Unlocking Desistance: The Role of Executive Functioning in the Rehabilitation of Correctional Populations

Danielle Nesi
LOYOLA UNIVERSITY CHICAGO

UNLOCKING DESISTANCE:
THE ROLE OF EXECUTIVE FUNCTIONING IN
THE REHABILITATION OF CORRECTIONAL POPULATIONS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN DEVELOPMENTAL PSYCHOLOGY

BY
DANIELLE J. NESI
CHICAGO, IL
MAY 2022
ACKNOWLEDGEMENTS

I would like to thank all of the people who made this dissertation possible, starting with my faculty mentors in the Psychology Department at Loyola University Chicago. Since our first meeting in 2014, Dr. James Garbarino has been my champion and a constant source of support and encouragement, always looking out for my best interests both professionally and personally. Since Dr. Garbarino’s retirement in 2020, Dr. Arthur Lurigio graciously stepped in as my primary research mentor, helping guide me through my dissertation and countless other academic endeavors to foster my professional development. Dr. Lurigio’s passion and drive is a constant inspiration and his knowledge of clinical psychology, forensic psychology and criminal justice and research in these fields has provided me with much perspective that has enriched my knowledge. It truly would not have been possible for me to complete this project without his guidance.

I also would like to thank Dr. Fred Bryant for his significant contributions in guiding me through the complex statistical analyses needed to complete this project. I consider myself fortunate to have taken two classes taught by Dr. Bryant during my time at Loyola. His optimism and passion for statistics and academic research is a gift to all students lucky enough to work with him. Although not a formal member of my dissertation committee, Dr. Bryant spent countless hours guiding me through complex statistical procedures that surpassed the lessons covered in his Advanced Statistics and Structural Equation Modeling courses. Dr. Bryant thoughtfully guided me through these procedures, making sure that my analyses were completed.
properly and that I truly understood every decision made and every result obtained. His mentorship heightened my ability to conduct statistical analyses beyond my expectations.

I would be remiss not to thank my dissertation committee members: Dr. David Olson, Dr. Noni Gaylord-Harden, and Dr. Troy Harden. Throughout the many iterations of this project, my committee has been steadfast and encouraging, providing their expertise to address the many unforeseen challenges that arose during the completion of this project. Thank you to all of my committee members who reviewed and critiqued this work. Your thoughtful insights are sincerely appreciated.

I must also thank my family and friends who have supported me throughout this entire process. Thank you for your patience, understanding, and encouragement. In particular, I want to thank my mom, who has endured countless phone calls regarding my dissertation. Not only did you allow me to vent (usually), but you really listened and provided me with advice that was truly helpful. Thank you, mom, for spending hours listening to me practice my defense, always participating as an engaged and thoughtful listener. I am lucky to have you as my mom.

Finally, I would also like to thank my dear friend, Kristin Hamilton. Thank you, Kristin for encouraging me to take that criminal justice elective at University of Wisconsin and for modeling what it means to be a good student and an even better person. I would also like to extend my appreciation to the entire Hamilton-Geil family for treating me as one of their own, especially since I moved to Chicago to begin my doctorate at Loyola. During the trials that have accompanied the completion of this degree, it was a blessing to be adopted by your family.
To my mom, dad, and grandma Frances for their endless love and support.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** iii

**LIST OF TABLES** viii

**LIST OF FIGURES** x

## CHAPTER ONE: INTRODUCTION
- Current Research in a Thumbnail 2
- Executive Functions 2
- Executive Functioning & Antisocial Behavior 6
- Executive Functioning & Recidivism 7
- Executive Functioning and Substance Use & Other Psychiatric Disorders 8
- Diminished Executive Capacity Due to Correctional Institutionalization 12
- Plasticity of Executive Functioning and Potential for ‘Critical Age’ for Remediation 12
- Executive Function Remediation Among Antisocial Populations 16

## CHAPTER TWO: THE CURRENT STUDY
- Purpose of the Current Study 18
- Methods of the Current Study 19
- Rationale for Reframing Fishbein and Sheppard (2006) 22
- Contribution of this Reanalysis 25
- Current Study Hypotheses 26

## CHAPTER THREE: METHODS
- Study Design 28
- Participants/Participant Characteristics 29
- Measures and Covariates 31
- Data Processing and Diagnostics 65

## CHAPTER FOUR: RESULTS
- Hypothesis 1: Baseline Executive Functioning and Recidivism and Hypothesis 1a: Mediating Effects of Treatment on Recidivism 74
- Path 1a: Baseline Executive Functioning, Recidivism and Treatment Group Mediation 75
- Path 1b: Baseline Executive Functioning Predicting Recidivism Mediated by Treatment Gain 85
- Hypothesis 2: Does Age Predict Change in Executive Functioning? Does Change in Executive Functioning Vary by Age Group? 91
- Hypothesis 3: Does Engagement in Treatment Predict Change in Executive Functioning? 100
- Hypothesis 4: Does Change in Executive Functioning Predict Recidivism? 105
LIST OF TABLES

Table 1. Means, Standard Deviations, and Ranges for Demographic Variables Assessed at Baseline 30

