. The applied temporal processing measures in the present study did not statistically correspond to create a construct when tested as a latent variable, suggesting that a re-evaluation of the combination of measures used is needed. The present findings support use of the popular TMQ and TSQ, and previous work has shown that these surveys possess reliable and valid estimates of general time management abilities (Alay & Koçak, 2002; Bond & Feather, 1988; Clasessens et al., 2007). Since these two were the only measures used, the connection between perceived abilities and execution of proper time management was not fully addressed. The battery should therefore include a time management task and a measure of executive function to bridge the gap between perceived and actual timing abilities.

Time management tasks have been implemented in longitudinal studies attempting to better understand the discrepancy between perceived ability and execution of action. For example, Sanna et al., (2005) had participants make assessments regarding how long their senior thesis would take to complete, then recorded the actual number of days their projects took to finish. Findings demonstrated that participants were overconfident in their predictions. On average participants rated that their projects would take 33.9 days to finish and the actual average time to completion was 55.5 days. A task such as this would offer strong considerations for how participants understand and apply their timing abilities. However, cognitive abilities may also explain the discrepancy found in Sanna et al. (2005). Previous literature states that executive functioning heavily contributes to timing in an applied setting (Level III) and therefore impacts daily life activities such as the ability to avoid procrastination, meet deadlines, and arrive on time. The considerable influence of cognitive function on time provides support for a battery that includes executive functioning tasks that analyze working memory and attention (Franssen, Vandierendonck, & Hiel, 2006; Gibbon et al., 1984; Kramer, Bressan, & Grassi, 2011). While the present study attempted to capture these facets of executive functioning with the LNS task, little variation was found among participants. Future studies should include additional working memory and attentional tasks to the proposed battery and consider a longitudinal time management task and a measure that addresses how often a person thinks about timing.

To understand applications of timing, a developed battery should aim to address how and if participants are introspective about their time management abilities. Although participants in the present study provided self-assessments on use of time and time management behaviors, they were not directly evaluated on these behaviors or thoughts about time. The relationship between thinking about time and executing time-related behaviors is important to consider (Francis-

Smyth & Robertson, 1999). Francis-Smyth and Robertson (1999) suggested that perhaps individuals who are hypervigilant about their abilities in applied timing settings would inaccurately report that they are late simply because they are overly critical in their self-evaluations compared to those who do not think about time. Although speculative, metacognitive beliefs about time could have been a confounding variable in the present study. Perhaps those in the calibration intervention became hyperaware of their own abilities due to the nature of the training. Participants in this intervention may have retained lower survey scores because they were more aware of the limitations of their abilities. Meanwhile, participants in other groups, such as the metronome intervention, did not receive feedback to calibrate their timing abilities nor were they required to engage in the same kind of introspection. Thus, an evaluation of metacognitive beliefs, specifically about time, could shed light on the discrepancy of responses on these surveys, providing more information on the findings in Hypothesis II.

While developing a time management battery was outside the scope of this study, this work aimed to lay down the fundamental relationship between tasks that are traditionally used to evaluate timing. Moving forward, future studies can dig further into the time continuum with multiple levels of temporal perception (i.e., psychophysical measures) as intermediate stages between the proposed levels of timing, and develop a battery of time management. Now that the relationship between traditionally used timing tasks has been explored, more comprehensive investigations aimed to evaluate timing can be explored to lead to a better understanding of the components of timing that have been assessed using traditional timing measures.

Overall, the present study laid the initial outline of the time continuum. The time continuum refers to a mediation process between neurophysiological timing, behavioral timing, and applied measures of timing. This hypothesized mediation across proposed levels is

restricted: as the causal process evolves down the time continuum (see Figure 11), the statistical effects are inherently diminished. Thus, two adjacent measures such as the neurophysiological and behavioral tasks would demonstrate larger effects compared to two distant measures (neurophysiology and applied). Thus, tasks, such as psychophysical tasks, evaluating a timescale that overlaps with the scale evaluated in neurophysiological measures of timing would offer an opportunity for larger effect sizes to be obtained. By closing the gaps in the timescales between tasks, more sensitive measures of temporal processing can be established. Furthermore, the psychophysical tasks offer a widow into the decision making processes that may underlie timing through use of signal detection theory and the time management battery would include direct and indirect measures to capture a more accurate assessment of one's timing. One example could be development and use of a phone application that tracks task duration. This application could include strategies to calibrate individuals as by inputting task categories and descriptions as well as user predictions of completion time.

Inclusion of psychophysical tasks and a time management battery would allow researchers to examine subtle differences between populations. For example, a participant who is identified as low-risk or in a prodromal stage of psychosis may benefit from the more precise measures of timing, as described above. By increasing the sensitivity of timing tasks, a finely-tuned explanation for the variations amongst healthy populations and better understanding of the origin of temporal information can both be obtained.

Limitations

Overall, the tasks used in this exploratory study demonstrated little variability across participants. In the neurophysiological measures, participants did not display much within-participant or between-group variations, such that there were limited differences across sessions

and across participants. Duration-MMN tasks have shown good test retest reliability (r = .76, Tervaniemi & Näätänen, 1999) when a deviant stimulus is 66% of the standard. This study and other work have also shown that dMMN tends to increase across time (Dalebout & Fox, 2001; Tervaniemi & Näätänen, 1999). The deviant stimuli used in the current work was 70% of the standard, which may be partly responsible for the miniscule attenuation in response across time.

The reliability of dMMN offers an explanation for minimal within-participant variation; however, neurophysiological responses between groups were also similar. Those who were selected for Intervention I (Neurophysiology) were directly trained to attend to the MMN paradigm to address Hypothesis II. These individuals were expected to display the largest increases in neurophysiological responses due to the intervention but did not display significant amplitude changes compared to the other groups. The similarity of responses across participants provides support for use of MMN as a biomarker in clinical populations. To use MMN in experiments examining psychopathology, other cognitive functions that are known to be deficient in these populations, such as attention, cannot modulate the response. Healthy individuals who were asked to pay attention to the paradigm did not demonstrate significantly better neurophysiological responses; therefore, the present study findings support the claim that attention does not moderate the MMN response and furthermore supports that this paradigm can be used in special populations as a biomarker (Garrido et al., 2009; Näätänen et al., 2001; Sussman et al., 2002; Campbell & Davalos, 2015).

The focus of this study was on auditory MMN as a pre-attentive measure of timing and therefore, results are constrained to timing in this sensory domain. Auditory MMN was chosen over other forms of MMN (e.g. visual MMN or oddball tasks) as this measure most likely relates to the other timing tasks used. These other types of MMN were not chosen, such as visual time-

based MMN, which is related to adapting to changes (e.g. location, motion direction, orientation, spatial frequency, contrast/luminance, color, shape, or size) in the visual environment (Kimura, Schröger, & Czigler, 2011). Due to the large number of other measures in this study that evaluate timing in the auditory domain, an auditory time-based MMN was appropriate. The use of this auditory tool corresponds specifically with the metronome intervention and the behavioral tasks, both of which required participants to listen to tones binaurally as was completed during the neurophysiological tasks (MMN). Thus, the findings of this study do not extend to other sensory modes of time (e.g. visual, tactile, etc.), but have important considerations for auditory-based timing.

There was also limited variation in the sample on both the behavioral and applied temporal processing measures. This suggests that college students may have been a restrictive population for the scope of this study. The largest variations were found in the surveys across time and in increases in accuracy on the behavioral tasks, but all of these changes were trivial. While there is a benefit in understanding how performance on these tasks change over time for healthy individuals, future studies should examine time differences in a general or clinical population where differences in timing ability may be more pronounced.

Summary

Once characterized as psychology's 'lost dimension,' the present study aimed to address the nuances of timing (Jones, 1976). Individually, each hypothesis assessed the relationship between the most regularly used timing tasks in the field of timing. Hypothesis I examined timing as one fluid field with subdomains. Each subdomain was created and defined by a task that was most often used within the field to assess timing in the brain, as it relates to behavior, or in applied settings. Findings from the current study suggest that high performance on one timing

task is not necessarily related to other timing tasks, which leaves the question – what does it mean to have good performance in each of these tasks? Future research aimed to address this question would benefit from developing a battery specific to applied timing that assesses an individual's timing-related executive functioning (i.e., attention and working memory), how often the individual thinks about time (i.e., metacognitive beliefs), ability to create and maintain a plan (i.e., time management tasks), and the person's perception of their own timing abilities (i.e., TMQ and TSQ).

The interest in whether timing abilities are malleable and can be improved has caught the attention of researchers for decades, since those who are inaccurate often also exhibit psychopathologies or poor performance in educational and occupational settings (Franssen, Vandierendonck, & Hiel, 2006; Gibbon et al., 1984; Kramer, Bressan, & Grassi, 2011; Umbrict & Krjles, 2005). Thus, separate from Hypothesis I, Hypothesis II proposed interventions to increase one's time accuracy. While none of the interventions were significantly more effective across all timing domains, giving feedback to participants showed the most promising results. Awareness of one's own timing ability seemed to allow participants to make quick calibrations in order to increase their accuracy when estimating and producing intervals.

Endorsing symptomology of psychosis, as measured by the PQ, demonstrated strong relationships with poor performance across many measures. Since inaccurate timing is related to symptoms of psychopathology and less effective time management strategies, future studies may offer foundations to better understand the relationship between inaccurate timing, psychopathology, and intelligence. Findings in both hypotheses collectively offer a foundational starting point to begin developing and partitioning out the time continuum.

Table 1a. *Means and standard deviations for all variables*

Variable		N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance
	Session	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Age	N/A	144	8	18	26	19.08	0.13	1.56	2.45
Average Neurophysiology	1	145	5.41	-5.47	-0.05	-1.49	0.09	1.09	1.18
Average Neurophysiology	2	145	5.1	-5.21	-0.11	-1.3	0.07	0.9	0.81
Brain Injury	N/A	144	1	1	2	1.14	0.03	0.35	0.12
Cz	1	101	5.89	-5.9	0	-1.6	0.15	1.5	2.25
Cz	2	99	4.97	-4.98	-0.01	-1.17	0.1	0.99	0.97
Estimation (z-score)	2	141	1.45	-0.62	0.83	-0.19	0.03	0.33	0.11
Estimation (z-score)	1	140	2.48	-0.62	1.87	-0.03	0.04	0.44	0.19
Fz	1	131	5.55	-5.58	-0.03	-1.66	0.11	1.24	1.53
Fz	2	132	5.1	-5.21	-0.11	-1.6	0.09	1.05	1.11
Handed	N/A	144	1	1	2	1.92	0.02	0.27	0.07
LNS (z-score)	1	145	13	0.6	13.6	6.3	0.2	2.43	5.92
LNS (z-score)	2	145	13	1.31	14.31	6.95	0.21	2.48	6.16
Musical Experience (Years)	N/A	144	4	1	5	2.28	0.12	1.43	2.05
Nicotine	N/A	144	0	1	1	1	0	0	0
PQscore	N/A	144	12	0	12	2.55	0.23	2.75	7.55
Production (*z*-score)	1	141	8.87	-1.86	7.01	0.6	0.16	1.85	3.41
Production (*z*-score)	2	139	5.79	-1.9	3.89	-0.13	0.09	1.1	1.22
Pz	1	125	5.93	-5.95	-0.02	-1.52	0.09	1.04	1.08
Pz	2	118	4.9	-4.92	-0.02	-1.12	0.09	0.96	0.92
Race	N/A	143	7	1	8	6.63	0.19	2.22	4.91
Sex	N/A	144	1	1	2	1.72	0.04	0.45	0.21
TMQ (z-score)	1	134	45	27.77	72.77	51	0.74	8.55	73.13
TMQ (z-score)	2	131	43	27.98	70.98	50.9	0.73	8.36	69.86
TSQ (z-score)	1	136	59	65.08	124.08	91.93	0.98	11.38	129.42
TSQ (z-score)	2	127	51	64.97	115.97	92.3	0.93	10.5	110.25
Unconscious	N/A	144	0	1	1	1	0	0	0

Note. The descriptive statistics for all variables, separated by session is presented, in Table 1. All electrodes are in the metric of microvolts (μV). PQ refers to the prodromal questionnaire score where higher scores reflect endorsement of symptoms associated with psychosis.

Table 1b. *Skew and kurtosis values for all variables*

Variable		Skewness		Kurtosis	
	Session	Statistic	Std. Error	Statistic	Std. Error
Age	N/A	2.23	0.2	5.45	0.4
Average Neurophysiology	1	-1.33	0.2	1.8	0.4
Average Neurophysiology	2	-1.82	0.2	4.28	0.4
Brain Injury	N/A	2.11	0.2	2.49	0.4
Cz	1	-1.15	0.24	0.64	0.48
Cz	2	-1.67	0.24	3.54	0.48
Estimation (z-score)	2	1.07	0.2	0.62	0.41
Estimation (z-score)	1	1.66	0.2	3.67	0.41
Fz	1	-0.91	0.21	0.37	0.42
Fz	2	-1.03	0.21	1.35	0.42
Handed	N/A	-3.22	0.2	8.51	0.4
LNS (z-score)	1	0.2	0.2	0.14	0.4
LNS (z-score)	2	0.3	0.2	-0.08	0.4
Musical Experience (Years)	N/A	0.79	0.2	-0.63	0.4
Nicotine	N/A	N/A	N/A	N/A	N/A
PQscore	N/A	1.46	0.2	1.81	0.4
Production (*z*-score)	1	1.67	0.2	2.24	0.41
Production (*z*-score)	2	1.17	0.21	1.84	0.41
Pz	1	-1.13	0.22	2.04	0.43
Pz	2	-1.65	0.22	3.11	0.44
Race	N/A	-1.17	0.2	-0.31	0.4
Sex	N/A	-0.96	0.2	-1.09	0.4
TMQ (z-score)	1	0.12	0.21	-0.36	0.42
TMQ (z-score)	2	-0.02	0.21	-0.41	0.42
TSQ (z-score)	1	0.24	0.21	-0.11	0.41
TSQ (z-score)	2	-0.04	0.21	-0.59	0.43
Unconscious	N/A				

Note. Normality was assessed using kurtosis and skew values of all scaled data. Kurtosis and skew values that fall between -2 and +2 are considered acceptable (George & Mallory, 2010; Gravetter & Wallnau, 2014). PQ refers to the prodromal questionnaire score where higher scores reflect endorsement of symptoms associated with psychosis.

Table 2. *Model fit and latent variable factor loadings of proposed model.*

	Chi-squared									
Factors	*DF*	$\chi 2$	*p*	CFI	TLI	SRMR	Estimate	S.E	Ratio	*p*
LNS	16	35.26	0.00	0.08	-0.61	0.09	1	0	0.99	0.99
TM1Z							0.02	0.04	0.62	0.54
TSQ1Z							-0.07	0.05	-1.62	0.11

Note. The original factor loadings for the proposed model are presented in Table 2. This table demonstrates that these loadings did meet criteria for latent variable modeling and single indicator models were adopted to address the hypothesis instead.

Table 3. *Model fit for single indicator models*

	Model Description	Chi-Squared *DF*	Chi-Squared $\chi 2$	Chi-Squared *p*	CFI	TLI	SRMR	Conclusion
Criteria	Standard	NA	NA	>.05	Close to 1	Close to 1	<.08	Excellent
LNS	Two mediator model	6	39.41	0	0.01	-1.47	0.12	Poor
	Estimation	3	2.77	0.43	1	1	0.04	Excellent
	Production	3	2.96	0.40	1	1	0.04	Excellent
TMQ	Two mediator model	6	38.16	0	0	-1.725	0.12	Poor
	Estimation	3	1.03	0.80	1	1	0.03	Excellent
	Production	3	0.84	0.84	1	1	0.02	Excellent
TSQ	Two mediator model	6	37.13	0	0	-1.60	0.12	Poor
	Estimation	3	0.62	0.89	1	1	0.02	Excellent
	Production	3	0.39	0.94	1	1	0.03	Excellent

Note. A reference is presented in the first row of the table with acceptable values for each of the model fit criteria (Hu & Butler, 1999).

Table 4. *Standardized and unstandardized effect sizes for the letter-number sequencing task.*

Mediator	Path Description	Path	Estimate	S.E	Ratio	*p*	Standardized Estimates	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Estimation	APT by LNS		1.00	0.00	999.00	999.00	0.60	1.00	1.00
	APT on Estimation	b	-0.10	0.59	-0.17	0.87	-0.03	-1.40	1.00
	APT on Neurophysiology	c	-0.09	0.23	-0.40	0.69	-0.07	-0.57	0.36
	PQG on APT	covar	-0.05	0.08	-0.59	0.56	-0.30	-0.19	0.00
	Estimation on Neurophysiology	a	0.01	0.04	0.26	0.80	0.03	-0.07	0.10
	PQG on Neurophysiology	covar	0.00	0.03	0.05	0.96	0.01	-0.05	0.06
	Music on Neurophysiology	covar	0.13	0.14	0.90	0.37	0.10	-0.16	0.41
	Indirect effect	ab	0.00	0.03	-0.04	0.97	0.00	-0.07	0.04
Production	APT by LNS		1.00	0.00	999.00	999.00	0.60	1.00	1.00
	APT on Production	b	-0.17	0.12	-1.33	0.18	-0.23	-0.42	0.07
	APT on Neurophysiology	c	-0.15	0.23	-0.66	0.51	-0.11	-0.60	0.31
	PQG on APT	covar	-0.04	0.07	-0.66	0.51	-0.27	-0.20	0.00
	Production on Neurophysiology	a	-0.35	0.21	-1.62	0.11	-0.19	-0.75	0.10
	PQG on Neurophysiology	covar	0.00	0.03	0.07	0.94	0.01	-0.05	0.06
	Music on Neurophysiology	covar	1.00	0.00	999.00	999.00	0.60	1.00	1.00
	Indirect effect	ab	-0.10	0.59	-0.17	0.87	-0.03	-1.40	1.00

Note. This is a single indicator model for LNS where the variable APT represents the reliability measure for LNS. All variables are session 1 variables. PQG stands for prodromal questionnaire group. Those with a score above 6 received a 2 to indicate *prodromal.* Those with a score less than 1 received a 1 to indicate *control*. Covar stands for covariate.