Table 2. Baseline Test Battery and Measures 64
Table 3. Follow-Up Test Battery and Measures 64
Table 4. Correlation Coefficient Matrix for Path 1a Variables 76
Table 5. Pearson Correlations, Means, and Standard Deviations for Path 1a Variables 77
Table 6. Direct Effects in Model 1a (LISREL parameter estimates) 83
Table 7. Indirect Effects in Model 1a (LISREL parameter estimates) 85
Table 8. Pearson Correlations, Means, and Standard Deviations for Path 1b Variables 86
Table 9. Direct Effects in Model 1b (LISREL parameter estimates) 90
Table 10. Indirect Effects in Model 1b (LISREL parameter estimates) 91
Table 11. Pearson Correlations, Means, and Standard Deviations for Hypothesis 2 Variables 93
Table 12. LISREL Parameter Estimates for Hypothesis 2: Does Age Predict Change in Executive Functioning (Model 2) 93
Table 13. Tests for Normality Among Hypothesis 2 Variables (Shapiro-Wilk) 95
Table 14. Median Scores for Change in Reaction Time by Age Group 96
Table 15. Median Scores for Change in Percentage of Safe Choices by Age Group 97
Table 16. Tests for Normality Among Baseline Executive Functioning Across Age Group (Shapiro-Wilk) 99
Table 17. Mean Scores in Baseline Executive Abilities by Age Group 100
Table 18. Pearson Correlations, Means and Standard Deviations for Hypothesis 3 Variables 103

Table 19. LISREL Parameter Estimates for Hypothesis 3: Do Treatment Gains Predict Change in Executive Functioning? 105

Table 20. Pearson Correlations, Means, and Standard Deviations for Model 4 Variables 107

Table 21. LISREL Parameter Estimates for Regression Coefficients – Predicting Change in Number of Infractions 109

Table 22. Pearson Correlations, Means, and Standard Deviations for Model 5 Variables 109

Table 23. LISREL Parameter Estimates for Regression Coefficients – Predicting Change in NAS Score 110

Table 24. Pearson Correlations, Means, and Standard Deviations for Model 6 Variables 110

Table 25. LISREL Parameter Estimates for Regression Coefficients – Predicting Change in PI Score 111
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Path Diagram of the Conceptual Model Tested in the Present Research (Model 1a)</td>
<td>82</td>
</tr>
<tr>
<td>2</td>
<td>Path Diagram of the Conceptual Model Tested in the Present Research (Model 1b)</td>
<td>87</td>
</tr>
<tr>
<td>3</td>
<td>Path Diagram of the Conceptual Model Tested in the Present Research (Hypothesis 2/Model 2)</td>
<td>92</td>
</tr>
<tr>
<td>4</td>
<td>Path Diagram of the Conceptual Model Tested in the Present Research (Model 4)</td>
<td>106</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

The overall goal of this research is to examine the relation between executive functioning and recidivism among individuals engaged in antisocial behavior. Previous work has demonstrated that executive function deficits are pervasive among antisocial populations (Morgan & Lilienfeld, 2000), leading many researchers to conclude that such deficits are important risk factors for, or correlates of, antisocial behavior (Elliot, 1978; Gorenstein, 1982; Raine, 1997). Executive functioning can be understood as “an umbrella term that refers to the cognitive processes that allow for future, goal-oriented behavior” (Morgan & Lilienfeld, 2000, p. 114). Broadly, executive functions refer to the management of cognitive processes that guide complex behavior (Banich, 2009; Miller & Cohen, 2001) and executive functioning is critical for activating, maintaining, and selecting different courses of action necessary to carry out the complex behaviors required to achieve various goals (Miyake & Friedman, 2012).

It has been suggested that, compared to the general population, individuals with executive function deficits are less able to override inclinations towards maladaptive responses required to maintain more appropriate and personally beneficial behavior (Zeier, Baskin-Sommers, Hiatt Racer, & Newman, 2011). Moreover, it is believed that executive function deficits may contribute to deviant behavior by decreasing impulse inhibition, sensitivity to reward and punishment, and the ability to plan and formulate behaviors that correspond with social demands (Ishikawa & Raine, 2003; Raine, 2002; Séguin, 2004). Consequently, researchers have proposed
that cognitive abilities, particularly executive functions, are of critical importance in regards to outcomes for antisocial individuals (Rocha, Marques, Fortuna, Antunes, & Hoaken, 2014).

**Current Research in a Thumbnail**

The current study is an exploration of the contributions and interrelations of age and executive function in treatment-related outcomes and continued antisocial behavior as indicated by future institutional infractions, a proxy for recidivism, among a correctional population. Specifically, the current study sought to identify whether baseline executive functions predict future antisocial behavior and the extent to which such relationships could be explained by individuals’ capacity to benefit from and engage in an intervention program designed for this population. In addition, the current study sought to explore the importance of age in relation to cognitive remediation among a correctional population. Specifically, following prior research indicating cognitive remediation may be more effective among younger adults, the current study sought to explore whether younger inmates were more amendable to cognitive remediation. Potential findings may elucidate important relations between executive functioning and recidivism that may be used to inform correctional treatment programs in order to best address critical risk factors among this population.

**Executive Functions**

As mentioned previously, executive functions encompass a wide array of cognitive functions, such as planning, inhibition, switching and monitoring (Diamond, 2013). These abilities, which are also sometimes referred to as executive control or cognitive control (Diamond, 2013), refer to a collection of top-down mental processes that are needed when one has to concentrate and pay attention, and when going on autopilot or relying on instinct or intuition would be ill-advised, insufficient, or impossible (Burgess & Simons, 2005; Espy, 2004;
Miller & Cohen, 2001). As explained by Diamond (2013), the use of executive functions is indeed effortful; it is easier for one to continue engaging in routinized behaviors or habits than it is to change; it is easier to give into temptation than to resist it, and it is easier to go on “automatic pilot” than it is to consider a future course of action (Diamond, 2013).

Although there seems to be a clear understanding of what executive abilities are, at least generally speaking, and why they are significant for individuals’ well-being and overall level of functioning, the conceptualization of executive functions and their structural organization has been contested by researchers over the past several decades. However, there is general agreement on the existence of three core executive functions (e.g., Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000): inhibition, working memory, and cognitive flexibility (also called set-shifting, mental flexibility, or mental set shifting).

Inhibition includes inhibitory control, such as self-control (e.g., behavioral inhibition) and interference control (e.g., selective attention and cognitive inhibition). This core executive function involves being able to control one’s attention, behavior, thoughts and/or emotions in order to override a strong internal predispositions or external incentives, and instead do what is more appropriate or necessary (Diamond, 2013). As Diamond (2013) explains, without these abilities, we would be at the mercy of impulses, long-standing habits of thought or actions (e.g., conditioned responses), and/or influential environmental stimuli that pull us in a certain direction. Thus, inhibitory control enables us to change and allows us to choose how we react and behave rather than being unthinking creatures of habits. In regards to attention, inhibitory control (e.g., interference control), or attentional inhibition, enables individuals to selectively attend, focusing on what they choose and suppressing attention to other stimuli (Diamond, 2013). In other words, attentional control allows individuals to focus their attention only to stimuli
relevant to current goals and also to ignore distractive stimuli that are task-irrelevant.

Another core executive function, working memory, involves holding information in mind and mentally working with it, or in other words, working with information that is no longer perceptually present (Baddeley & Hitch, 1994; Smith & Jonides, 1999). Thus, working memory refers to the small amount of information that can be held in the mind and subsequently used in the execution of cognitive tasks (Cowan, 2014). There are two types of working memory that are distinguished by content-verbal and nonverbal (e.g., visual-spatial) working memory (Diamond, 2013). Working memory enables individuals to do mental tasks such as doing math in their head, mentally re-ordering items (for example, reorganizing a to-do list), and translating instructions into action plans (Diamond, 2013). Moreover, working memory is critical for reasoning and for creativity, as it allows individuals to see connections between seemingly unrelated things and to separate elements from an integrated whole (Diamond, 2013).

Lastly, the third core executive function, cognitive flexibility, also referred to as shifting, refers to the ability to switch between different mental sets, tasks, or strategies. Thus, cognitive flexibility has to do with one’s capacity for flexible thinking. Cognitive flexibility allows individuals to change perspectives spatially, for example, “What would this look like if I were to view it from another direction,” and interpersonally (e.g., “Let me see if I can see this from another person’s point of view”) (Diamond, 2013). Moreover, cognitive flexibility enables us to change how we think about something, for example, when a certain method of problem solving is ineffective, is someone able to come up with a new way to address the issue, or to come up with a solution that had not been considered previously. Cognitive flexibility also enables individuals to adjust to changing demands or priorities, to admit when they are wrong, and to take advantage of sudden, unexpected opportunities; for example, when an individual may have
been planning to do something (X), but an opportunity arises (Y), cognitive flexibility enables them to take advantage of such an opportunity. From these core executive functions, higher order executive functions are formed such as reasoning, problem solving, and planning (Collins & Koechlin, 2012; Lunt et al., 2012).

Regarding the structural organization of executive functions, there have been different theories proposed. One of the most well-known is the unity and diversity theory which was first proposed by Teuber (1972) and is now commonly associated with the work of Miyake and Friedman (see Miyake et al., 2000; Friedman et al., 2008; Miyake & Friedman, 2012). This structural conceptualization is based on studies that have found that individual components of executive function correlate with one another, which has been used as evidence to suggest there is some common underlying ability (e.g., unity), and that there is also some separation of these abilities (e.g., diversity). Thus, unity-diversity theories espouse that there are some underlying common elements and also specific elements of executive functions. This general unity-diversity conceptualization has been shared by several different models of executive functioning (e.g., Duncan, Johnson, Swales, & Freer, 1997; Repovs & Baddeley, & 2006), although such models have focused on different executive components and different levels of analysis (e.g., behavior vs. neural).