Table 5. *Standardized and unstandardized effect sizes for TSQ.*

Mediator	Path Description	Path	Estimate	S.E	Ratio	*p*	Standardized Estimates	95% Confidence Interval	
								Lower Bound	Upper Bound
Estimation	APT by TSQ		1.00	0.00	999.00	999.00	0.99	1.00	1.00
	APT on Estimation	b	0.00	0.29	-0.01	0.99	0.00	-0.66	0.47
	APT on Neurophysiology	c	-0.06	0.10	-0.64	0.53	-0.06	-0.24	0.14
	PQ Score on APT	covar	0.04	0.02	1.61	0.11	0.19	0.00	0.10
	Estimation on Neurophysiology	a	0.01	0.04	0.26	0.80	0.03	-0.07	0.10
	PQ Score ON Neurophysiology	covar	0.01	0.02	0.35	0.72	0.04	-0.04	0.05
	Music on Neurophysiology	covar	0.13	0.14	0.90	0.37	0.10	-0.16	0.41
	Indirect effect	ab	0.00	0.01	0.00	1.00	0.00	-0.04	0.03
Production	APT by TSQ		1.00	0.00	999.00	999.00	0.99	1.00	1.00
	APT on Production	b	0.01	0.07	0.20	0.85	0.03	-0.15	0.14
	APT on Avg Neurophysiology	c	-0.06	0.10	-0.54	0.59	-0.06	-0.26	0.14
	PQ Score on APT	covar	0.04	0.02	1.61	0.11	0.19	0.00	0.10
	Production on Neurophysiology	a	-0.36	0.21	-1.66	0.10	-0.20	-0.76	0.09
	PQ Score on Neurophysiology	covar	0.01	0.02	0.35	0.73	0.04	-0.04	0.05
	Music on Neurophysiology	covar	1.00	0.00	999.00	999.00	0.99	1.00	1.00
	Indirect effect	ab	0.00	0.29	-0.01	0.99	0.00	-0.66	0.47

Note. This is a single indicator model for TSQ where the variable APT represents the reliability measure for TSQ. All variables are session 1 variables. PQ Score refers to the participant's score on the prodromal questionnaire group. Those with a score above 6 received a 2 to indicate *prodromal.* Those with a score less than 1 received a 1 to indicate *control*. Covar stands for covariate.

Table 6a. *Standardized and unstandardized effect sizes for TMQ with assumed reliability scores*

Mediator	Path Description	Path	Estimate	S.E	Ratio	*p*	Standardized Estimates	95% Confidence Interval	
								Lower Bound	Upper Bound
Estimation	APT by TMQ		1.00	0.00	999.00	999.00	0.65	1.00	1.00
	APT on Estimation	b	-0.15	0.20	-0.77	0.44	-0.13	-0.60	0.18
	APT on Neurophysiology	c	-0.02	0.07	-0.30	0.76	-0.05	-0.16	0.11
	PQ Score on APT	covar	0.04	0.07	0.54	0.59	0.08	-0.05	0.22
	Estimation on Neurophysiology	a	0.01	0.04	0.23	0.82	0.03	-0.07	0.10
	PQ Score on Neurophysiology	covar	0.01	0.03	0.27	0.79	0.03	-0.05	0.05
	Music on Neurophysiology	covar	0.13	0.14	0.90	0.37	0.10	-0.16	0.41
	Indirect effect	ab	0.00	0.01	-0.15	0.88	0.00	-0.03	0.01
Production	APT by TMQ		1.00	0.00	999.00	999.00	0.65	1.00	1.00
	APT on Production	b	-0.06	0.05	-1.23	0.22	-0.22	-0.14	0.04
	APT on Neurophysiology	c	-0.04	0.08	-0.55	0.58	-0.09	-0.20	0.09
	PQ Score on APT	covar	0.04	0.06	0.68	0.50	0.09	-0.04	0.20
	Production on Neurophysiology	a	-0.37	0.22	-1.71	0.09	-0.20	-0.78	0.07
	PQ Score on Neurophysiology	covar	0.01	0.03	0.28	0.78	0.03	-0.05	0.05
	Music on Neurophysiology	covar	1.00	0.00	999.00	999.00	0.65	1.00	1.00
	Indirect effect	ab	-0.15	0.20	-0.77	0.44	-0.13	-0.60	0.18

Note. This is a single indicator model for TMQ where the variable APT represents the reliability measure for TMQ. The values used as the reliability measure for TMQ were not available in literature; therefore, Table 6a displaces the values for TMQ when assuming .1 as the reliability score for the single indicator model. This was assumed because there was no literature for test-retest reliability of the TMQ, only a Cronbach's alpha value for internal consistency. There was no support for using Cronbach's alpha to account for measurement error in this type of model. Thus, the observed score for TMQ was used in the final model is provided in the next table for comparison, but the differences among the results are miniscule. All variables are session 1 variables. PQ score refers to the participant's score on the prodromal questionnaire. Those with a score above 6 received a 2 to indicate *prodromal.* Those with a score less than 1 received a 1 to indicate *control*. Covar stands for covariate.

Table 6b. *Standardized and unstandardized effect sizes for the observed score of TMQ*

Mediator	Path Description	Path	Estimate	S.E	Ratio	*p*	Standardized Estimates	95% Confidence Interval	
								Lower Bound	Upper Bound
Estimation	TMQ on Estimation	b	-0.15	0.20	-0.77	0.44	-0.09	-0.62	0.18
	TMQ on Neurophysiology	c	-0.02	0.07	-0.30	0.76	-0.03	-0.16	0.11
	PQ Score on TMQ	covar	0.02	0.02	0.82	0.42	0.05	-0.05	0.05
	Estimation on Neurophysiology	a	0.01	0.04	0.24	0.81	0.03	-0.07	0.10
	PQ Score on Neurophysiology	covar	0.01	0.02	0.26	0.79	0.03	-0.01	0.06
	Music on Neurophysiology	covar	0.13	0.14	0.90	0.37	0.10	-0.16	0.41
	Indirect effect	ab	0.00	0.01	-0.15	0.88	0.00	-0.03	0.01
Production	TMQ on Production	b	-0.06	0.05	-1.20	0.23	-0.14	-0.14	0.04
	TMQ on Neurophysiology	c	-0.04	0.08	-0.55	0.59	-0.06	-0.20	0.10
	PQ Score on TMQ	covar	0.02	0.02	0.81	0.42	0.05	-0.01	0.06
	Production on Neurophysiology	a	-0.37	0.22	-1.70	0.09	-0.20	-0.78	0.07
	PQ Score on Neurophysiology	covar	0.01	0.02	0.26	0.79	0.03	-0.05	0.05
	Music on Neurophysiology	covar	0.13	0.14	0.90	0.37	0.10	-0.16	0.41
	Indirect effect	ab	0.02	0.02	0.88	0.38	0.05	-0.01	0.09

Note. Table 6b shows the observed score for TMQ, as opposed to a single indicator model with .1 as an assumed reliability (see Table 6a). The differences between the results are miniscule. APT stands for applied temporal processing latent variable. In this case, the APT latent variable for TMQ with the reliability for TMQ as the single indicator. All variables are session 1 variables. PQ score refers to the participant's score on the prodromal questionnaire. Those with a score above 6 received a 2 to indicate *prodromal.* Those with a score less than 1 received a 1 to indicate *control*. Covar stands for covariate.

Table 7a. *Levene's test of equality of error variances for neurophysiology*

Variable	Session	F	df1	df2	Sig.
Average Response	1	0.07	3	140	0.98
Average Response	2	1.85	3	140	0.14

Note. Significant values at the .05 level indicate that the assumption has not been met.

Table 7b. *Levene's test of equality of error variances for behavioral timing*

Variable	Session	F	df1	df2	Sig.
Estimation	1	0.78	3	128	0.51
Estimation	2	5.20	3	128	0.00*
Production	1	1.13	3	128	0.34
Production	2	1.73	3	128	0.16

Note. Significant values at the .05 level indicate that the assumption has not been met and are denoted with an asterisk. These values correspond to variables that did not retain equality for equal error variances, suggesting opportunity for a Type I error. Therefore, the Greenhouse-Geisser statistic, a stringent assessment of significance, was used when interpreting these models.

Table 7c. *Levene's test of equality of error variances for applied temporal processing*

Variable	Session	F	df1	df2	Sig.
TSQ	1	1.42	3	106	0.24
TSQ	2	0.62	3	106	0.60
TMQ	1	1.10	3	106	0.35
TMQ	2	6.16	3	106	0.00*
LNS	1	0.17	3	106	0.92
LNS	2	0.48	3	106	0.70

Note. Significant values at the .05 level indicate that the assumption has not been met and are denoted with an asterisk. These values correspond to variables that did not retain equality for equal error variances, suggesting opportunity for a Type I error. Therefore, the Greenhouse-Geisser statistic, a stringent assessment of significance, was used when interpreting these models.

Table 8. *Hypothesis II descriptive statistics for average neurophysiological response by intervention*

Session	Intervention	Mean	Std. Deviation	*N*
1	Control	-1.50	1.12	38
	Metronome	-1.56	1.21	39
	Calibration	-1.50	1.05	33
	Neurophysiology	-1.39	0.97	34
	Total	-1.49	1.09	144
2	Control	-1.24	0.71	38
	Metronome	-1.08	0.70	39
	Calibration	-1.42	1.04	33
	Neurophysiology	-1.49	1.12	34
	Total	-1.30	0.90	144

Note. This table displays the descriptive statistics for each intervention separated by time for the neurophysiological tasks only.

Table 9. *Hypothesis II main effects for neurophysiology*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	20.61	1	20.61	17.14	0.00	0.11
Prodromal score	1.10	1	1.10	0.92	0.34	0.01
Musical experience	0.33	1	0.33	0.27	0.60	0.00
Intervention	1.05	3	0.35	0.29	0.83	0.01
Error	165.95	138	1.20			

Note. The significance values for each main effect is displayed in Table 9. None of the main effects for prodromal score, musical experience, or intervention were significantly different across time when the neurophysiological task was evaluated.

Table 10. *Hypothesis II interaction effects and partial eta-squared values for neurophysiology*

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	3.94	1	3.94	4.88	0.03	0.03
Time * Prodromal Score	2.52	1	2.52	3.12	0.08	0.02
Time * Music (Years)	0.24	1	0.24	0.29	0.59	0.00
Time * Intervention	4.08	3	1.36	1.69	0.17	0.04
Error(Time)	111.25	138	0.81			

Note. The main effects for time and the interaction effects of time*intervention, and time by each of the covariates is presented in Table 10. The interaction of time and prodromal score had significantly different neurophysiological responses across time.

Table 11a. *Hypothesis II estimated marginal means for neurophysiology for time only*

			95% Confidence Interval	
Time	Mean	Std. Error	Lower Bound	Upper Bound
1	-1.49	0.09	-1.67	-1.31
2	-1.31	0.08	-1.46	-1.16

Note. The estimated marginal means, for time only, are presented in Table 11a. These are adjusted values that account for variables that may affect neurophysiological responses (covariates = prodromal scores and years of musical experience).

Table 11b. *Hypothesis II estimated marginal means for neurophysiology for intervention only*

Intervention	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Control	-1.39	0.13	-1.64	-1.14
Metronome	-1.31	0.13	-1.55	-1.06
Calibration	-1.44	0.14	-1.71	-1.17
Neurophysiology	-1.46	0.13	-1.73	-1.20

Note. The estimated marginal means, for intervention only, are presented in Table 11b. These are adjusted values that account for variables that may affect neurophysiological responses (covariates = prodromal scores and years of musical experience).

Table 11c. *Hypothesis II estimated marginal means for neurophysiology for time and intervention*

Intervention	Session	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Control	1	-1.50	0.18	-1.86	-1.15
	2	-1.29	0.15	-1.58	-1.00
Metronome	1	-1.56	0.18	-1.91	-1.21
	2	-1.06	0.14	-1.35	-0.78
Calibration	1	-1.50	0.20	-1.89	-1.11
	2	-1.37	0.16	-1.68	-1.05
Neurophysiology	1	-1.40	0.19	-1.77	-1.02
	2	-1.52	0.15	-1.82	-1.21

Note. The estimated marginal means, for time and intervention, are presented in Table 11c. These are adjusted values that account for variables that may affect neurophysiological responses (covariates = prodromal scores and years of musical experience).

Table 12a. *Hypothesis II pairwise comparison for neurophysiology for time only*

(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
1	2	-0.18	0.11	0.09	-0.39	0.03
2	1	0.18	0.11	0.09	-0.03	0.39

Note. The pairwise comparisons are made using the estimated marginal means, for time only. These comparisons are made on adjusted values that account for variables that may affect neurophysiological responses (covariates = prodromal scores and years of musical experience).

Table 12b. *Hypothesis II pairwise comparison for neurophysiology for intervention only*

(I) Intervention	(J) Intervention	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Control	Metronome	-0.09	0.18	0.63	-0.44	0.27
	Calibration	0.05	0.19	0.81	-0.33	0.42
	Neurophysiology	0.07	0.18	0.71	-0.29	0.43
Metronome	Control	0.09	0.18	0.63	-0.27	0.44
	Calibration	0.13	0.18	0.47	-0.23	0.50
	Neurophysiology	0.16	0.18	0.40	-0.21	0.52
Calibration	Control	-0.05	0.19	0.81	-0.42	0.33
	Metronome	-0.13	0.18	0.47	-0.50	0.23
	Neurophysiology	0.02	0.19	0.90	-0.36	0.41
Neurophysiology	Control	-0.07	0.18	0.71	-0.43	0.29
	Metronome	-0.16	0.18	0.40	-0.52	0.21
	Calibration	-0.02	0.19	0.90	-0.41	0.36

Note. The pairwise comparisons are made using the estimated marginal means, for intervention only. These comparisons are made on adjusted values that account for variables that may affect neurophysiological responses (covariates = prodromal scores and years of musical experience).

Table 13. *Hypothesis II descriptive statistics for estimation and production tasks by intervention*

Task by Session	Intervention	Mean	Std. Deviation	N
Production at Time 1	Control	0.21	1.39	35
	Metronome	0.66	1.91	35
	Calibration	0.39	1.23	31
	Neurophysiology	0.20	1.45	31
	Total	0.37	1.52	132
Production at Time 2	Control	0.14	1.31	35
	Metronome	0.02	1.06	35
	Calibration	-0.75	0.85	31
	Neurophysiology	-0.08	0.83	31
	Total	-0.15	1.08	132
Estimation at Time 1	Control	-0.09	0.46	35
	Metronome	-0.03	0.49	35
	Calibration	-0.04	0.39	31
	Neurophysiology	-0.10	0.34	31
	Total	-0.06	0.42	132
Estimation at Time 2	Control	-0.16	0.26	35
	Metronome	-0.15	0.38	35
	Calibration	-0.34	0.20	31
	Neurophysiology	-0.21	0.27	31
	Total	-0.21	0.29	132

Note. This table displays the descriptive statistics for each intervention separated by time for the behavioral tasks only. All values in the production and estimation task are z-scores.

Table 14. *Hypothesis II main and interaction effects with partial eta-squared for behavioral tasks*

Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	Production	18.544	1	18.544	14.162	0	0.1
	Estimation	1.565	1	1.565	17.522	0	0.12
Time * Intervention	Production	10.738	3	3.579	2.734	0.046	0.06
	Estimation	0.532	3	0.177	1.985	0.119	0.044
Error(Time)	Production	167.603	128	1.309			
	Estimation	11.433	128	0.089			

Note. The significance values for the main effect of time and the interaction of time and intervention is displayed in Table 14. The main effect for time was significant for the behavioral tasks and the interaction was significant for the production task only.

Table 15a. *Hypothesis II estimated marginal means for behavioral tasks for time only*

Measure	Session	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Estimation	1	-0.06	0.04	-0.14	0.01
	2	-0.22	0.03	-0.27	-0.17
Production	1	0.36	0.13	0.10	0.63
	2	-0.17	0.09	-0.35	0.01

Note. The estimated marginal means, for time only, are presented in Table 15a. These are adjusted values that account for variables that may affect behavioral responses.

Table 15b. *Hypothesis II estimated marginal means for behavioral tasks for intervention only*

Measure	Intervention	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Production	Control	0.18	0.17	-0.17	0.52
	Metronome	0.34	0.17	-0.01	0.68
	Calibration	-0.18	0.18	-0.55	0.18
	Neurophysiology	0.06	0.18	-0.30	0.43
Estimation	Control	-0.12	0.05	-0.22	-0.02
	Metronome	-0.09	0.05	-0.19	0.01
	Calibration	-0.19	0.05	-0.30	-0.08
	Neurophysiology	-0.15	0.05	-0.26	-0.05

Note. The estimated marginal means, for intervention only, are presented in Table 15b. These are adjusted values that account for variables that may affect behavioral responses.

Table 15c. *Hypothesis II estimated marginal means for behavioral tasks for time and intervention*

Measure	Intervention	Time	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Production	Control	1	0.21	0.26	-0.30	0.72
		2	0.14	0.18	-0.21	0.49
	Metronome	1	0.66	0.26	0.15	1.17
		2	0.02	0.18	-0.33	0.36
	Calibration	1	0.39	0.27	-0.15	0.93
		2	-0.75	0.19	-1.12	-0.38
	Neurophysiology	1	0.20	0.27	-0.34	0.74
		2	-0.08	0.19	-0.44	0.29
Estimation	Control	1	-0.09	0.07	-0.23	0.06
		2	0.03	0.05	-0.25	-0.06
	Metronome	1	-0.03	0.07	-0.17	0.11
		2	-0.16	0.05	-0.25	-0.06
	Calibration	1	-0.04	0.08	-0.19	0.12
		2	-0.34	0.05	-0.45	-0.24
	Neurophysiology	1	-0.10	0.08	-0.25	0.06
		2	-0.21	0.05	-0.31	-0.11

Note. The estimated marginal means, for time and intervention, are presented in Table 15c. These are adjusted values that account for variables that may affect behavioral responses.

Table 16a. *Hypothesis II pairwise comparisons of behavioral tasks for time only*

Measure	(I) Time	(J) Time	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Production	1	2	0.53*	0.14	0	0.25	0.81
	2	1	-0.53*	0.14	0	-0.81	-0.25
Estimation	1	2	0.15*	0.04	0	0.08	0.23
	2	1	-0.15*	0.04	0	-0.23	-0.08

Note. The pairwise comparisons are made using the estimated marginal means, for time only. These comparisons are made on adjusted values that account for variables that may affect behavioral responses.

Table 16b. *Hypothesis II pairwise comparisons of behavioral tasks for intervention only*

Measure	(I) Intervention	(J) Intervention	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Production	Control	Metronome	-0.16	0.25	0.51	-0.65	0.32
		Calibration	0.36	0.25	0.16	-0.14	0.86
		Neurophysiology	0.12	0.25	0.65	-0.39	0.62
	Metronome	Control	0.16	0.25	0.51	-0.32	0.65
		Calibration	.518*	0.25	0.04	0.02	1.02
		Neurophysiology	0.28	0.25	0.28	-0.22	0.78
	Calibration	Control	-0.36	0.25	0.16	-0.86	0.14
		Metronome	-.518*	0.25	0.04	-1.02	-0.02
		Neurophysiology	-0.24	0.26	0.35	-0.76	0.27
	Neurophysiology	Control	-0.12	0.25	0.65	-0.62	0.39
		Metronome	-0.28	0.25	0.28	-0.78	0.22
		Calibration	0.24	0.26	0.35	-0.27	0.76
Estimation	Control	Metronome	-0.03	0.07	0.68	-0.17	0.11
		Calibration	0.07	0.07	0.36	-0.08	0.21
		Neurophysiology	0.03	0.07	0.68	-0.12	0.18
	Metronome	Control	0.03	0.07	0.68	-0.11	0.17
		Calibration	0.10	0.07	0.19	-0.05	0.24
		Neurophysiology	0.06	0.07	0.42	-0.09	0.21
	Calibration	Control	-0.07	0.07	0.36	-0.21	0.08
		Metronome	-0.10	0.07	0.19	-0.24	0.05
		Neurophysiology	-0.04	0.08	0.63	-0.19	0.11
	Neurophysiology	Control	-0.03	0.07	0.68	-0.18	0.12
		Metronome	-0.06	0.07	0.42	-0.21	0.09
		Calibration	0.04	0.08	0.63	-0.11	0.19

Note. The pairwise comparisons are made using the estimated marginal means, for intervention only. These comparisons are made on adjusted values that account for variables that may affect behavioral responses.