In considering the structure of executive functions, another important distinction lies in the distinction between “hot” and “cool” executive functions (Zelazo & Müller, 2002; see Peterson & Welsh, 2014, for a review). This distinction stems from evidence based on human and nonhuman animal research that the top-down, neurocognitive processes (i.e., executive function abilities) needed in a given scenario vary depending on the presence of high levels of emotion. “Hot” and “cool” executive functions, which ordinarily work together to aid in solving
real-world problems, are both forms of deliberate, effortful, self-regulatory processing that depend on prefrontal regions in the brain; however, these abilities vary in the extent to which they require managing emotion and motivation and also in the specific neural circuits involved (Zelazo, Blair, & Willoughby, 2016). “Cool” executive functions are evoked under relatively abstract, non-affective situations, whereas “hot” executive functions are evoked in conditions that are motivationally significant and affective (Zelazo & Carlson, 2012; Zelazo & Müller, 2002). When individuals are confronted with an affective or personally meaningful problem that they are motivated to solve, the affective, “hot” aspects are most likely to be elicited. Thus, “hot” executive functions are elicited when people care about a problem they are attempting to solve (Zelazo, Qu, & Müller, 2005). In contrast, most laboratory-based, behavioral measures of executive functioning assess cool executive function using abstract, decontextualized tasks that lack a significant affective or motivation component for participants (Zelazo et al., 2016).

**Executive Functioning & Antisocial Behavior**

The relationship between antisocial behavior and executive functioning has gained widespread acceptance in recent years in part due to meta-analytic reviews such as that by Morgan and Lilienfeld (2000), which concluded the existence of a “robust and statistically significant relationship between antisocial behavior and executive functioning deficits” (p. 128). In this review, it was reported that antisocial individuals performed .62 standard deviations worse on tests of executive functioning than comparison groups. Subsequent research has strengthened the body of evidence demonstrating significant executive function deficits among antisocial individuals (see Ogilvie, Stewart, Chan, & Shum, 2011; Zeier et al., 2011).

Although the mechanisms for how executive function deficits influence antisocial behavior are not fully understood, discerning the function of executive abilities may elucidate
how deficits in executive functioning contribute to antisocial behavior. Models explaining the relationship between executive functioning deficits and antisocial behavior suggest that compared to the general population, these individuals are less able to override inclinations towards maladaptive responses required to maintain more appropriate and personally beneficial behavior (Zeier et al., 2011), and consequently, individuals with executive deficits are at a high risk for persistent rule breaking and violent behavior (Zeier et al., 2011).

**Executive Functioning & Recidivism**

Unsurprisingly, as executive functioning deficits are related to antisocial behaviors and greater risk of criminal conduct, offenders with executive deficits demonstrate higher rates of recidivism as compared to other offenders (Langevin & Curnoe, 2011; Ross & Hoaken, 2011). As previously discussed, executive functions encompass higher order cognitive abilities including planning, working memory, taking initiatives, set-shifting, attention, and impulse control (Jurado & Rosselli, 2007; Diamond 2013). According to Hofmann, Schmeichel and Baddeley (2012), these higher order cognitive functions are crucial for self-regulation. As explained by Meijers and colleagues (2015), during the process of reentering into society, offenders face numerous obstacles that place high demands on one’s executive abilities. For example, released offenders have to take initiatives and be able to engage in planning and reasoning in order to do things like secure housing and employment, the absence of which are both risk factors for reoffending (Luther, Reichert, Holloway, Roth, & Aalsma, 2011). Thus, it is possible that executive dysfunction may cause an increase in reoffending through failures in one’s ability to engage in self-regulation (Meijers et al., 2015).
Executive Functioning and Substance Use & Other Psychiatric Disorders

In addition to being related to antisocial behaviors, deficits in executive functioning have also been found to be related to substance use disorder (Verdejo-Garcia, Bechara, Recknor, & Perez-Garcia, 2006) and risk for substance dependence (Giancola & Tarter, 1999). The relationship between executive functioning deficits and substance use is particularly noteworthy however, as individuals with executive deficits are more likely to use substances and there is evidence that this relationship is multidirectional.

Research has shown that youth with neurobehavioral disinhibition are at increased risk for later substance use disorder (Tarter et al., 2003) and previous research with young adults with parental history of substance use disorder found that abnormal tests of executive function were predictive of alcohol consumption at a 3-year follow up (Deckel & Hesselbrock, 1996). Another prospective study found that lower executive function ability was predictive of later alcohol-related problems and the number of illicit substances used, controlling for other factors including parent alcoholism and IQ (Nigg et al., 2006). Thus, it seems that executive function deficits are associated with future problems with substance use, which is consistent with the view that low levels of executive functioning serve as risk factors for substance use and dependence.

Although executive function deficits may predict problematic substance use, research has also shown that substance use can have an adverse effect on executive functioning. Reviews by Silveri (2012) and Jacobus and Tapert (2013) provided evidence suggesting the existence of alcohol-related neurobehavioral deficits. In particular, these studies suggested that alcohol use, particularly heavy episodic, or binge drinking, is associated with various cognitive deficits including reduced attention, poorer working memory, and impaired visual-spatial skills. Moreover, such drinking behaviors were found to be associated with structure and functional
changes in the brain, such as reductions in frontal, parietal, temporal and cerebellar volumes, and
in changes in white matter volume and integrity. Brain activation patterns and connectivity were
also found to be altered among individuals with these patterns of alcohol consumption. In
addition, other research has found that polysubstance users demonstrate cognitive deficits similar
to those observed among problem drinkers (Fernandez-Serrano, Perez-Garcia, & Verdejo-Garcia,
2011; Grant & Judd, 1976; Vik, Cellucci, Jarchow, & Hedt, 2004).