Table 17. *Hypothesis II Descriptive statistics for applied temporal processing tasks by intervention*

Task	Intervention	Mean	Std. Deviation	N
LNS at Time 1	Control	6.27	2.41	30
	Metronome	6.42	2.50	27
	Feedback	6.30	2.51	23
	Neurophysiology	6.20	2.43	30
	Total	6.29	2.43	110
LNS at Time 2	Control	7.01	2.41	30
	Metronome	7.35	2.39	27
	Feedback	6.18	2.10	23
	Neurophysiology	6.91	2.47	30
	Total	6.90	2.36	110
TMQ at Time 1	Control	50.93	6.36	30
	Metronome	48.91	7.86	27
	Feedback	52.85	8.14	23
	Neurophysiology	50.23	8.99	30
	Total	50.65	7.89	110
TMQ at Time 2	Control	51.11	5.66	30
	Metronome	49.87	9.71	27
	Feedback	52.41	9.52	23
	Neurophysiology	50.88	7.96	30
	Total	51.02	8.18	110
TSQ at Time 1	Control	88.78	8.87	30
	Metronome	94.08	12.11	27
	Feedback	88.12	11.39	23
	Neurophysiology	95.28	11.00	30
	Total	91.71	11.15	110
TSQ at Time 2	Control	92.54	8.43	30
	Metronome	94.31	10.43	27
	Feedback	90.97	10.63	23
	Neurophysiology	93.51	11.28	30
	Total	92.91	10.14	110

Note. This table displays the descriptive statistics for each intervention separated by time for the applied temporal processing tasks only. All values are z-scores.

Table 18. *Hypothesis II Main effects for applied temporal processing tasks*

Source	Task	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	LNS	318.47	1	318.47	35.16	0.00	0.25
	TMQ	19778.70	1	19778.70	167.46	0.00	0.62
	TSQ	51566.99	1	51566.99	271.99	0.00	0.72
PQ score	LNS	5.82	1	5.82	0.64	0.43	0.01
	TMQ	158.22	1	158.22	1.34	0.25	0.01
	TSQ	949.87	1	949.87	5.01	0.03	0.05
Sex	LNS	5.76	1	5.76	0.64	0.43	0.01
	TMQ	106.74	1	106.74	0.90	0.34	0.01
	TSQ	1.78	1	1.78	0.01	0.92	0.00
Intervention	LNS	10.69	3	3.56	0.39	0.76	0.01
	TMQ	273.69	3	91.23	0.77	0.51	0.02
	TSQ	1151.34	3	383.78	2.02	0.12	0.06
Error	LNS	941.96	104	9.06			
	TMQ	12283.18	104	118.11			
	TSQ	19717.67	104	189.59			

Note. The significance values for each main effect is displayed in Table 18. The only main effect that was significant was for TSQ and prodromal group. None of the other main effects (i.e., sex or intervention) were significantly different across time when the applied temporal processing tasks were evaluated.

Table 19. *Hypothesis II Interaction effects and partial eta-squared for applied temporal processing tasks*

Source	Task	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	LNS	0.70	1	0.70	0.26	0.61	0.00
	TMQ	3.96	1	3.96	0.33	0.57	0.00
	TSQ	72.48	1	72.48	2.73	0.10	0.03
Time * Prodromal Score	LNS	0.01	1	0.01	0.00	0.95	0.00
	TMQ	1.33	1	1.33	0.11	0.74	0.00
	TSQ	53.49	1	53.49	2.01	0.16	0.02
Time * Sex	LNS	0.00	1	0.00	0.00	0.99	0.00
	TMQ	0.85	1	0.85	0.07	0.79	0.00
	TSQ	3.28	1	3.28	0.12	0.73	0.00
Time * Intervention	LNS	7.62	3	2.54	0.95	0.42	0.03
	TMQ	12.22	3	4.07	0.34	0.80	0.01
	TSQ	291.92	3	97.31	3.66	0.02	0.10
Error(Time)	LNS	279.46	104	2.69			
	TMQ	1265.92	104	12.17			
	TSQ	2763.24	104	26.57			

Note. The significance values for the main effect of time and the interaction of time and intervention is displayed in Table 19. The interaction of time and intervention was significant for the TSQ task only. No other main effects reached significance.

Table 20a. *Hypothesis II estimated marginal means for applied temporal processing tasks for time only*

				95% Confidence Interval	
Measure	Session	Mean	Std. Error	Lower Bound	Upper Bound
LNS	1	6.29	0.24	5.82	6.76
	2	6.86	0.23	6.41	7.31
TMQ	1	50.75	0.76	49.25	52.25
	2	51.09	0.79	49.52	52.66
TSQ	1	91.50	1.02	89.49	93.52
	2	92.80	0.98	90.86	94.73

Note. The estimated marginal means, for time only, are presented in Table 20a. These are adjusted values that account for variables that may affect responses on the applied temporal processing (covariates = prodromal scores and biological sex).

Table 20b. *Hypothesis II estimated marginal means for applied temporal processing tasks for intervention only*

Measure	Intervention	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
LNS	Control	6.68	0.39	5.91	7.46
	Metronome	6.84	0.41	6.03	7.66
	Calibration	6.20	0.45	5.31	7.10
	Neurophysiology	6.58	0.39	5.80	7.35
TMQ	Control	50.81	1.41	48.01	53.61
	Metronome	49.58	1.49	46.64	52.53
	Calibration	52.84	1.62	49.62	56.06
	Neurophysiology	50.45	1.41	47.65	53.24
TSQ	Control	91.03	1.79	87.48	94.58
	Metronome	93.98	1.88	90.24	97.71
	Calibration	88.82	2.06	84.74	92.90
	Neurophysiology	94.78	1.79	91.23	98.32

Note. The estimated marginal means, for intervention only, are presented in Table 20b. These are adjusted values that account for variables that may affect responses on the applied temporal processing (covariates = prodromal scores and biological sex).

Table 20c. *Hypothesis II estimated marginal means for applied temporal processing tasks for time and intervention*

Measure	Intervention	Session	Mean	Std. Error	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
LNS	Control	1	6.31	0.45	5.41	7.21
		2	7.06	0.44	6.19	7.92
	Metronome	1	6.37	0.48	5.43	7.32
		2	7.31	0.46	6.40	8.22
	Calibration	1	6.26	0.52	5.22	7.29
		2	6.15	0.50	5.16	7.14
	Neurophysiology	1	6.22	0.45	5.32	7.12
		2	6.93	0.44	6.07	7.80
TMQ	Control	1	50.73	1.45	47.85	53.60
		2	50.89	1.52	47.88	53.89
	Metronome	1	49.10	1.52	46.08	52.12
		2	50.06	1.60	46.89	53.22
	Calibration	1	53.02	1.67	49.72	56.33
		2	52.66	1.74	49.20	56.11
	Neurophysiology	1	50.14	1.45	47.27	53.01
		2	50.75	1.51	47.75	53.75
TSQ	Control	1	89.23	1.95	85.36	93.09
		2	92.83	1.87	89.12	96.54
	Metronome	1	93.83	2.05	89.76	97.89
		2	94.13	1.97	90.23	98.03
	Calibration	1	87.21	2.24	82.76	91.65
		2	90.43	2.15	86.16	94.69
	Neurophysiology	1	95.76	1.95	91.90	99.62
		2	93.79	1.87	90.09	97.50

Note. The estimated marginal means, for intervention and time, are presented in Table 20c. These are adjusted values that account for variables that may affect responses on the applied temporal processing (covariates = prodromal scores and biological sex).

Table 21a. *Hypothesis II pairwise comparisons for applied temporal processing tasks for time only*

Measure	Session (I)	Session (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference Lower Bound	95% Confidence Interval for Difference Upper Bound
LNS	1	2	-.571*	0.22	0.01	-1.01	-0.13
	2	1	.571*	0.22	0.01	0.13	1.01
TMQ	1	2	-0.34	0.47	0.48	-1.28	0.60
	2	1	0.34	0.47	0.48	-0.60	1.28
TSQ	1	2	-1.29	0.70	0.07	-2.68	0.10
	2	1	1.29	0.70	0.07	-0.10	2.68

Note. The pairwise comparisons are made using the estimated marginal means, for time only. These comparisons are made on adjusted values that account for variables that may affect responses on the applied temporal processing tasks (covariates = prodromal score and biological sex).

Table 21b. *Hypothesis II pairwise comparisons for applied temporal processing tasks for intervention only*

Measure	Intervention (I)	Intervention (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
LNS	Control	Metronome	-0.16	0.57	0.78	-1.29	0.97
		Calibration	0.48	0.60	0.42	-0.71	1.67
		Neurophysiology	0.11	0.55	0.85	-0.98	1.20
	Metronome	Control	0.16	0.57	0.78	-0.97	1.29
		Calibration	0.64	0.61	0.30	-0.57	1.84
		Neurophysiology	0.27	0.57	0.64	-0.86	1.39
	Calibration	Control	-0.48	0.60	0.42	-1.67	0.71
		Metronome	-0.64	0.61	0.30	-1.84	0.57
		Neurophysiology	-0.37	0.60	0.54	-1.56	0.82
	Neurophysiology	Control	-0.11	0.55	0.85	-1.20	0.98
		Metronome	-0.27	0.57	0.64	-1.39	0.86
		Calibration	0.37	0.60	0.54	-0.82	1.56
TMQ	Control	Metronome	1.23	2.06	0.55	-2.86	5.31
		Calibration	-2.03	2.16	0.35	-6.32	2.26
		Neurophysiology	0.36	1.99	0.86	-3.58	4.30
	Metronome	Control	-1.23	2.06	0.55	-5.31	2.86
		Calibration	-3.26	2.19	0.14	-7.61	1.09
		Neurophysiology	-0.86	2.05	0.67	-4.93	3.20
	Calibration	Control	2.03	2.16	0.35	-2.26	6.32
		Metronome	3.26	2.19	0.14	-1.09	7.61
		Neurophysiology	2.39	2.17	0.27	-1.90	6.69
	Neurophysiology	Control	-0.36	1.99	0.86	-4.30	3.58
		Metronome	0.86	2.05	0.67	-3.20	4.93
		Calibration	-2.39	2.17	0.27	-6.69	1.90
TSQ	Control	Metronome	-2.95	2.61	0.26	-8.12	2.22

	Calibration	2.21	2.74	0.42	-3.23	7.65
	Neurophysiology	-3.75	2.52	0.14	-8.74	1.25
Metronome	Control	2.95	2.61	0.26	-2.22	8.12
	Calibration	5.16	2.78	0.07	-0.35	10.67
	Neurophysiology	-0.80	2.60	0.76	-5.95	4.35
Calibration	Control	-2.21	2.74	0.42	-7.65	3.23
	Metronome	-5.16	2.78	0.07	-10.67	0.35
	Neurophysiology	-5.96*	2.74	0.03	-11.40	-0.52
Neurophysiology	Control	3.75	2.52	0.14	-1.25	8.74
	Metronome	0.80	2.60	0.76	-4.35	5.95
	Calibration	5.96*	2.74	0.03	0.52	11.40

Note. The pairwise comparisons are made using the estimated marginal means, for intervention only. These comparisons are made on adjusted values that account for variables that may affect responses on the applied temporal processing tasks (covariates = prodromal score and biological sex).

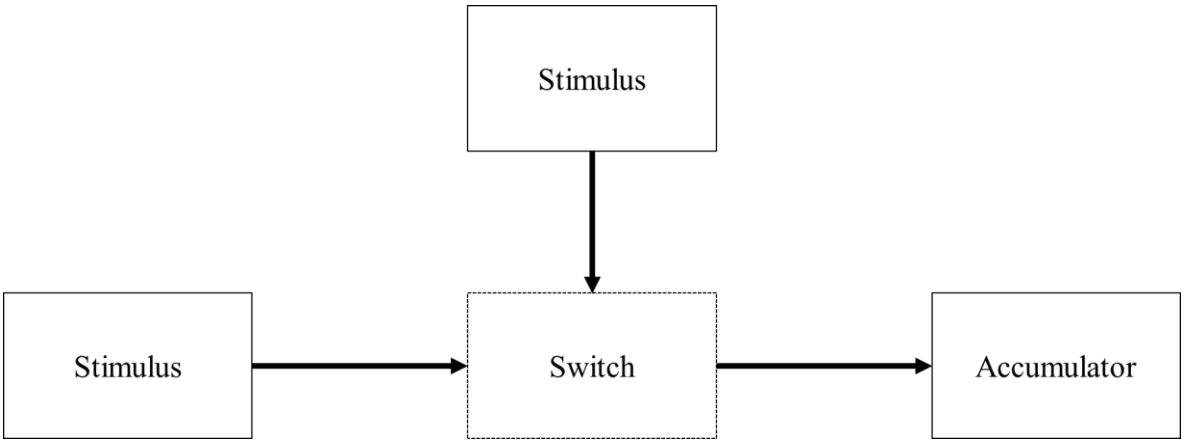


Figure 1. Internal Clock Hypothesis Diagram. When a stimulus is present, the accuracy of the switch is influenced and therefore the number of subjective time units (STUs) counted in the accumulator is also affected. Adopted from Grondin (2010).

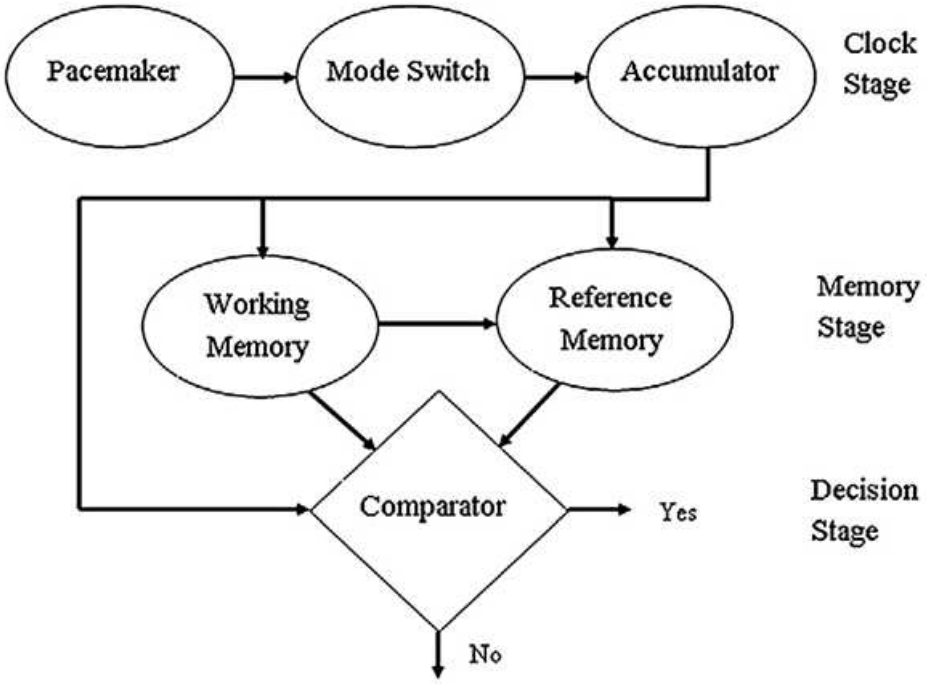


Figure 2. A Comprehensive Design of the Stages of the Scalar Expectancy Model from Gibbon (1984). Diagram adopted from Grondin (2010). This model adds a memory component to the previously established internal clock theory, thus providing an explanation for how individuals can compare and understand time relatively.

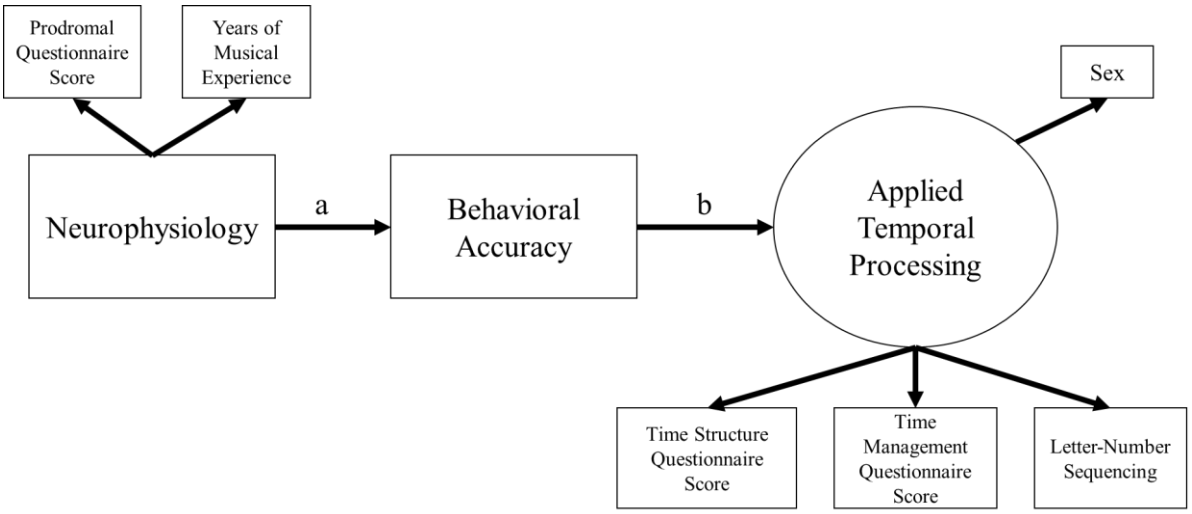


Figure 3. Conceptual Model of Hypothesis I. This model tests the hierarchical relationship of timing using a mediation model. Observed variables are placed in boxes whereas unobserved variables or factors are indicated with circles. Single-headed arrows indicate direct effects and double-headed arrows indicate covariance. This represents a bottom-up relationship where timing abilities are derived from neurophysiology and any deficits presumably manifest in applied temporal processing. This path is mediated by one's ability to perform on behavioral timing tasks.

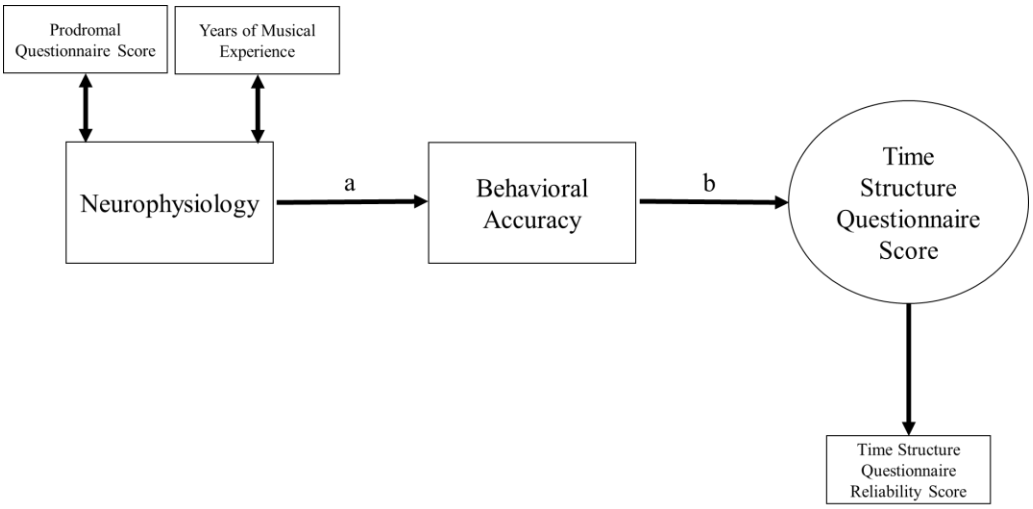


Figure 4a. New Hypothesis I conceptual models for TSQ. The observed measures combine in the proposed latent variable, applied temporal processing, were separated into 3 mediation models. This model consists of the observed measure and a reliability score found for TSQ in literature as a single indicator to control for measurement error.

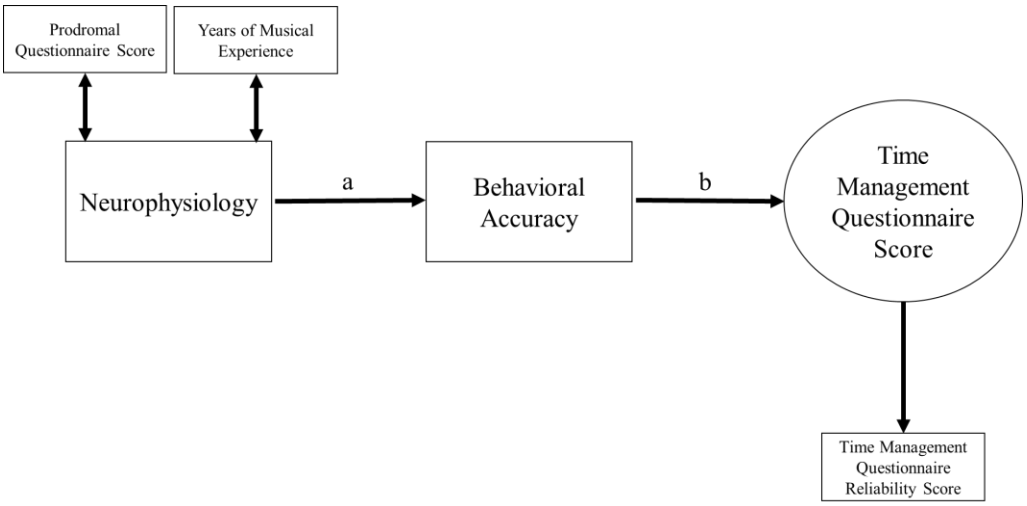


Figure 4b. New Hypothesis I conceptual models for TMQ. The observed measures combine in the proposed latent variable, applied temporal processing, were separated into 3 mediation models. This model consists of the observed measure and a reliability score found for TMQ in literature as a single indicator to control for measurement error.

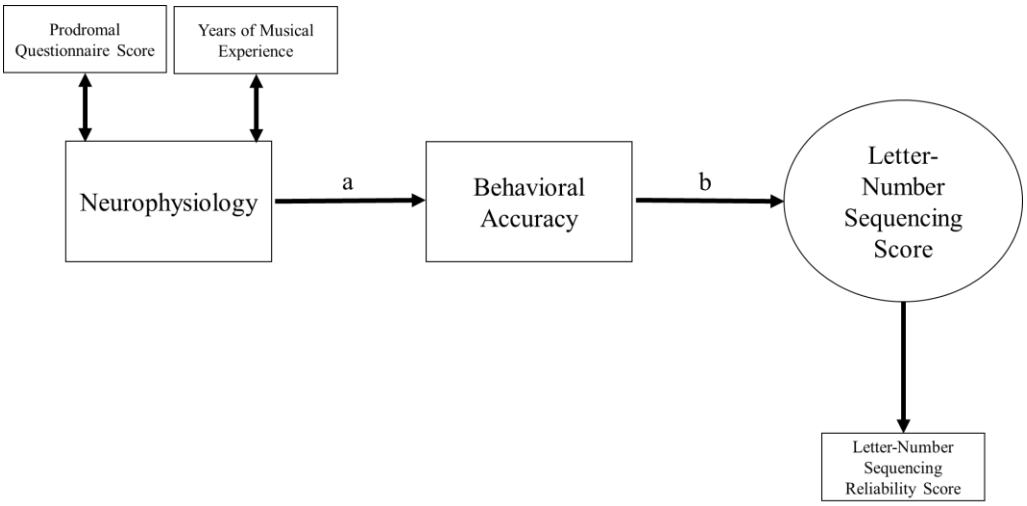


Figure 4c. New Hypothesis I conceptual models for LNS. The observed measures combine in the proposed latent variable, applied temporal processing, were separated into 3 mediation models. This model consists of the observed measure and a reliability score found for LNS in literature as a single indicator to control for measurement error.

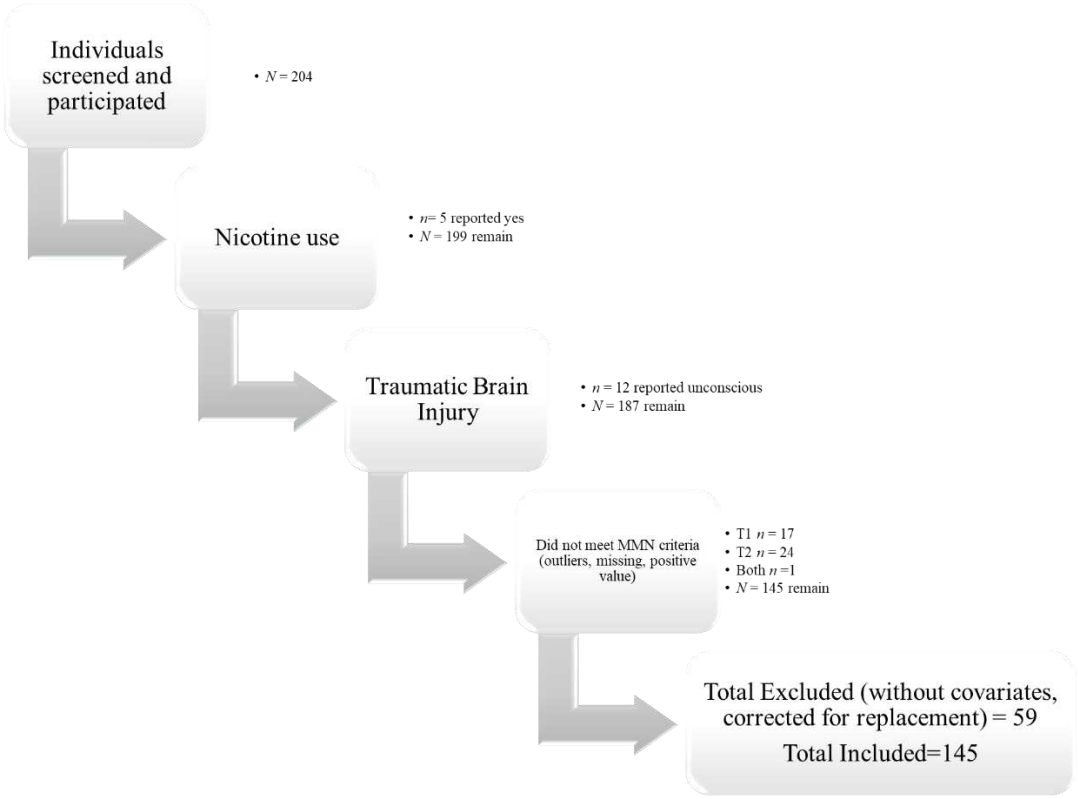


Figure 5. Inclusion/Exclusion Criteria for Participant Data. Of the original 204 participants with collected data, a total of 59 were removed. Five participants indicated nicotine use and 12 indicated they had been knocked on conscious suggesting a traumatic brain injury. Both of these factors can heavily implicate final results and therefore, the data from these individuals were removed entirely. By definition, any positive MMN amplitudes are required to be removed and a z-score transformation was applied to amplitudes to test outliers (values that were greater than 2 standard deviations above or below the mean were removed). Removing these outliers also increased normality assessments of these variables (kurtosis and skewness retained values suggesting these variables were normally distributed. At Time = 1 there were 9 participants who had positive MMN values and 8 participants who were outliers. There were 13 participants with positive MMNs and 11 with outliers at Time = 2. After removing these values there was only 1 participant with positive MMN amplitudes at both timepoints. This resulted in 145 total participants, which met criteria for a powered study.

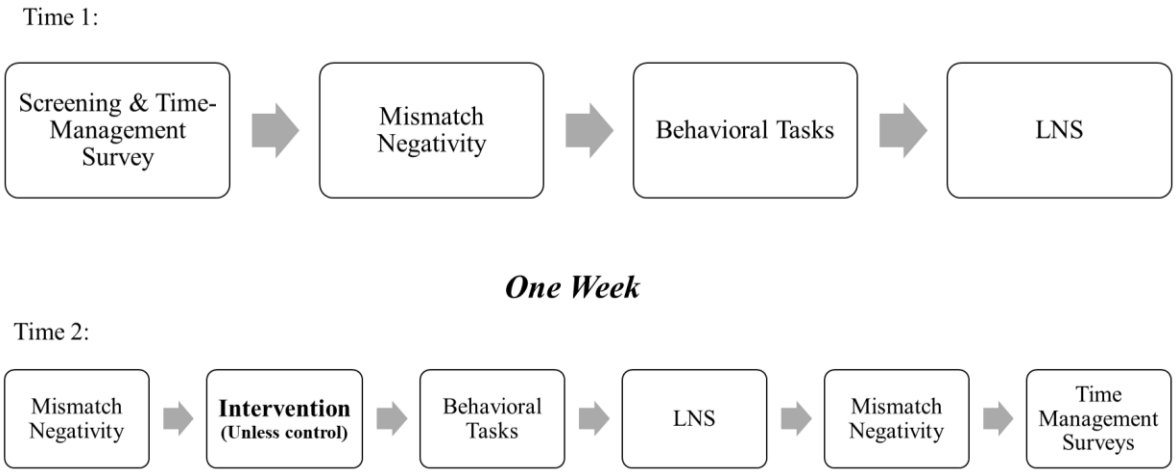


Figure 6. General Procedure. Time 1 and Time 2 sessions were completed 1 week a part at the same time of day and in the same room. The first MMN in Time 2 was used to control if there are changes following behavioral tasks. An additional MMN was given at the end of Time 2 to investigate any changes at the neurophysiological level at the conclusion of the study and was used in analysis. Participants are randomly selected for an intervention or control group. Those in the interventions completed the training between the first MMN of Time 2 and the behavioral tasks. Interventions included attention to MMN (Intervention I), metronome (Intervention II), or calibration (Intervention III). Participants randomly chosen as control completed Time 2 in a similar manner as Time 1. Time management surveys will be given at the beginning and end of the study.

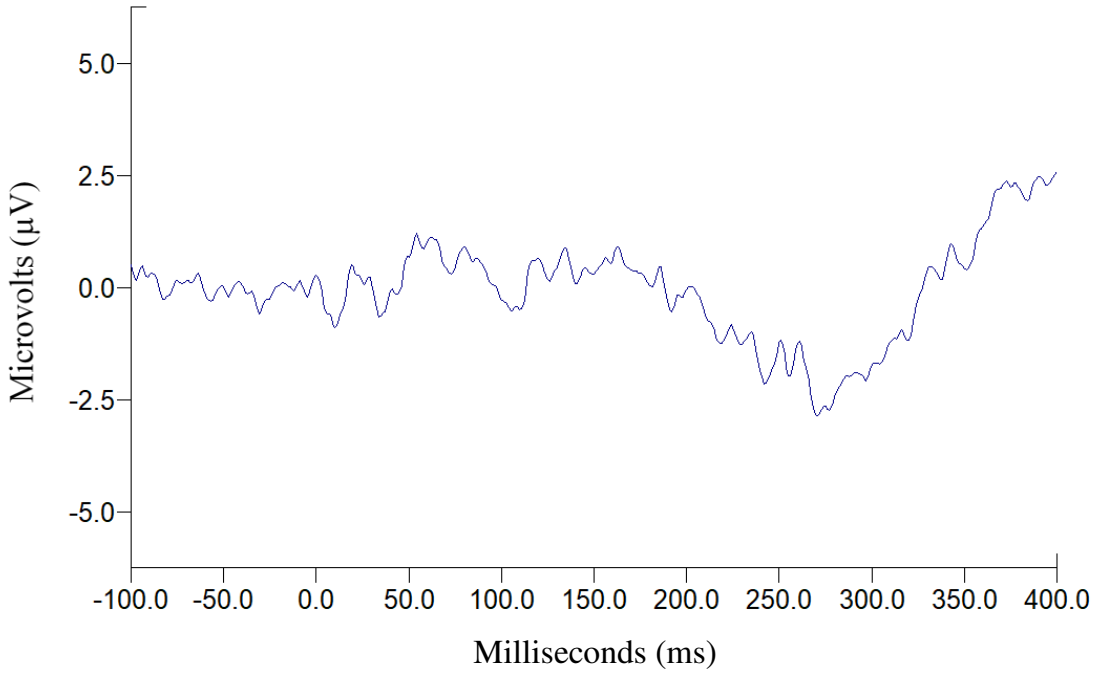


Figure 7a. Sample standard waveform for electrode placement Cz. This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents all responses to the standard stimuli (500 ms tones).

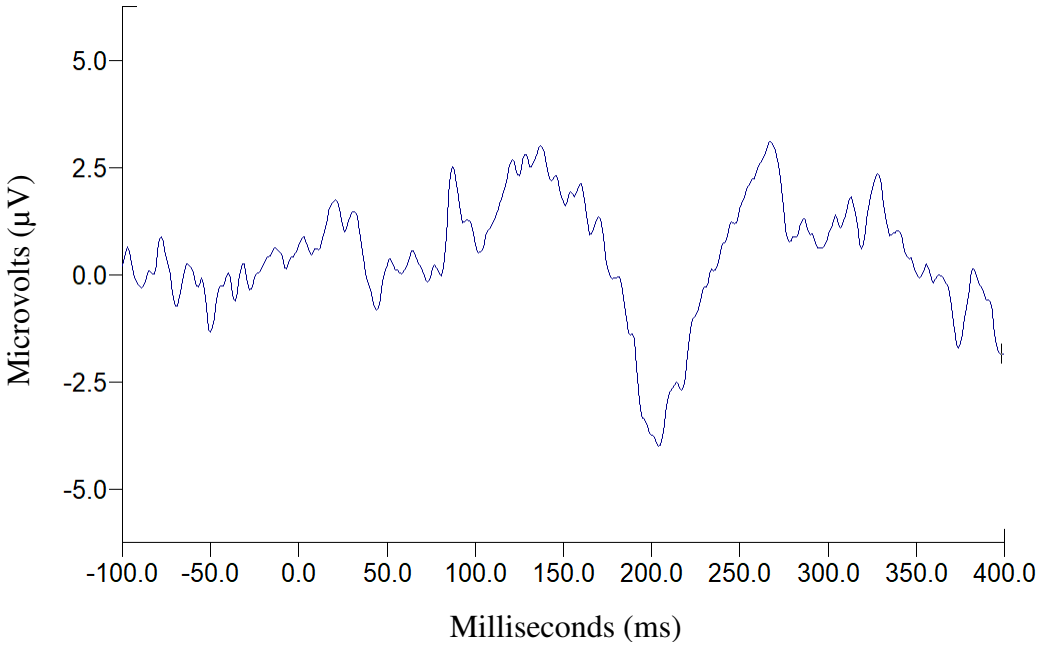


Figure 7b. Sample deviant waveform for electrode placement *Cz.* This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents all responses to the deviant stimuli (350 ms tones).

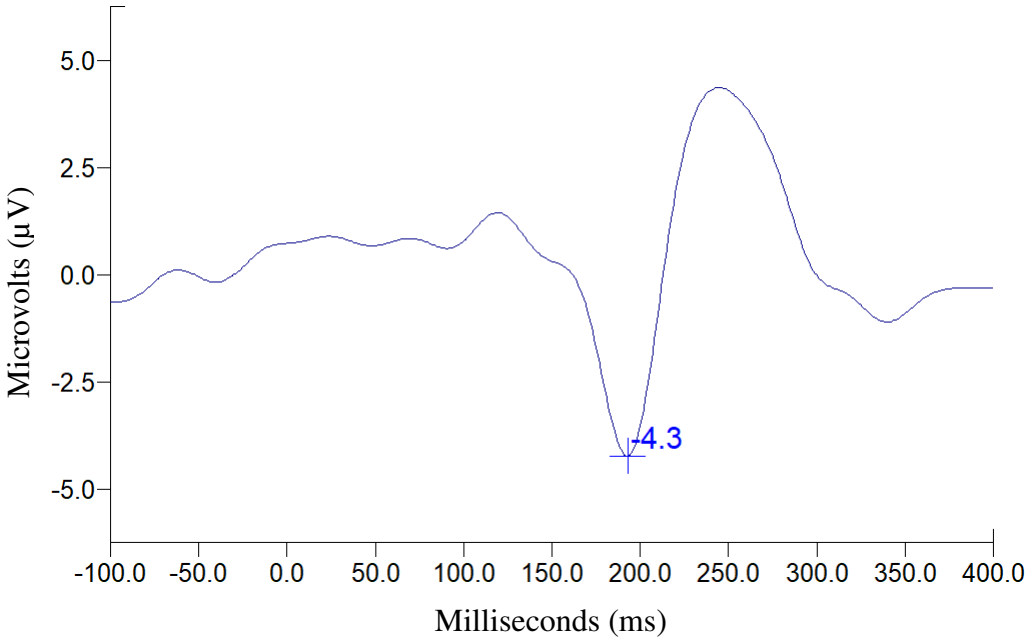


Figure 7c. Sample average waveform for electrode placement *Cz.* This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents the subtraction between the waveform created from all responses to the standard stimuli (500 ms tones; Figure xa) and the waveform created from all responses to the deviant stimuli (350 ms tones; Figure xb). The peak detection for MMN is set to the minimum amplitude between 140 ms and 210 ms following the event onset.