Concerns of the interactional relationship between drug and alcohol use and executive
functioning deficits are especially relevant when considering the behaviors of antisocial
individuals due to the high levels of substance use and abuse among this population. A review by
Fazel, Bains and Doll (2006) estimated the prevalence of drug use to be 10-48% among male
prisoners and 30-60% among female prisoners and the prevalence of alcohol abuse or
dependence to be 18-30% percent among male prisoners and 10-24% among female prisoners.
To put these numbers in context, the National Institute on Drug Abuse (NIDA) estimated that in
2013, 6.6% of Americans (17.3 million individuals) were dependent on alcohol, or had
experienced problems related to their alcohol use (alcohol abuse) (NIDA, 2015).

Moreover, other estimates of national drug abuse and dependence can be derived from
several ongoing, large-scale epidemiologic studies, such as the National Epidemiologic Survey
on Alcohol and Related Conditions (NESARC), which have assessed diagnostic categories of
drug abuse and dependence as defined in the DSM-IV (Schulden, Thomas, & Compton, 2009).
Results from the NESARC indicated that approximately 2% of adults living in US households
met DSM-IV criteria for having a drug use disorder in the past 12 months (1.4% met diagnostic
criteria for abuse, 0.6% met criteria for dependence) and 10.3% reported having a drug use
disorder at any point over the course of their lifetime (7.7% abuse, 2.6% dependence) (Compton,
Thomas, Stinson, & Grant, 2007). Thus, correctional populations have demonstrated far higher rates of substance use/abuse than have been observed among the general population. The prevalence of significant substance use and abuse among correctional populations may exacerbate the effects of executive functioning deficits that are common among this population.

Beyond substance use disorders, executive deficits are also pervasive in other psychiatric disorders. Previous research has shown that executive function deficits often accompany Bipolar Disorder. For example, Soraggi-Frez and colleagues (2017) found that working memory deficits are often observed among individuals with Bipolar disorder, even in the euthymic stage (e.g., when individuals are in normal mood state i.e. not in manic or depressive states). Other studies with elder Bipolar populations found comprehensive cognitive deficits, including impairments in: working memory, processing speed, inhibitory control, and cognitive flexibility (Caixeta et al. 2017, see also Cotrena, Branco, Shansis, & Fonseca 2016).

Executive deficits have also been seen among those with depressive disorders. For example, work by Cotrena and colleagues (2016) found individuals with Major Depressive Disorder (MDD) demonstrate poor attention and impairment on timed tasks, suggesting poor efficiency in executive processing. Executive function impairments are also some of the most commonly observed deficits among individuals with Schizophrenia (see Orellana & Slachevsky, 2013). In addition to these conditions, executive functioning deficits seem to be a hallmark of Attention-Deficit/Hyperactivity Disorder (AD/HD), with numerous studies reporting that significant executive function deficits are pervasive among this population (see Adler et al., 2017; Barkley, 1997; Barkley, Murphy & Fisher, 2008; Barkley & Murphy, 2010, 2011; Barkley & Fischer, 2011; Biederman, Faraone, Spencer, Mick, & Monuteaux, 2006; Brown, 2005, 2006).
As previously discussed, executive deficits are also prevalent among individuals with Antisocial Personality Disorder (ASPD), particularly in regards to response inhibition and poor, or risky, decision making (Fossati et al., 2004; Chamberlain, Derbyshire, Leppink, & Grant, 2016) and cognitive control (Zeier et al., 2011). Other studies have demonstrated those with ASPD have more pervasive executive impairment, such as deficits in visual short-term and working memory, inhibitory control, attention (Balouis, Duggan, McCarthy, Huband, & Völlm, 2019) and in planning, cognitive flexibility (set-shifting), sustained attention (Dolan & Park, 2002).

Notably, the specific executive impairments associated with ASPD may vary depending on whether one has comorbid psychopathy. Although a study of violent offenders with ASPD found that both those with and without comorbid psychopathy demonstrated similar impairments in verbal working memory and adaptive decision-making and did not differ in regards to levels of risk taking, impulsivity, etc. (De Brito, Viding, Kumari, Blackwood, & Hodgins, 2013), other research has found that offenders with ASPD and comorbid psychopathy do not present impairments in ‘cool EFs’ such as attentional set-shifting, planning and verbal working memory (see Blair et al., 2006; Hare, 1984; Hart, Forth, & Hare, 1990; Lapierre, Braun, & Hodgins, 1995; Mitchell, Colledge, Leonard, & Blair, 2002; Mitchell, Avny, & Blair, 2006; Smith, Arnett, & Newman, 1992) and instead present deficits in other ‘cool EFs’ such as response inhibition (Lapierre et al., 1995) and in ‘hot EFs’ such as affective decision-making, behavioral extinction (see Mitchell et al., 2002; Mitchell et al., 2006; Budhani, Richnell, & Blair, 2006; Newman, Patterson, & Kosson, 1987).
**Diminished Executive Capacity Due to Correctional Institutionalization**

Although neural plasticity can have positive implications as it suggests opportunities for intervention to alter negative trajectories and enhance cognitive abilities among individuals who display deficits, the effects of neural plasticity are not always positive. As cognitive remediation can improve cognitive abilities, scenarios in which there is a deprivation of cognitive stimuli can result in deterioration of abilities.

In regards to correctional populations in particular, work by Umbach, Raine and Leonard (2018) explored the effects of incarceration on cognitive functioning and found that the experience of incarceration resulted in significant decline in cognitive control and emotion regulation at a four-month follow up. In the same study, an experimental group who were provided with a cognitive behavioral therapy/mindfulness training (CBT/MT) intervention showed no significant decline in either outcome. These results are striking as they suggest that the experience of incarceration can serve to exacerbate deficits in cognitive abilities such as cognitive control and emotion regulation, which are known to be risk factors for antisocial behavior and recidivism. These findings suggest that the experience of incarceration may actually make individuals more likely to engage in antisocial behavior post-release. Moreover, this research indicates the negative effects of incarceration on cognitive abilities can manifest quickly, within a matter of months (Umbach et al., 2018). On a more optimistic note, this work also suggests that there is potential for CBT/MT therapies to protect against the cognitive declines that may typically accompany experiences of incarceration.

**Plasticity of Executive Functioning and Potential for ‘Critical Age’ for Remediation**

Evidence has shown that executive abilities are not fixed and can be improved using various remediation methods such as cognitive training programs specifically designed to
improve executive functioning (see Au et al., 2015; Klingberg, 2010; Kueider, Parisi, Gross, & Rebok, 2012; Spierer, Chavan, & Manuel, 2013; von Bastian & Oberauer, 2013). Other studies have also shown that physical activity programs can also improve EF (see Angevaren, Aufdemkampe, Verhaar, Aleman, & Vanhees, 2008; Audiffren & André, 2015; Barenberg, Berse, & Dutke, 2011; Etnier, Nowell, Landers, & Sibley, 2006; Hindin & Zelinski, 2012; Kramer & Erickson, 2007; Moreau & Conway, 2013; Streiner, 2009); however, such programs have been shown to produce little improvement when they do not include a cognitive component (Angevaren et al., 2008; Smith et al., 2010).

Although studies have demonstrated programs designed to improve executive functions can work across the lifespan, there is a paucity of research on whether such programs are more beneficial when administered to particular age groups. Research has demonstrated that among individuals with schizophrenia, cognitive remediation produces greater gains in younger individuals (see Wykes et al., 2009; Bowie, Grossman, Gupta, Oyewumi, & Harvey, 2014; Corbera, Wexler, Poltorak, Thime, & Kurtz, 2017; Golas et al., 2015; Kontis, Huddy, Reeder, Landau, & Wykes, 2013; Thomas, Puig, & Twamley, 2017). It should be noted, however, the effects of age in such studies are confounded by factors such as the duration of schizophrenic symptoms or time elapsed since the onset of the disorder.

Sensitive periods (SPs) are points in development during which there is heightened neural sensitivity to specific environmental stimuli. During such periods, exposure to specific external stimuli is necessary for typical developmental processes to occur (Thompson & Steinbeis, 2020). After a sensitive period has ended, individuals still have the ability to learn new skills in a given domain. However, the potential for new skill acquisition in that domain is constrained by the experiences that took place during the SP. This is in contrast to ‘critical periods.’ Once a critical
period has closed, further development in the relevant domain is not possible (Nelson, Zeanah, & Fox, 2019).

In relation to executive functioning, it has been suggested that the ‘shared’ and ‘distinct components’ of executive functioning may be associated with separate sensitive periods (Thompson & Steinbeis, 2020). There is evidence that the first few months of life may represent one sensitive period in executive functioning. As noted by Thompson and Steinbeis (2020), the development of sensorimotor systems in the first few months of life must occur in order to enable an infant to receive accurate visual input. It is only if this skill is sufficiently developed, that infants are able to begin regulating their attention to visual inputs (Rosen, Amso, & McLaughlin, 2019; Amso & Scerif, 2015). This ability to regulate attention to visual inputs is considered a developmental prerequisite of future executive functioning (Blankenship et al., 2019; Cuevas & Bell, 2014; Rose, Feldman, & Jankowski, 2012).

Moreover, infants’ attention is influenced by interactions with caregivers in particular as parents provide scaffolding (Bernier, Carlson, & Whipple, 2010), whereby a parent directs an infant’s attention to particular items in his or her environment (Bibok, Carpendale, & Müller, 2009). As suggested by Thompson and Steinbeis (2020), this sort of interaction is likely influenced by factors such as caregiver attachment and parental attentiveness. Sensitive and stimulus contingent caregiving is needed for the development of parent-child attachment (Beijersbergen, Juffer, Bakermans-Kranenburg, & van IJzendoorn, 2012). Moreover, this type of caregiving serves to provide infants with greater perceptions of instrumental control over their environment, which is strongly related to individual differences in executive functioning (Henderson, Snyder, Gupta, & Banich, 2012; Sanchís-Ollé et al., 2019).
As early experiences in the first few months of life have a profound impact on future development of executive functioning, Thompson and Steinbeis (2002) have suggested this time constitutes a sensitive period in executive function development. They note, however, that executive functioning encompasses a collection of different cognitive abilities that are served by partially overlapping brain regions and that have subtly distinct developmental trajectories. Consequently, they suggest that there may be several sensitive periods in the formation and maturation of executive functions.