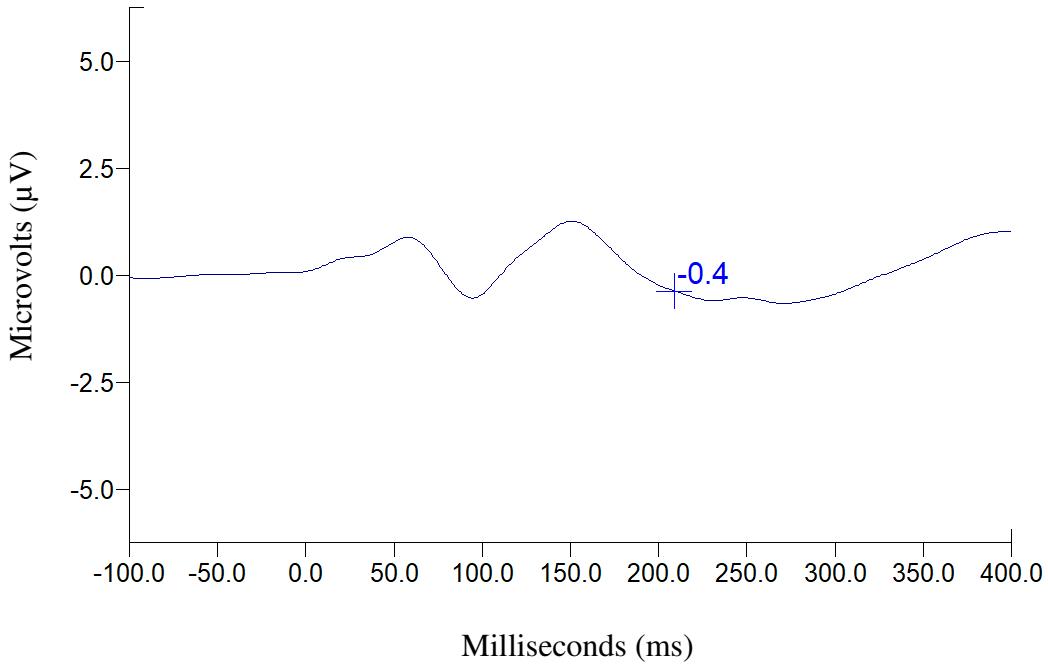


Figure 8a. Group average waveform for electrode placement *Cz.* This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents the subtraction between the waveform created from all responses to the standard stimuli (500 ms tones) and the waveform created from all responses to the deviant stimuli (350 ms tones) for all participants at Session 1 (pre-intervention). The peak detection for MMN is set to the minimum amplitude between 140 ms and 210 ms following the event onset.

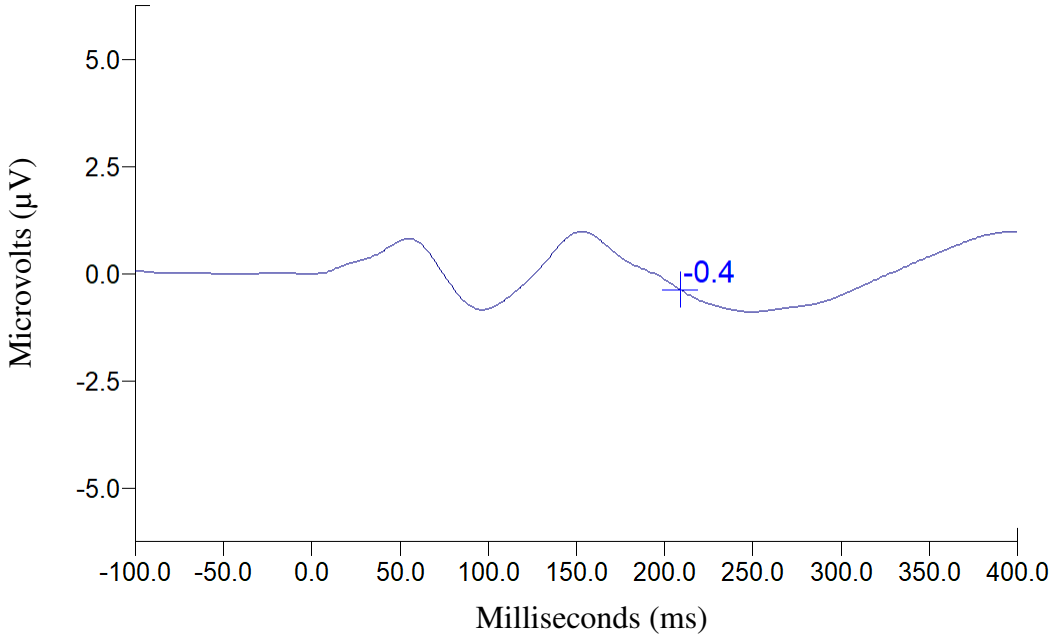


Figure 8b. Group average waveform for electrode placement *Cz*. This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents the subtraction between the waveform created from all responses to the standard stimuli (500 ms tones) and the waveform created from all responses to the deviant stimuli (350 ms tones) for all participants at Session 1 (post-intervention). The peak detection for MMN is set to the minimum amplitude between 140 ms and 210 ms following the event onset.

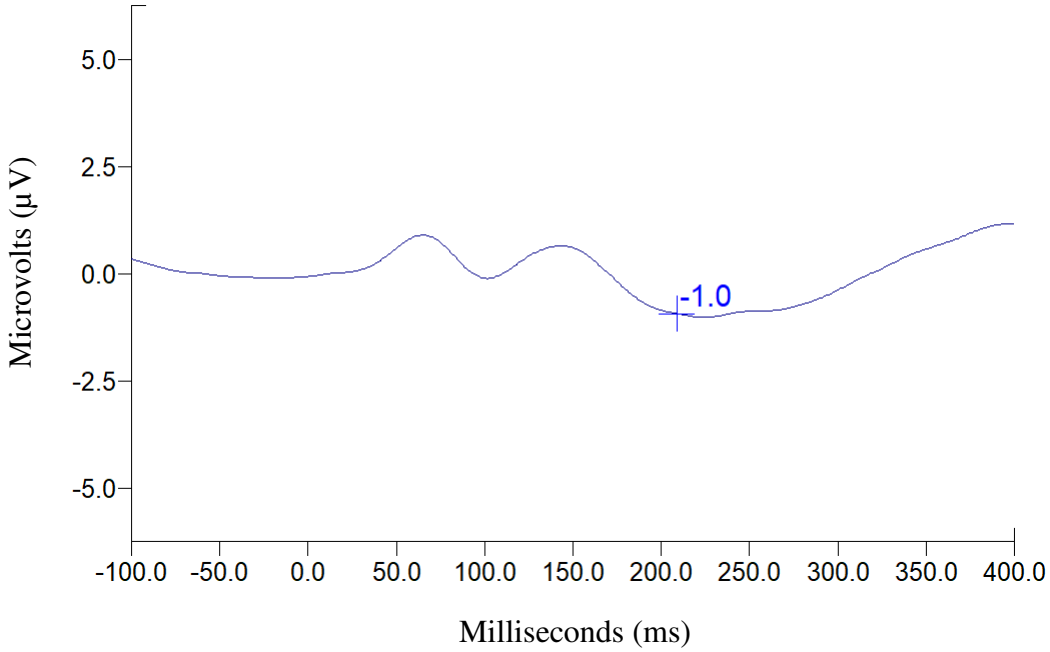


Figure 8c. Group average waveform for electrode placement *Fz.* This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents the subtraction between the waveform created from all responses to the standard stimuli (500 ms tones) and the waveform created from all responses to the deviant stimuli (350 ms tones) for all participants at Session 1 (pre-intervention). The peak detection for MMN is set to the minimum amplitude between 140 ms and 210 ms following the event onset.

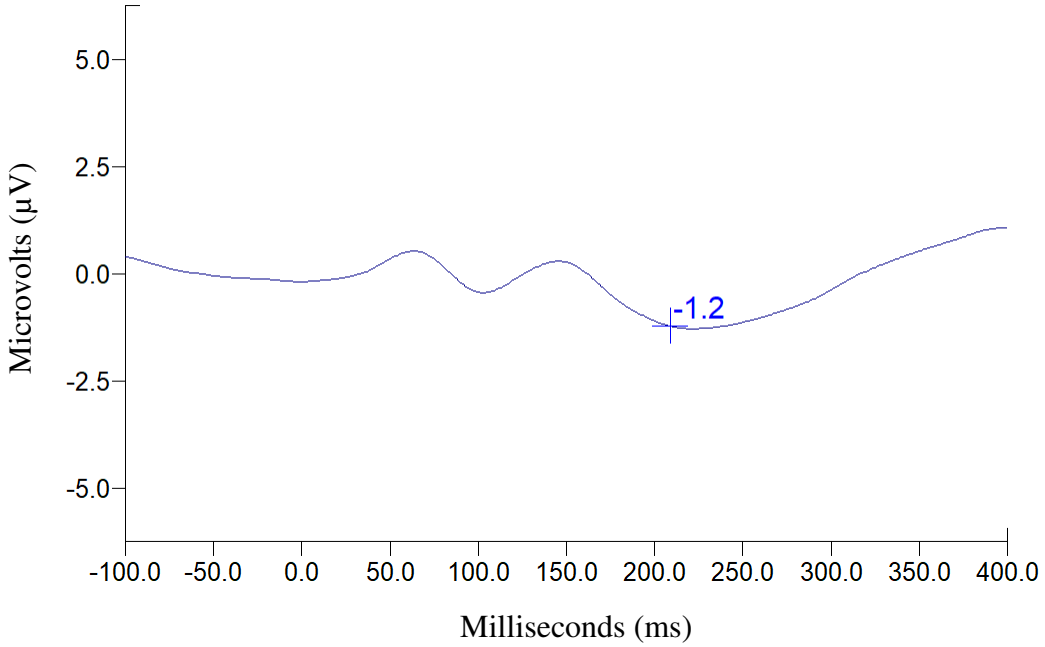


Figure 8d. Group average waveform for electrode placement *Fz.* This data shows milliseconds (ms) on the x-axis and microvolts (µV) on the y-axis. This waveform represents the subtraction between the waveform created from all responses to the standard stimuli (500 ms tones) and the waveform created from all responses to the deviant stimuli (350 ms tones) for all participants at Session 2 (post-intervention). The peak detection for MMN is set to the minimum amplitude between 140 ms and 210 ms following the event onset.

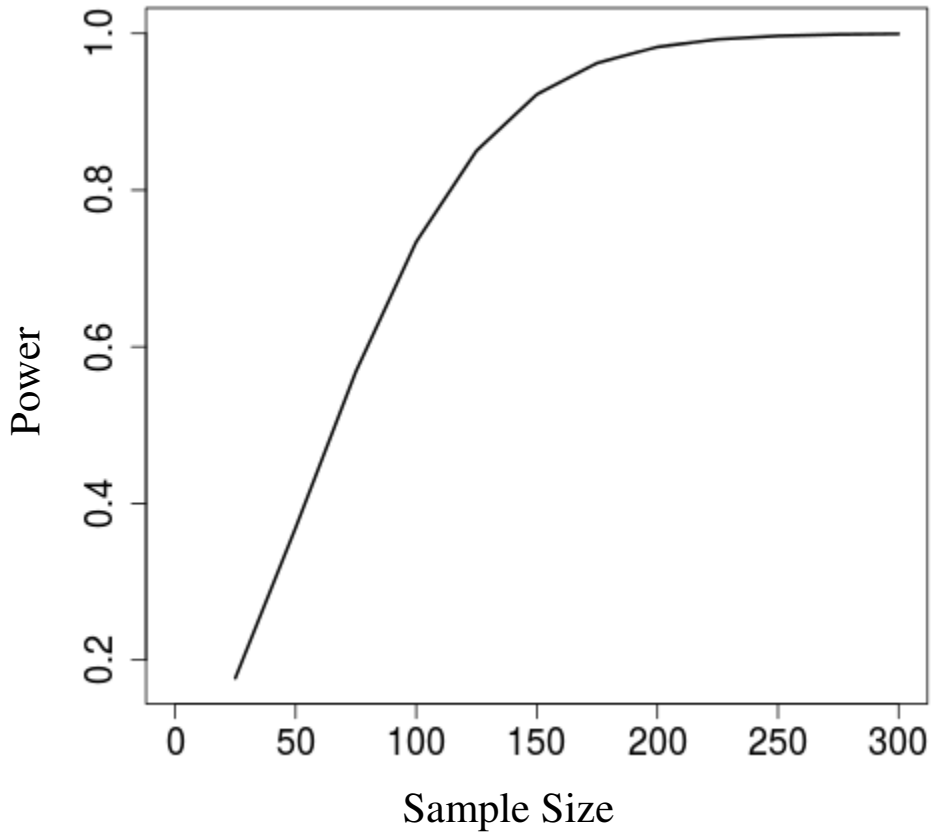


Figure 9. Power Curve for Hypothesis I. This figure was generated from Schoemann et al. (2010) using these inputs: alpha = .05, Degrees of Freedom = 18, determined from model outlined in Figure 3., step size =30, Null RMSEA = 0, Alt. RMSEA = .1.

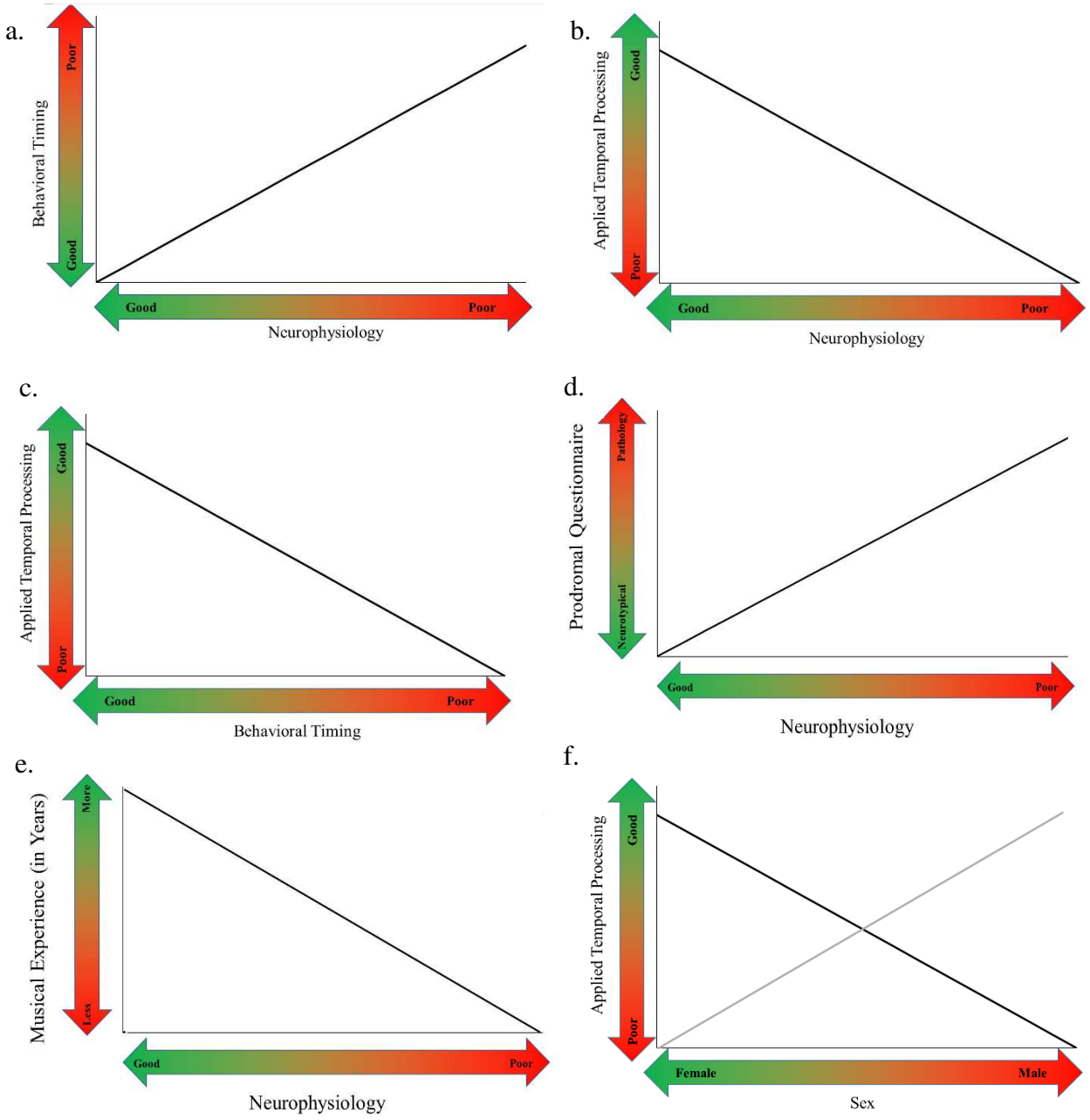


Figure 10a-f. Hypothesis I Conceptual Models of Expected Results. *Figures a-c* represent the expected direction of direct relationships. *Figures d-f* demonstrate anticipated interactions between variables and covariates. For *figure f.* the gray line represents the relation between the LNS task and sex whereas the black line represents the relation between time management survey responses and sex.

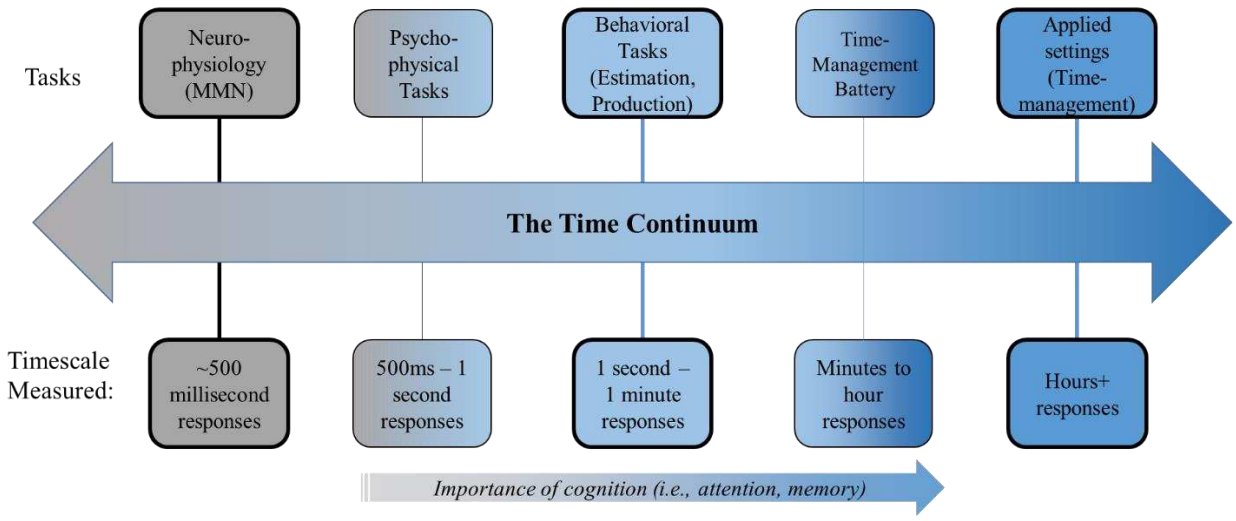


Figure 11. The time continuum. The time continuum refers to the mediation process beginning in the brain (neurophysiology), going to the lab (behavioral timing), and ending in an applied environment. In linear modeling, effects are reduced with increased distance between factors. Thus, two adjacent measures, such as neurophysiological and behavioral tasks would have comparatively larger effects. This phenomenon is due to the way indirect effects are calculated, by multiplying two standardized betas together. Since standardized betas are fractions, this will always result in a smaller fraction or effect. If this linear model is correct, the effects have to become weaker over the course of the time continuum. Therefore, adding in adjacent measures, such as psychophysical tasks and time-management tasks offer an opportunity to close the gap of timescales between tasks and obtain larger effects. Tasks with bolded box outlines were used in the present study and tasks without bolded box outlines are proposed for future research.