From a developmental perspective, the fact that higher order (e.g., planning, decision-making, etc.) and “hot” executive functions (e.g., planning, decision-making, impulse control in emotionally laden situations or “real-world” scenarios) continue to develop into early adulthood suggests that this time may be another sensitive period in executive function development. During the period of early adulthood, individuals are becoming increasingly independent and gaining autonomy (Settersten, Ottusch, & Schneider, 2015). The increasing autonomy granted to young adults provides them with more opportunities wherein they are required to implement executive abilities. These opportunities can be viewed as “inputs” needed for executive functions to mature. From this view, it is possible that opportunities to practice using executive abilities or interventions designed to improve such skills may be more effective when completed during early adulthood.

This line of thought is supported by research that demonstrates that one of the fundamental features of effective executive remediation programs and is that executive abilities are continually challenged (Diamond & Ling, 2016). To see improvement, it is not enough for an individual to practice using such skills; they need to be pushed beyond their current level of functioning (Diamond & Ling, 2016). This is essentially what happens in the natural world as the
increasing autonomy provided to young and emerging adults provides them with greater tests of executive abilities as they navigate the world as independent adults.

As previously discussed, studies of cognitive remediation in other populations have demonstrated that younger adults show greater improvement in executive functioning as compared to older adults (see Wykes et al., 2009; Bowie et al., 2014; Corbera et al., 2017; Golas et al., 2015; Kontis et al., 2013; Thomas et al., 2017). Such findings provide additional evidence for early-adulthood as a potential sensitive period in the development of executive functioning.

**Executive Function Remediation Among Antisocial Populations**

As noted, a large body of research has demonstrated that criminality, antisocial behavior, aggression, and violence are associated with neurocognitive deficits and reduced functional brain capacity. In particular, the presence of impairments in executive functions is thought to exert a considerable effect in the development of antisocial behavior by decreasing impulse inhibition, sensitivity to reward and punishment, and the ability to plan and formulate behaviors that are in accordance with social demands (Ishikawa & Raine, 2003; Raine, 2002; Séguin, 2004). Acknowledging this evidence, researchers such as Rocha and colleagues (2014) concluded that improving neuro-cognition in general and executive functioning in particular could be beneficial for inmates. Specifically, it was thought that improving offenders’ executive capacities would be helpful to their overall functioning as these skills are crucial for the demands of everyday life (Rocha et al., 2014). Moreover, it has been suggested that improving these skills would make offenders better able to take advantage of existing correctional remediation programs (Ross & Hoaken, 2011).

Evidence suggests that efforts to improve cognitive capacities are beneficial for antisocial individuals including offenders. A pilot study by Rocha and colleagues (2014) exploring the
efficacy of cognitive remediation for female inmates found a cognitive remediation designed to treat neurocognitive deficits including cognitive flexibility, memory, and planning produced significant positive changes across several neurocognitive domains. Furthermore, female inmates who received the cognitive remediation also had decreases in negative emotional states (e.g., depression, anxiety, tension/stress) and demonstrated significant improvements in prison behavior. The work by Rocha and colleagues (2014) demonstrates that it is possible to improve cognitive capacities among offenders. Perhaps more importantly, however, this work demonstrates the widespread benefits of cognitive remediation for this population as the training produced significant improvements among inmates’ problematic behaviors in addition to their psychological well-being.

The question remains, however, if there is a sensitive period for executive function remediation in general, and in particular among antisocial individuals and correctional populations. Though, it would seem that early adulthood may represent such a period. This may be especially true for correctional populations who are often cognitively immature in regards to capacities for decision-making, impulse control etc. as compared to non-offending peers.

Identification of a potential sensitive period in executive development and/or remediation is critical in identifying optimal times for interventions aimed at reducing underlying executive deficits among these populations.
CHAPTER TWO

THE CURRENT STUDY

Purpose of the Current Study

The purpose of the current study was to explore the relations between age, executive functioning, and recidivism (continued antisocial behavior) among offenders. Specifically, the current research sought to parse out how much executive functioning contributes to future antisocial behavior and whether deficits in executive functions, which are pervasive among antisocial individuals, improved following correctional treatment. Moreover, this study sought to identify whether changes in executive functioning vary by age as it was expected that younger inmates would be more amendable to cognitive remediation than older inmates.

The goal of this research was to identify whether it is possible to improve executive functioning in a correctional population and whether such gains may serve as an important component in reducing recidivism and future antisocial behavior. Possible findings derived from this research could highlight key executive abilities that should be targeted as part of correctional treatment program, thereby improving treatment programs provided to offenders.

In addition, by exploring the relations of age and changes in executive function, this study sought to explore whether a potential sensitive period exists in regards to executive function remediation among antisocial individuals. Potential findings regarding a sensitive period for executive remediation could serve to inform treatment programs provided to this population by differentiating who may benefit from such interventions and who may be less
responsive to such programming.