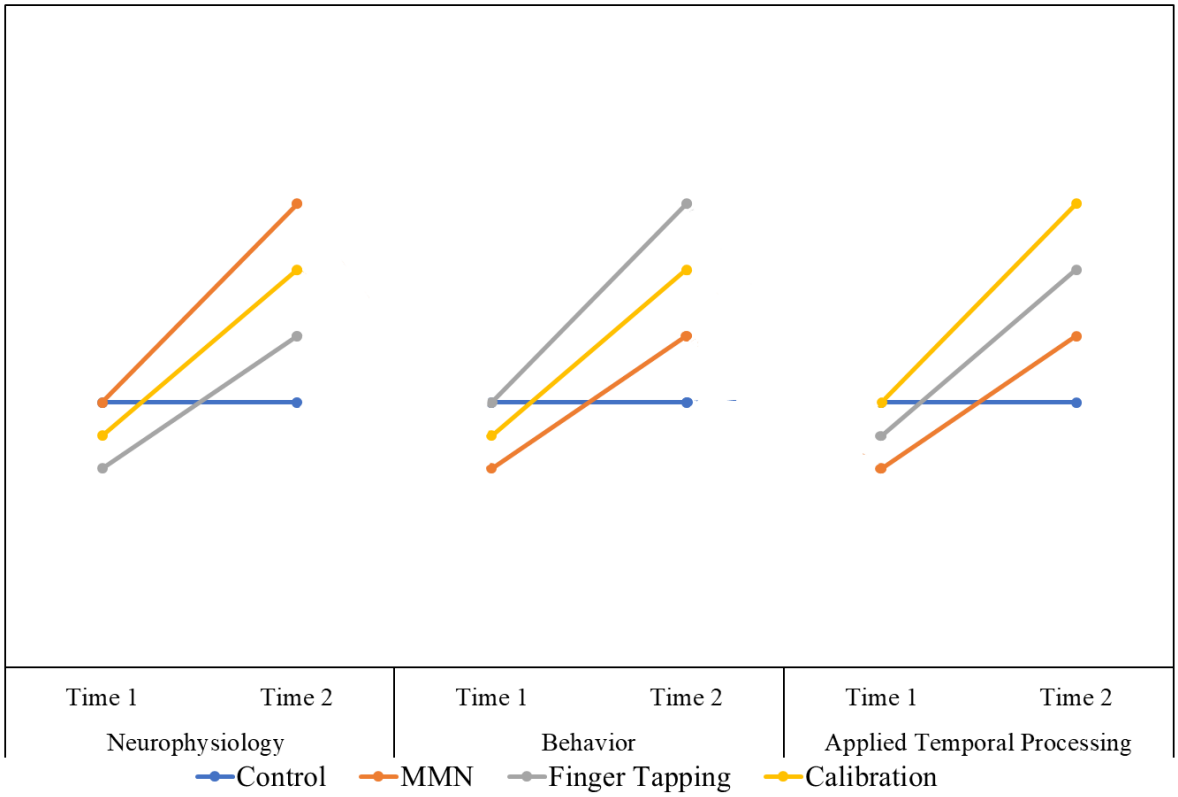


Figure 12. Hypothesis II Model of Expected Results. This figure does not demonstrate real data but portrays expected trends of change. Values were not included on the y-axis because these are expected trends. Here, a point high on the y-axis represents a better score and a low point low on the y-axis represent a low score on the respective tasks. Individuals in the control group should demonstrate no significant change across time. Participants in an intervention group should perform better at Time 2 in the task that is associated with the intervention. For example, participants in the MMN intervention are expected to improve their physiology across time, those in the metronome should increase their behavioral accuracy across time, and those in the calibration intervention should demonstrate an increase in the applied temporal processing tasks across time.

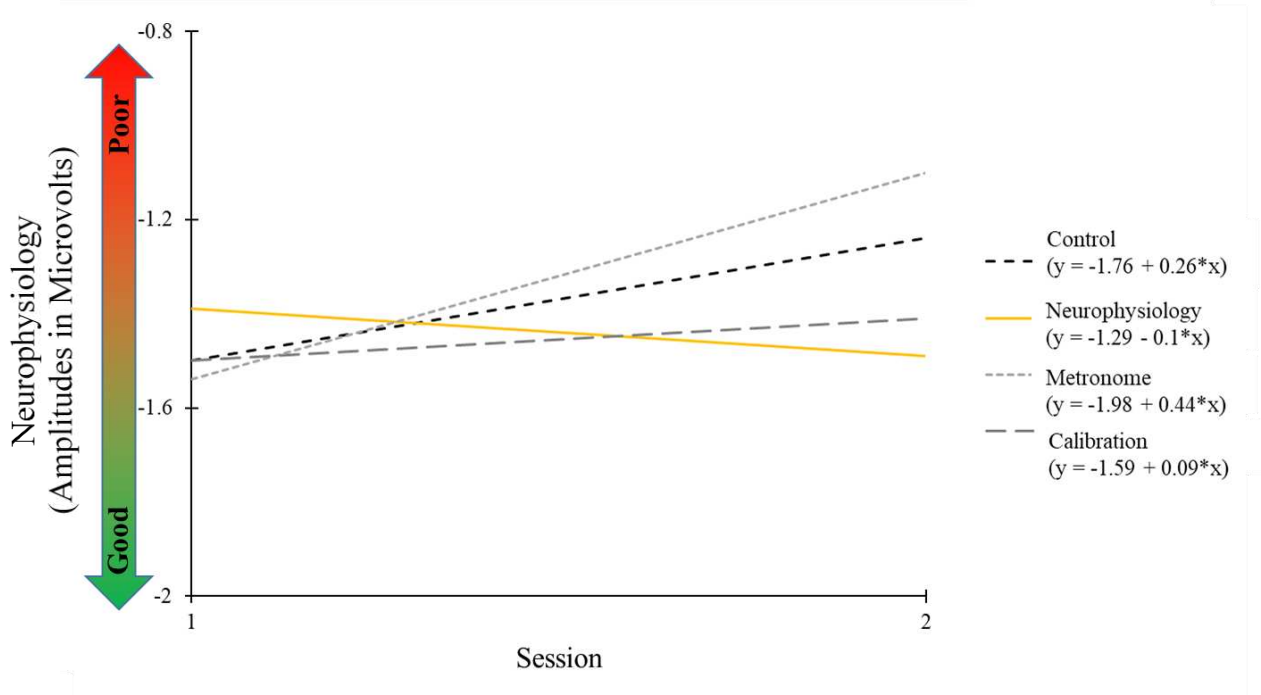


Figure 13. Intervention I (Neurophysiology) is compared to all other interventions on the neurophysiological tasks. The two time points are on the x-axis, neurophysiological amplitudes in microvolts are on the y-axis where more negative values relate to accurate timing, and each line represents a different intervention. The expectation was that those who were directly trained on MMN paradigm (Intervention I: Neurophysiology) would demonstrate the largest change in amplitude. The yellow line demonstrates that the expected trend was observed, but no significant change was evident, providing further support for this paradigm as a pre-attentive biomarker of pathology.

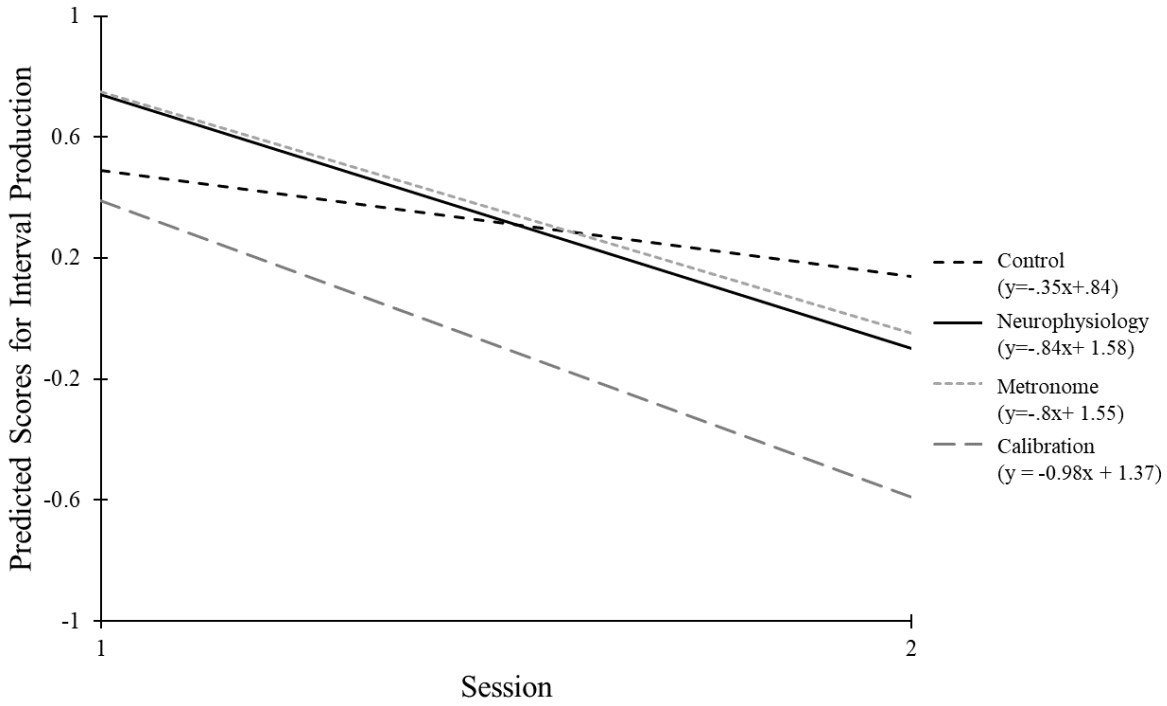


Figure 14. Simple slopes for interval production by session and intervention. Interval production was pulled into a regression framework to examine and compare simple slopes due to the significant interaction found between time and intervention. The *y*-axis demonstrates the predicted scores for interval production that are based on the original *z*-scores. Therefore, a more positive number refers to less accuracy whereas a more negative *z*-score refers to more accuracy on the production task. The *x*-axis shows the two timepoints: Time 1: pre-intervention and Time 2: post-intervention. All participants improved their accuracy on this task across time regardless of intervention, as indicated by the negative slopes. The control group demonstrated the least change across sessions, indicated by the smallest slope. The neurophysiology and metronome interventions showed similar baseline and final accuracy scores on the production task. The calibration task showed the largest change across time and the highest accuracy at the conclusion of the study.

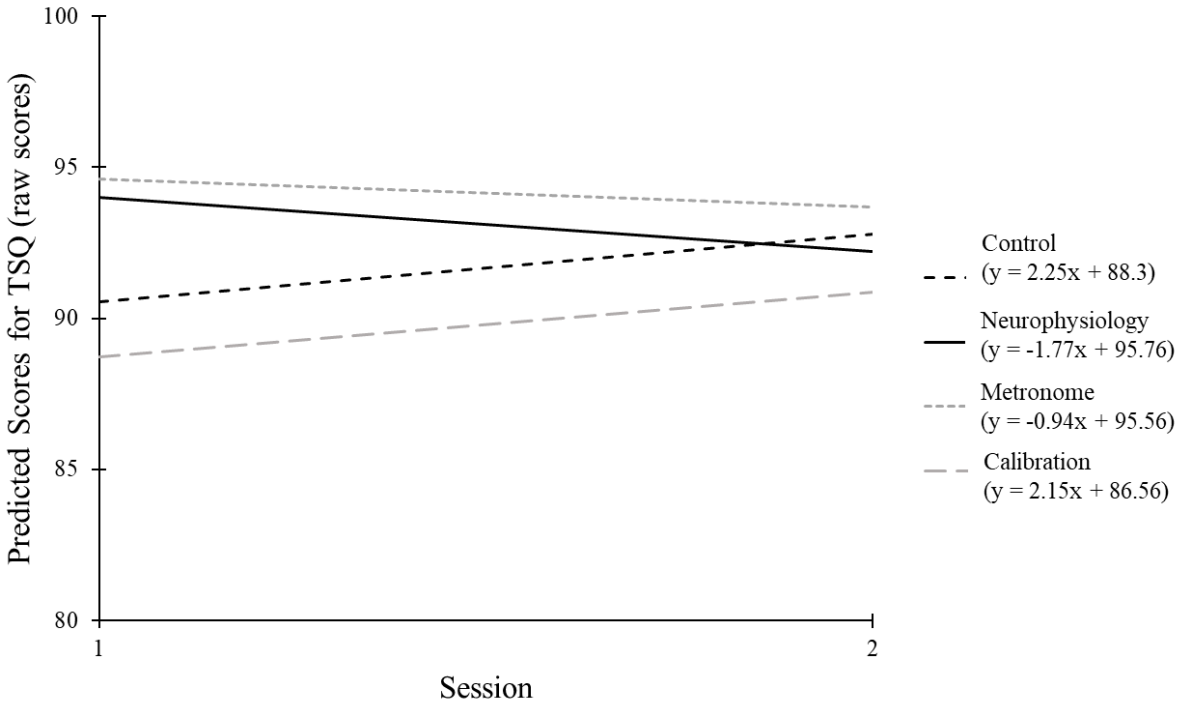


Figure 15. Simple slopes for the TSQ by session and intervention. A significant interaction of time and intervention was demonstrated when the TSQ was analyzed. Therefore, this task was pulled into a regression framework to examine and compare simple slopes. The *y*-axis demonstrates the predicted scores for the TSQ and is based on the observed scores for interpretability. Therefore, a higher value on the TSQ is indicative of higher self-assessments of time usage whereas lower scores indicate that an individual does not view themselves as efficient users of time. The *x*-axis shows the two timepoints: Time 1: pre-intervention and Time 2: post-intervention. Those in the control group and calibration intervention improved their mean TSQ, as indicated by the positive slopes. The metronome and neurophysiology interventions did not improve their TSQ mean score over time. Those in the calibration intervention were expected to outperform all other intervention groups but demonstrated the lowest mean score at baseline and at the conclusion of the study. However, these individuals retained the second largest improvement across sessions.

REFERENCES

Alay, S. & Koçak, S. (2002) Validity and reliability of time management questionnaire. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *22*, 9-13.

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the automated working memory assessment. *Educational Psychology, 28*(7), 725-734. doi:10.1080/01443410802243828

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. *American Psychologist 47*, 1597–1611.

Andersson, S., Posserud, M., & Lundervold, A. J. (2013). Early and late auditory event-related potentials in cognitively high functioning male adolescents with autism spectrum disorder. *Research In Autism Spectrum Disorders, 7*(7), 815-823. doi:10.1016/j.rasd.2013.03.007

Atienza, M. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning & Memory*, 138-150.

Audacity Team (2014). Audacity(R): Free audio editor and recorder [Computer program]. http://audacity.sourceforge.net/.

Balci, F., Freestone, D., Simen, P., Desouza, L., Cohen, J. D., & Holmes, P. (2011). Optimal temporal risk assessment. *Frontiers in Integrative Neuroscience, 5*. doi:10.3389/fnint.2011.00056

Bartha-Doering, L., Deuster, D., Giordano, V., am Zehnhoff-Dinnesen, A., & Dobel, C. (2015). A systematic review of the mismatch negativity as an index for auditory sensory memory:

From basic research to clinical and developmental perspectives. *Psychophysiology*, *52*(9), 1115–1130. https://doi.org/10.1111/psyp.12459

Bausenhart, K. M., Luca, M. D., & Ulrich, R. (2018). Assessing duration discrimination: psychophysical methods and psychometric function analysis. *Timing and Time Perception: Procedures, Measures, and Applications*, 52-78. doi:10.1163/9789004280205_004

Block, RA., Hancock, PA., Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychol; 134*(3): 330-43.

Bodatsch, M., Ruhrmann, S., Wagner, M., Mller, R., Schultze-Lutter, F., Frommann, I., … Brockhaus-Dumke, A. (2011). Prediction of psychosis by mismatch negativity. *Biological Psychiatry*, *69*(10), 959–966. https://doi.org/10.1016/j.biopsych.2010.09.057

Britton, B. K., & Tesser, A. (1991). Effects of time management practices on college grades. *Journal of Educational Psychology*, *83*(3), 405–410. https://doi.org/10.1037/0022-0663.83.3.405

Bond, M. J., & Feather, N. T. (1988). Time structure questionnaire. PsycTESTS Dataset. doi:10.1037/t01384-000

Box, G. E. P., 1949. A general distribution theory for a class of likelihood criteria. *Biometrika, 36*: 317–346.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755-765. doi:10.1038/nrn1764

Caldara, R., Deiber, MP., Andrey, C., Michel, GM., Thut, G., Hauert, CA. (2004) Actual and mental motor preparation and execution: A spatiotemporal ERP study. *Experimental Brain Research.159*: 389–399.

Campbell, A. M., & Davalos, D. B. (2015). Levels of attention and task difficulty in the modulation of interval duration mismatch negativity. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.01619

Caruso, V. C., & Balaban, E. (2015). Auditory perceptual category formation does not require perceptual warping. *Journal of Cognitive Neuroscience*, *27*(8), 1659-1673. doi:10.1162/jocn_a_00806

Claessens, B., Wendelien van Eerde, Rutte, C., Roe, R. (2007) A review of the time management literature, *Personnel Review*, *36*(2), pp.255-276, https://doi.org/10.1108/00483480710726136

Cohen, J. (1988). Statistical power analysis for the behavioral sciences 2nd edn.

Cohen, B. (2008). Explaining Psychological Statistics. John Wiley & Sons.

Ciullo, V., Spalletta, G., Caltagirone, C., Jorge, R. E., & Piras, F. (2015). Explicit time deficit in schizophrenia: Systematic review and meta-analysis indicate it is primary and not domain specific. *Schizophrenia Bulletin*, *42*(2), 505-518. doi:10.1093/schbul/sbv104

Dalebout, S. D., & Fox, L. G. (2001). Reliability of the mismatch negativity in the responses of individual listeners. *J Am Acad Audiol,* 12: 245-253.

Davalos, D. B. , Kisley , M. A. , & Freedman , R. (2005). Behavioral and electrophysiological indices of temporal processing dysfunction in schizophrenia. *Journal of Neuropsychiatry and Clinical Neurosciences*, *17*, 517 – 525.

Delevoye-Turrell, Y., Wilquin, H., and Giersch, A. (2012). A ticking clock for the production of sequential actions: where does the problem lie in schizophrenia? *Schizophr. Res. 135*, 51–54. doi: 10.1016/j.schres.2011.12.020

Eagleman, D. M., & Holcombe, A. O. (2002). Causality and the perception of time. *Trends in*

Cognitive Sciences, 6(8), 323-325.

Eagleman, D. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology*, 18(2), 131-136.

Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap. New York: Chapman & Hall.

Elvevåg, B., Mccormack, T., Gilbert, A., Brown, G. D., Weinberger, D. R., & Goldberg, T. E. (2003). Duration judgements in patients with schizophrenia. *Psychological Medicine*, *33*(7), 1249-1261. doi:10.1017/s0033291703008122

Foster, S., Kisley, M., Davis, H., Diede, N., Campbell, A., & Davalos, B. (2013). Cognitive function predicts neural activity associated with pre - attentive temporal processing . *Neuropsychologia*, *51*(2), 211–9. https://doi.org/10.1016/j.neuropsychologia.2012.09.017.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology, 35*, 1-36.

Francis-Smythe, J. A., & Robertson, I. T. (1999). On the relationship between time management and time estimation. *British Journal of Psychology*, *90*, 333–347. https://doi.org/10.1348/000712699161459

Franssen, V., Vandierendonck, A., & Van Hiel, A. (2006). Duration estimation and the phonological loop: articulatory suppression and irrelevant sounds. *Psychol. Res*. *70*, 304–316.

Fritz, M. S., & Mackinnon, D. P. (2007). Required Sample Size to Detect the Mediated Effect. *Psychological Science*, *18*(3), 233-239. doi:10.1111/j.1467-9280.2007.01882.x

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*(2), 289-344. doi:10.1037/0033-295x.107.2.289

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin Neurophysiol*, *120*(3), 453–463. https://doi.org/10.1016/j.clinph.2008.11.029

Gebuis, T., & Reynvoet, B. (2013). The neural mechanisms underlying passive and active processing of numerosity. *NeuroImage, 70*, 301-307. doi:10.1016/j.neuroimage.2012.12.048

George, D., & Mallery, M. (2010). SPSS for windows step by step: A simple guide and reference, 17.0 update (10a ed.) Boston: Pearson

Gibbon, J., Church, R. M., & Meck, W. (1984). Scalar timing in memory. In J. Gibbon and L. Allan (Eds*.), Annals of the New York Academy of Sciences*, 423: *Timing and time perception*: 52-77. New York: New York Academy of Sciences.

Gravetter, F., & Wallnau, L. (2014). Essentials of Statistics for the Behavioral Sciences (8th ed.). Belmont, CA: Wadsworth.

Gomot, M., Blanc, R., Clery, H., Roux, S., Barthelemy, C., & Bruneau, N. (2011). Candidate electrophysiological endophenotypes of hyper-reactivity to change in autism. *Journal Of Autism & Developmental Disorders*, *41*(6), 705-714 10p. doi:10.1007/s10803-010-1091-y

Grondin S. (2008) Methods for studying psychological time. *Psychology of time*: Emerald; 51–74.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561-582.

Gropper, R. J., & Tannock, R. (2009). A pilot study of working memory and academic

achievement in college students with ADHD. *Journal of Attention Disorders*, *12*(6), 574-581. doi:10.1177/1087054708320390

Gros, A., Giroud, M., Bejot, Y., Rouaud, O., Guillemin, S., Eboulã©, C. A., . . . Martin, M. L. (2015). A time estimation task as a possible measure of emotions: Difference depending on the nature of the stimulus used. *Frontiers in Behavioral Neuroscience*, *9*. doi:10.3389/fnbeh.2015.00143

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408-420. doi:10.1080/03637750903310360

Heim, S., Choudhury, N., & Benasich, A. A. (2016). Electrocortical dynamics in children with a language-learning impairment before and after audiovisual training. *Brain Topography, 29*, 459–476. http://doi.org/10.1007/s10548-015-0466-y

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structure analysis conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.

Ising, H. K., Veling, W., Loewy, R. L., Rietveld, M. W., Rietdijk, J., Dragt, S., ... & van der Gaag, M. (2012). The validity of the 16-item version of the prodromal questionnaire (PQ-16) to screen for ultra-high risk of developing psychosis in the general help-seeking population. *Schizophrenia bulletin*, *38*(6), 1288-1296.

Ivry, R. B., & Spencer, R. M. (2004). The neural representation of time. *Current Opinion in Neurobiology*, *14*(2), 225-232. doi:10.1016/j.conb.2004.03.013

Jones, M.R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*, 323-355.

Kamiyama, K. S., & Okanoya, K. (2014). Synchronized tapping facilitates learning sound

sequences as indexed by the P300. *Frontiers in Human Neuroscience*, *8*. doi:10.3389/fnhum.2014.00826

Kharlamov, V., Campbell, K., & Kazanina, N. (2011). Behavioral and electrophysiological evidence for early and automatic detection of phonological equivalence in variable speech inputs. *Journal of Cognitive Neuroscience*, 23, 3331-3342.

Kramer, P., Bressan, P., and Grassi, M. (2011). Time estimation predicts mathematical intelligence. *PLoS One* 6: e28621. doi: 10.1371/journal. pone.0028621

Kraus, N., Mcgee, T., Carrell, T., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity associated with speech discrimination training. *Journal of cognitive neuroscience,* 25-32.

Kuznetsov, N. A., & Wallot, S. (2011). Effects of accuracy feedback on fractal characteristics of time estimation. *Frontiers in Integrative Neuroscience*, 5. doi:10.3389/fnint.2011.00062

Light, G. A., & Näätänen, R. (2013). Mismatch negativity is a breakthrough biomarker for understanding and treating psychotic disorders. *Proc Natl Acad Sci U S A*, *110*(38), 15175

Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students time management: Correlations with academic performance and stress. *Journal of Educational Psychology*, *82*(4), 760-768. doi:10.1037//0022-0663.82.4.760

Martin, L., Davalos, D., & Kisley, M (2009) Nicotine enhances automatic temporal processing as measured by the mismatch negativity waveform. *Nicotine & Tobacco Research*, *11*(6), 698-706. doi:10.1093/ntr/ntp052

Matthews, W. J., & Meck, W. H. (2014). Time perception: The bad news and the good. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(4), 429-446. doi:10.1002/wcs.1298

Miller, G. A. (1956) The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97

Miller, S., Zhang, Y., & Nelson, P. (2016). Neural correlates of phonetic learning in postlingually deafened cochlear implant listeners. *Ear and Hearing*, *37*(5), 514-528. doi:10.1097/aud.0000000000000287

Mioni, G., Stablum, F., Mcclintock, S. M., & Grondin, S. (2014). Different methods for reproducing time, different results. *Attention, Perception, & Psychophysics*, *76*(3), 675-681. doi:10.3758/s13414-014-0625-3

Mioni, G. (2018). Methodological issues in the study of prospective timing. *Timing and Time Perception: Procedures, Measures, and Applications*, 79-97. doi:10.1163/9789004280205_005

Muthén, L. K., & Muthén, B. O. (1998-2011). Mplus user's guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, *29*(1), 81-117. doi:10.2333/bhmk.29.81

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544-2590.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, & Winkler I. (2001) Primitive intelligence in the auditory cortex. *Trends Neurosci*, 24: 283–8.

Näätänen (1984) In search of a short-duration memory trace of a stimulus in the human brain. *Human action and personality. Essays in honour of Martti Takala.* 29–43.

Oberfeld, D., Thönes, S., Palayoor, B. J., & Hecht, H. (2014). Depression does not affect time

perception and time-to-contact estimation. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.00810

Okano, M., Shinya, M., & Kazutoshi, Kudo. (2017) Paired synchronous rhythmic finger tapping without an external timing cue shows greater speed increases relative to those for solo tapping. *Scientific Report,* 7:43987.

Pantlin, L. and Davalos, D. (2016) Neurophysiology for detection of high-risk for psychosis. *Schizophrenia Research and Treatment.* vol. 2016, Article ID 2697971, 5 pages.

Petersen, B., Weed, E., Sandmann, P., Brattico, E., Hansen, M., Sã¸rensen, S., & Vuust, P. (2015). Brain responses to musical feature changes in adolescent cochlear implant users. *Front. Hum. Neurosci.*

Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition*, 11, 77-82.

Razali, S. N., Rusiman, M. S., Gan, W. S., & Arbin, N. (2018). The Impact of time management on students' academic achievement. *Journal of Physics: Conference Series*, 995, 012042. doi:10.1088/1742-6596/995/1/012042

Rammsayer, T., & AltenmüLler, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, *24*(1), 37-48. doi:10.1525/mp.2006.24.1.37

Rogers, M., Hwang, H., Toplak, M., Weiss, M., & Tannock, R. (2011). Inattention, working memory, and academic achievement in adolescents referred for attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *17*(5), 444-458. doi:10.1080/09297049.2010.544648

Russo, N., Nicol, T., Zecker, S., Hayes, E., & Kraus, N. (2004). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, 95-103.

Sanna, LJ., Parks, Craig D., Chang, Edward C., Carter, Seth E. (2005). The hourglass is half full or half empty: Temporal framing and the group planning fallacy. *Group Dynamics: Theory, Research, and Practice. 9*(3): 173–188. doi:10.1037/1089-2699.9.3.173.

Schröger, E., Paavilainen, P., & Näätänen, R. (1994). Mismatch negativity to changes in a continuous tone with regularly varying frequencies. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *92*(2), 140-147. doi:10.1016/0168-5597(94)90054-x

Schoemann, A. M., Preacher, K. J., & Coffman, D. L. (2010, April). Plotting power curves for RMSEA [Computer software]. Available from http://quantpsy.org/.

Sculthrope L. D., Ouellet D. R., & Campbell K. B. (2009) MMN elicitation during natural sleep to violations of an auditory pattern. *Brain Res*; 1290: 52-62. doi: 10.1016/j.brainres.2009.06.013.

Shelley, A. M., Ward, P. B., Catts, S. V, Michie, P. T., Andrews, S., & McConaghy, N. (1991). Mismatch negativity: an index of a preattentive processing deficit in schizophrenia. *Biol Psychiatry*, *30*(10)*,* 1059–1062.

Silva, M. A. (2008). Development of the WAIS-III: A brief overview, history, and description. *Graduate Journal of Counseling Psychology*, *1*(1).

SPSS Inc. Released 2017. PASW statistics for windows, Version 25.0. Chicago: SPSS Inc.

Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., & Näätänen, R. (2002). Top-down effects on the initially stimulus-driven auditory organization. *Cognitive Brain Research*, *13*, 393–405.

Tamminen, H., Peltola, M., Kujala, T., & Näätänen, R. (2015). Phonetic training and non-native speech perception — New memory traces evolve in just three days as indexed by the

mismatch negativity (MMN) and behavioral measures. *International Journal of Psychophysiology,* 23-29.

Tamminen, H. & Peltola, M. (2015) Non-native memory traces can be further strengthened by short term phonetic training. *The Scottish Consortium for ICPhS,* 285.

Tao, L., Sun, H., Li, Q., & Chen, X. (2015). Mismatch negativity, social cognition, and functional outcomes in patients after traumatic brain injury. *Neural Regeneration Research, 10*(4), 618. doi:10.4103/1673-5374.155437

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs*, *77*, 576.

Tremblay, K., Kraus, N., Carrell, T.D., & McGee, T. (1997). Central auditory system plasticity: Generalization to novel stimuli following listening training, *J. Acoust. Soc. Am. 102*, 3762-3773.

Tong, Y., Melara, R., & Rao, A. (2009). P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Research,* 80-88.

Umbricht, D., & Krljes, S. (2005). Mismatch negativity in schizophrenia: a meta-analysis. *Schizophr Res*, *76*(1), 1–23. https://doi.org/10.1016/j.schres.2004.12.002

Vanneste, S., Pouthas, V., Wearden, J.H., 2001. Temporal control of rhythmic performance: a comparison between young and old adults. *Exp. Aging Res. 27*(1), 83–102.

Vuoksimaa, E. P. (2004). Sex differences in cognitive functions: A study of same sex and opposite-sex twin pairs. *Master's Thesis*, 1-42.

Wearden, J., & Farrar, R. (2007). Effects of feedback and calibration on the verbal estimation of the duration of tones. *Acta Psychologica, 126*(1), 1-17. doi:10.1016/j.actpsy.2006.10.003

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: conformity and

violations. *Quarterly Journal of Experimental Psychology, 61*(4), 569-587. doi:10.1080/17470210701282576

Wearden, J. (2014). Modelling behaviour on classical timing tasks: verbal estimation, production, and reproduction. *Procedia - Social and Behavioral Sciences*, *126*, 34-35. doi:10.1016/j.sbspro.2014.02.307

Weber, P., Hetzel, P., Fahnenstich, H., & Lütschg, J. (2006). Mismatch detection: An auditory event-related study to detect differences in pre-attentive and learning capability of fullterm and preterm newborns. *Neuropediatrics*, 37(06). doi:10.1055/s-2006-974107

Wechsler, D. (2008). Wechsler adult intelligence scale WAIS-IV; technical and interpretive manual. San Antonio, Tex.: Pearson.

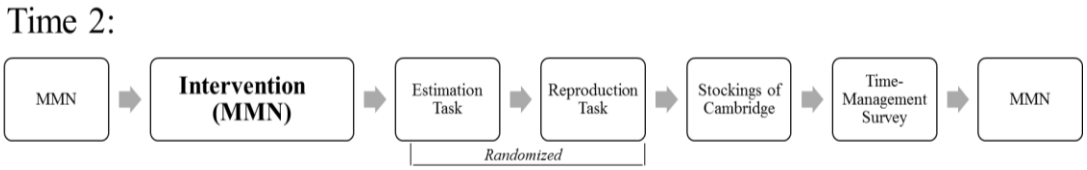
Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2010). Training the brain to weight speech cues differently: A study of Finnish second-language users of English. *Journal of Cognitive Neuroscience, 22*(6), 1319-1332. doi:10.1162/jocn.2009.21272

Zakay, D. (1990). The evasive art of subjective time measurement: Some methodological dilemmas. *Cognitive models of psychological time*, 59-84. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Zeiler, M. D., & Hoyert, M. S. (1989). Temporal reproduction. *Journal of the Experimental Analysis of Behavior*, *52*(2), 81–95. http://doi.org/10.1901/jeab.1989.52-81

APPENDIX A

Directions for Intervention I: Attention to MMN



1. Do not remove electrodes, do not remove the participant from the room
2. Change the blue HDMI monitor cords from the neuroscan to the eprime computer so you can duplicate the screen onto the participant's monitor
3. Pull up files
    a. On eprime folder: Pantlin → Pilot Study
        i. Filename: Lara Behavioral Intervention III Attention
    b. Pull up new recording in neuroscan
    c. Save file in Folder: Pantlin → Behavioral → Pilot study → Intervention III
        i. Filename: ID#IV3MMN

Script:

"We are going to run this task again.

What you were/will listen to is a series of beeping.

Some of the beeps are different from other beeps.

When you hear a beep that differs, I would like you to press mouse key.

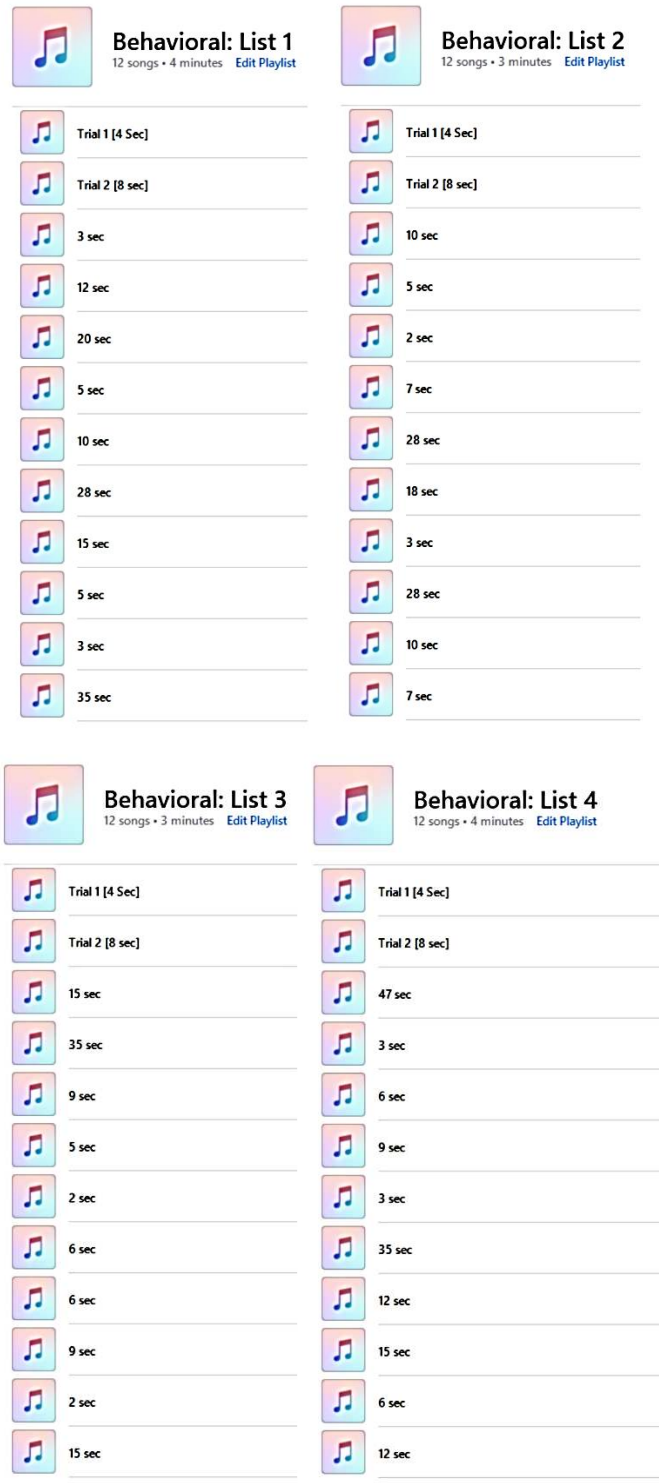
Do you have any questions before we begin?

- Run the participant on MMN for 6 minutes
- Take the participant into the behavioral room with the electrodes still on and run the participant on the behavioral tasks

    - Use the T1 Script
- When they have completed the behavioral tasks, run the participant on 12 more minutes of MMN

APPENDIX B

Interval Lists



APPENDIX C

Directions for Verbal Estimation Task

Directions:

1. Make sure you know which task they are getting first/second
2. Prepare your recording sheet
3. Prepare the playlist (should be on shuffle)
4. Read script:

Script:

RA will read the directions:

"I am going to play tones out loud.

There are two tones for each trial.

I want you to listen carefully and then tell me how much time you believe has elapsed between the two tones, as accurately as possible.

I cannot answer any questions or repeat the process, so make sure you pay attention.

Let's try one."

- RA completes first trial of verbal estimation with the participant (4 seconds).
- If the participant understands, you do not need to do the second trial (8 seconds).
- If they want another, give the participant a second trial.
- **Remember you cannot tell the participant how well they did or if they got it right.**

"Very good, now we are going to begin the experiment.

I cannot tell you how well you did or show you any results until after you complete the entire study (both time frames).

Remember that I also cannot answer any questions and I cannot repeat any trials.

Do you have any questions before we begin?

I will say 'okay, ready? (or whatever you want, just okay or ready will work)*' between each trial.*

Ready? Begin."

- RA has computer press *start* on first play list. Pause the playlist so it doesn't go to the next one.

RA asks participant "How much time has passed?"

- Record the interval that you played and the answer the participant gave.
- Press play on the next interval on the list. *Repeat*
- When you are nearing the end, you can say things like "Only 3 more trials left or last one."

Notes: Make sure you hit play/pause between EVERY TRIAL. Make sure that you are on shuffle and you write down each interval as it appears and write down the exact response from the participant. They are allowed to give you decimals, if they ask that. If something happens where shuffle plays one interval 3 times or doesn't play each one 2 times, manually adjust and make a note of it in participant notes.

APPENDIX D

Directions for Interval Production Task

Directions:

1. Make sure you know which task they are supposed to get first/second
2. Prepare your recording sheet
    1. Prepare the computer screen. Use full screen mode: https://www.timeanddate.com/stopwatch/
    2. Cover with sheet of paper - make sure the paper is not going to fall off between trials
3. Remember not to let the participant see how they did when you record
4. Read the script:

Script:

RA will give directions:

"I'm going to give you the laptop and read to you a time. I would like you to reproduce the interval on the computer for me by pressing the start/stop button.

Show where those buttons are on the screen

There will be a sheet of paper that covers up the clock. You are not allowed to move the paper at any time. After each trial, I will turn the computer, record, clear, and return it to you. I cannot tell you how well you did or show you any results until after you complete the entire study (This includes both time frames). Now let's practice. I would like you to reproduce 4 seconds on the clock"

Complete the trial, if the participant understands move on. If not, complete trial 2 (use 8 seconds). Do not tell the participant how they did.

"Very good, now we will start the actual test. Remember I cannot answer any questions or repeat any trials. Do you have any questions before we begin? Ready? Begin."

Make sure you take the computer and you write down the entire number each time. Clear the stopwatch and turn it back. Make sure the paper doesn't fall off. You cannot tell the participant how they did.

When you are nearing the end, you can say things like "Only 3 more trials left or last one."

Notes: Do not say anything like: 'I would like you to reproduce this interval **again."**

1. We don't want the participant to know that some intervals are twice. If they pick up on it, that's fine.

b. If they ask you questions remind the participant that you cannot answer questions.

APPENDIX E

Directions for Intervention II: Metronome

1. While the participant is doing MMN, queue up the metronome site in the behavioral room. http://a.bestmetronome.com/
2. Set the volume at 70db.
3. Set the metronome to 60 bpm.
4. Read from script:

Script:

You are going to listen to a metronome for 2 minutes. During this time, I would like you to alternate fingers and finger tap along with the metronome.

RA unplugs headphones, plays metronome, and demonstrates what to do

You are going to use this training later for a task so make sure you pay attention. Ready? Begin.

1.5 minutes

Now we are going to take a 1 minute break. Starting now

Time 1 minute

We are going to do the training one more time. Are you ready?

1 minute

5. During this time, check to make sure the participant is getting either interval production or verbal estimation first. And prepare that specific task. REMEMBER THESE TASKS AND THE INTERVALS ARE RANDOMIZED.

6. **MAKE SURE THE PARTICIPANT USES THE INTERVENTION (FINGER TAPPING) IN THE VERBAL ESTIMATION AND INTERVAL PRODUCTION TASKS**.

Script for Verbal Estimation:

RA will read the directions:
"I am going to play tones out loud. There are two tones for each trial. I want you to listen carefully and then tell me how much time you believe has elapsed between the two tones.
Try to be as accurate as possible and use the finger tapping method you just learned.
I cannot answer any questions or repeat the process, so make sure you pay attention.
Make sure that you use the finger tapping you just learned during these trials.
Let's try one."

- RA completes first trial of verbal estimation with the participant (4 seconds).
- If the participant understands, you do not need to do the second trial (8 seconds).
- If they want another, give the participant a second trial.
- Remember you cannot tell the participant how well they did or if they got it right.

"Very good, now we are going to begin the experiment. I cannot tell you how well you did or show you any results until after you complete the entire study (both time frames).
Remember that I also cannot answer any questions and I cannot repeat any trials.
Do you have any questions before we begin?
Remember to finger tap for every trial.
I will say 'okay, ready? (or whatever you want, just okay or ready will work)*' between each trial.*
Ready? Begin."

- RA has computer press *start* on first play list. Pause the playlist so it doesn't go to the next one.

RA asks participant "How much time has passed?"

- Record the interval that you played and the answer the participant gave.
- Press play on the next interval on the list. *Repeat*
- When you are nearing the end, you can say things like "Only 3 more trials left or last one."

Notes: Make sure you hit play/pause between EVERY TRIAL. Make sure that you are on shuffle and you write down each interval as it appears and write down the exact response from the participant. They are allowed to give you decimals, if they ask that. If something happens where shuffle plays one interval 3 times or doesn't play each one 2 times, manually adjust and make a note of it in participant notes.

Script for Interval Production:

RA will give directions:

"I'm going to give you the laptop and read to you a time.

I would like you to reproduce the interval on the computer for me by pressing the start/stop button ***with one hand, while finger tapping with the other.***

- Show where those buttons are on the screen

"There will be a sheet of paper that covers up the clock.

You are not allowed to move the paper at any time.

After each trial, I will turn the computer, record, clear, and return it to you.

I cannot tell you how well you did or show you any results until after you complete the entire study (This includes both time frames).

Now let's practice. I would like you to reproduce 4 seconds on the clock"

- Complete the trial, if the participant understands move on.
- **You can move the computer to make sure the participant is comfortable.**
- **Use the trials to allow the participant to get used to finger tapping and stopping the clock.**

"Very good, now we will start the actual test.

Remember I cannot answer any questions or repeat any trials ***and I want you to finger tap for every trial****.*

Do you have any questions before we begin?

Ready? Begin."

- Make sure you take the computer and you write down the entire number each time.
- Clear the stopwatch and turn it back.
- Make sure the paper doesn't fall off.
- You cannot tell the participant how they did.
- When you are nearing the end, you can say things like "Only 3 more trials left or last one."

Notes: Do not say anything like: 'I would like you to reproduce this interval **again."**

1. We don't want the participant to know that the intervals are twice. If they pick up on it, that's fine.
2. If they ask you questions remind the participant that you cannot answer questions.

APPENDIX F

The Letter Number Sequencing Task

- Read each item sequence at the rate of one character per second. You may use your stopwatch to ensure a steady pace.
- Be sure to clearly enunciate each letter and number so that the participant hears you correctly. Items may not be repeated.
- Write the participant's exact response on the line below each item. The correct answers are in parentheses below the line.
- If the participant gives the wrong sequence, circle the item. If the participant gives the correct sequence, do nothing more than record the response on the line below the item.
- A participant may correct herself if the administrator has not yet moved on to the next sequence. If so, write "self-correct" next to the circled item. Errors are only those mistakes that the participant does not correct spontaneously.
- If you accidentally administer the items in the wrong order, make a note of the incorrect sequence you gave and continue the test using the sequence that you gave.

Say to the participant, ***"I am going to say a list of numbers and letters. When I am through, I want you to first tell me the numbers in order from smallest to biggest. Then I want you to tell me the letters in alphabetical order. So, for example, if I say A4, the answer is 4A. The number goes first, then the letter. If I say 8B2, you answer 28B, numbers first in order, then letters. Try these."***

If the participant gets any item wrong, correct her and say the following, ***"Remember to first tell me the numbers in order from smallest to biggest and then the letters in alphabetical order."***

8. Letter–Number Sequencing

DISCONTINUE RULE: After scores of 0 for all three trials of an item.

RECORDING: All responses verbatim

SCORING RULE: 0–1 pt. for each trial

Item/Trial		(Correct Response)/Response	Score 0 or 1
1.	Trial 1	L – 2 (2 – L)	
	Trial 2	6 – P (6 – P)	
	Trial 3	B – 5 (5 – B)	
2.	Trial 1	F – 7 – L (7 – F – L)	
	Trial 2	R – 4 – D (4 – D – R)	
	Trial 3	H – 1 – 8 (1 – 8 – H)	
3.	Trial 1	T – 9 – A – 3 (3 – 9 – A – T)	
	Trial 2	V – 1 – J – 5 (1 –5 – J – V)	
	Trial 3	7 – N – 4 – L (4 – 7 – L – N)	
4.	Trial 1	8 – D – 6 – G – 1 (1 – 6 – 8 – D – G)	
	Trial 2	K – 2 – C – 7 – S (2 – 7 – C – K – S)	
	Trial 3	5 – P – 3 – Y – 9 (3 – 5 – 9 – P – Y)	
5.	Trial 1	M – 4 – E – 7 – Q – 2 (2 – 4 – 7 – E – M – Q)	
	Trial 2	W – 8 – H – 5 – F – 3 (3 – 5 – 8 – F – H – W)	
	Trial 3	6 – G – 9 – A – 2 – S (2 – 6 – 9 – A – G – S)	
6.	Trial 1	R – 3 – B – 4 – Z – 1 – C (1 – 3 – 4 – B – C – R – Z)	
	Trial 2	5 – T – 9 – J – 2 – X – 7 (2 – 5 – 7 – 9 – J – T – X)	
	Trial 3	E – 1 – H – 8 – R – 4 – D (1 – 4 – 8 – D – E – H – R)	
7.	Trial 1	5 – H – 9 – S – 2 – N – 6 – A (2 – 5 – 6 – 9 – A – H – N – S)	
	Trial 2	D – 1 – R – 9 – B – 4 – K – 3 (1 – 3 – 4 – 9 – B – D – K – R)	
	Trial 3	7 – M – 2 – T – 6 – F – 1 – Z (1 – 2 – 6 – 7 – F – M – T – Z)	

Total Score Range = 0 to 21

APPENDIX G

Time Management Questionnaire

Adopted from Britton and Tesser (1991)

(1 = never; 2 = infrequently; 3 = sometimes; 4 = frequently; 5 = always)

1. Do you make a list of the things you have to do each day?
2. Do you plan your day before you start it?
3. Do you make a schedule of the activities you have to do on work days?
4. Do you write a set of goals for yourself for each day?
5. Do you spend time each day planning?
6. Do you have a clear idea of what you want to accomplish during the next week?
7. Do you set and honor priorities?
8. Do you often find yourself doing things which interfere with your school-work simply because you hate to say "No" to people?
9. Do you feel you are in charge of your own time, by and large?
10. On an average class day do you spend more time with personal grooming than doing schoolwork?
11. Do you believe that there is room for improvement in the way you manage your time?
12. Do you make constructive use of your time?
13. Do you continue unprofitable routines or activities?

APPENDIX H

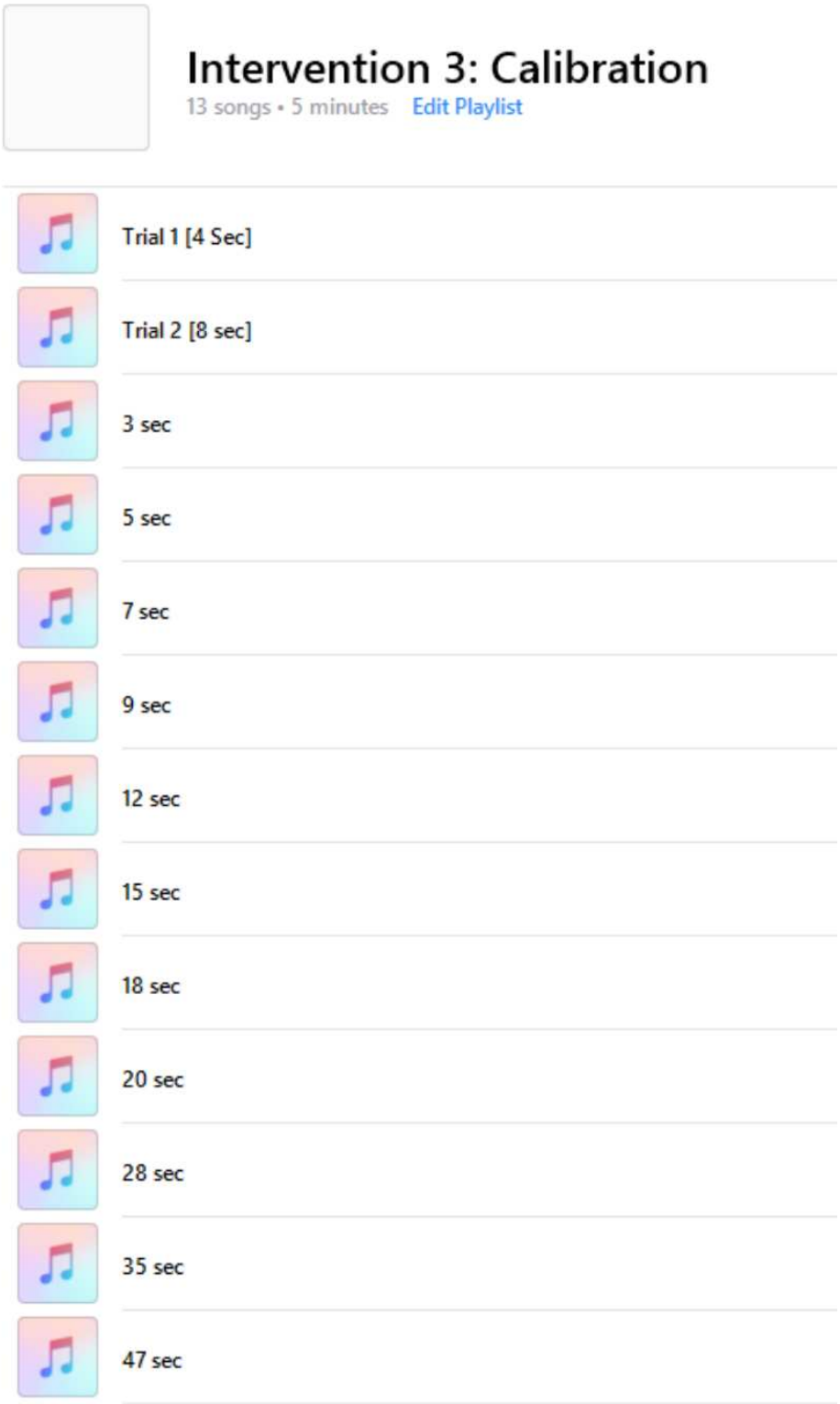
Time Structure Questionnaire
Adopted from Bond & Feather (1988)

	Yes, Always						**No, Never**
1. Do you ever have trouble organizing the things you have to	1	2	3	4	5	6	7
2. Do you have a daily routine which you follow?	1	2	3	4	5	6	7
3. Do you often feel that your life is aimless, with no definite	1	2	3	4	5	6	7
4. Many of us tend to daydream about the future. Do you find	1	2	3	4	5	6	7
5. Once you've started an activity do you persist at it until you've completed it?	1	2	3	4	5	6	7
6. Do you ever feel that the things you have to do during	1	2	3	4	5	6	7
7. Do you plan your activities from day to day?	1	2	3	4	5	6	7
8. Do you find that during the day you are often not sure what to do next?	1	2	3	4	5	6	7
9. Do you take a long time to "get going"?	1	2	3	4	5	6	7
10. Do you tend to change rather aimlessly from one activity to	1	2	3	4	5	6	7
11. Do you give up easily once you've started something?	1	2	3	4	5	6	7
12. Do you plan your activities so that they fall into a	1	2	3	4	5	6	7
	Would have no idea					**Yes,**	
13. Can you tell how many useful hours you put in last week?	1	2	3	4	5	6	7
	Yes, always						**No, never**
14. Do you get bored with your day-to- day activities?	1	2	3	4	5	6	7
15. Looking at a typical day in your life, do you think that most things you do have some purpose?	1	2	3	4	5	6	7
	No structure at all					**Very**	

16. Do your main activities during the day fit together in a	1	2	3	4	5	6	7
	Yes, Always						**No,**
17. Do you have any difficulty in finishing activities once you have started them?	1	2	3	4	5	6	7
18. Do you spend time thinking about opportunities that you have	1	2	3	4	5	6	7
19. Do you ever feel that the way you fill your time has little use or	1	2	3	4	5	6	7
20. Do you spend time thinking about what your future might be	1	2	3	4	5	6	7

APPENDIX I

Feedback Interval List



APPENDIX J

Directions for Intervention III: Calibration Feedback

1. Go through the verbal estimation and interval production trials in the order they are specified in 'procedure' tab of the schedule
2. Get the worksheet for the feedback intervals.
3. Look at the 'procedure' tab for which **Feedback List** the participant is getting
    a. One will be for verbal estimation and one will be for interval production, these are randomized across participants
        i. For example, one participant will get list 1 for verbal estimation and then list 2 for interval production.
        ii. No participants will get 1 list for both ***during the intervention.***
        iii. **These are different lists from the lists used during the typical behavioral (verbal estimation/interval production) tasks.**
4. Read the script for the task (next page)
5. Interval production: Train the participant on each interval 3 times, then move on to the next one.
    a. Example for Interval production task:
        i. The interval is 8 seconds, they reproduce 9 seconds
        ii. You say: you overestimated by 1 second.
            1. Show the participant how they did on the screen for interval production
            2. Have the participant redo the trial if they are +/- .5 seconds off twice, then move on
            3. For the longer intervals 20+ do not repeat trials
6. Verbal estimation: Ask for response, then tell the participant how much they over/underestimated by
    a. Example for Verbal estimation task:
        i. The interval is 8 seconds, the participant says 7 seconds
        ii. You say: You overestimated by 1 second.

iii. Do not repeat verbal estimation trials (they could just lie).

7. Once they have completed the intervention tasks, you will move on to the behavioral tasks
8. Check to see which one they get first (should be the opposite of whatever they got in time 1)
9. Prep your recording list and check to make sure the list is different from what they got in time
10. **Do not give the participant feedback once you have moved on to the behavioral tasks**

Scripts for Verbal Estimation and Interval Production Tasks in Intervention II (Feedback):

Script for Verbal Estimation:

"I am going to play tones out loud.
There are two tones for each trial.
I want you to listen carefully and then tell me how much time you believe has elapsed between the two tones, as accurately as possible.
I cannot answer any questions or repeat the process, so make sure you pay attention.
Let's try one."

- RA completes first trial of verbal estimation with the participant (4 seconds).
- If the participant understands, you do not need to do the second trial (8 seconds).
- If they want another, give the participant a second trial.
- Remember you cannot tell the participant how well they did or if they got it right.

"Very good, now we are going to begin the experiment.
After each trial, I'm going to tell you how close you were to the actual interval.
Do you have any questions before we begin? I will say 'okay, ready? (or whatever you want, just okay or ready will work)*' between each trial.*
Ready? Begin."

- RA has computer press *start* on first play list. Pause the playlist so it doesn't go to the next one.

RA asks participant "How much time has passed?"

Example response:

If **overestimates** (i.e. interval was 3, participant responds 5): *you overestimated by 2 seconds.*

If **underestimates** (i.e. interval was 3, participant responds 1): *you underestimated by 2 seconds.*

If **correct** (i.e. interval was 3, participant responds 3): *correct.*

Script for Interval Production:

"I'm going to give you the laptop and read to you a time. I would like you to reproduce the interval on the computer for me by pressing the start/stop button.

- Show where those buttons are on the screen

"There will be a sheet of paper that covers up the clock.

You are not allowed to move the paper at any time.

After each interval, I will tell you how well you did.

Now let's practice. I would like you to reproduce 4 seconds on the clock"

- Complete the trial, if the participant understands move on.
- If not, complete trial 2 (use 8 seconds).
- **Tell the participant how they did and tell the participant.**

"Very good, now we will start the actual test. Remember I cannot answer any questions or repeat any trials. Do you have any questions before we begin? Ready? Begin."

- If **overestimates** (i.e. interval was 3, participant responds 5): *you overestimated by 2 seconds.*

- If **underestimates** (i.e. interval was 3, participant responds 1): *you underestimated by 2 seconds.*
- If **correct** (i.e. interval was 3, participant responds 3): *correct.*

Other Notes:

When you then move on to the next task (either verbal estimation or interval production), **do not give the participant feedback anymore. You can tell the participant that they are not going to receive feedback.**