**Methods of the Current Study**

The current study consisted of a secondary data analysis of data collected by Fishbein and Sheppard (2006). The purpose of the work done by Fishbein and Sheppard (2006) was primarily to “identify fundamental differences in ECF function [executive cognitive functions] and emotional perception and regulation between inmates who respond favorably to standard correctional therapies, which are often cognitive-behavioral in approach, and those who respond poorly” (Fishbein & Sheppard, 2006, p. 3). More specifically, the project was designed to identify executive cognitive functions and emotional deficits in inmates with poor behavioral outcomes in the context of a standard but high-quality correctional therapy program.

**Selection of Fishbein and Sheppard (2006) for Reanalysis**

This data set was selected for reanalysis for several reasons. For one, it provided measures of various baseline executive functions and follow-up assessment of executive functions among a sample of adult offenders (aged 21 to 49 years) following participation in a correctional treatment program. (Note the inclusion of follow-up assessment of executive abilities was unique in this study as many other studies investigating executive functioning among antisocial and/or correctional populations only assessed executive functions at baseline as the primary focus was identifying whether such abilities (or deficits in such abilities) affect outcomes such as success in treatment programming, recidivism, etc.). As the data provided both baseline and follow-up assessment of executive functioning, it enabled me to investigate how executive functions may change over time. Moreover, variation in the ages within the sample allowed me to assess for possible age-related or developmental differences in executive
functioning and the applicability of potential interventions at various ages to remediate cognitive deficits, which was the primary interest of this study.

In addition, the Fishbein and Sheppard’s (2006) participants were enrolled in a high-quality Cognitive Behavior Therapy (CBT) program. CBT programs have demonstrated effectiveness for correctional populations (see Pearson, Lipton, Cleland, & Yee, 2002; Wilson, Bouffard, & MacKenzie, 2005) and are widespread in the U.S. correctional system (Feucht & Holt, 2016). Moreover, CBT is a therapeutic model that focuses on changing non-adaptive or destructive thinking patterns that in turn lead to the negative emotional states and undesired behavior (American Psychological Association [APA], 2017). CBT is a unique therapeutic approach as it is a “problem-oriented strategy” (Wenzel, 2017). This therapeutic module focuses on dealing with current problems with the goal being to help people to help themselves by changing maladaptive thinking patterns and developing healthy coping strategies (Beck, 2011; APA, 2017; DiGiuseppe, Venezia, & Gotterbarn, 2019).

The focus on changing maladaptive thought patterns and behaviors in CBT suggests this therapeutic model taps into many of core executive deficits observed among antisocial individuals (e.g., poor decision-making, impulsivity, response inhibition). Moreover, the focus on changing cognitions in this therapeutic approach may foster more mature ways of thinking that may facilitate improvement in other executive function deficits common among antisocial populations (e.g., planning, cognitive flexibility). Therefore, it was thought that participation in such programming would lead to improvements across various executive functions among this sample.

The CBT program provided to the inmates included in the study consisted of a “series of
3 groups that [met] for 90 minutes twice a week, totaling 50 sessions” (Fishbein & Sheppard, 2006, p. 22). In this adaptation of CBT, the first group, “Entry Point” involved curricula on “Thinking, Deciding, Changing” (Fishbein & Sheppard, 2006, p. 22). According to Fishbein and Sheppard (2006), the first group blended a decision making and a cognitive restructuring modality (e.g., a self-reflective process to search for triggers of misconduct) into a cognitive-behavioral modality (e.g., an external, skill building process) for self-change. The second group, “Communication,” was designed to orient members to the importance of effective communication in one’s everyday life and encouraged the experience of group cohesion and ownership by enabling members to create their own goals and rules through the use of dyad interviews (Fishbein & Sheppard, 2006). The third and final group, “Relationships,” focused on an examination of the way elements in the environment are dealt with or related to (Fishbein & Sheppard, 2006). Therefore, this third group focused on how program participants would implement skills leaned in the first and second groups in their everyday lives.

The specific content of the CBT program provided to the inmates seemed to address many of the executive deficits common in correctional populations. Moreover, the structure of the program also met program/intervention requisites needed to facilitate improvement in executive functioning. Specifically, research on executive remediation has demonstrated that improvement is dependent on the amount of time individuals spend practicing (Diamond & Ling, 2016). Among studies of various remediation types, including computerized cognitive training approaches and mindfulness programs, the length of training (e.g., duration) was associated with greater improvement (Diamond & Ling, 2016). Among computerized cognitive training programs (ranging from 2 to 14 weeks in duration), program length has been found to be associated with greater executive function improvement even when controlling for dose (e.g.,
session length) and frequency (e.g., number of sessions) (Basak, Boot, Voss, & Kramer, 2008; Diamond & Ling, 2016; Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). The CBT program provided to inmates in Fishbein and Sheppard (2006) consisted of approximately 25 weeks of sessions during which time inmates participated in two 90-minute sessions per week for a total of 50 sessions. Therefore, the program provided ample time needed to practice executive functioning skills targeted in this intervention.

It should be noted that failure to observe changes in executive functioning among individuals in the sample would also be an important discovery. As previously discussed, CBT is pervasive among correctional institutions. The finding that CBT programming fails to produce meaningful changes in executive functioning would suggest an important factor contributing to antisocial behavior and its continuation (e.g., recidivism) is not being addressed in programs designed to rehabilitate this population.

**Rationale for Reframing Fishbein and Sheppard (2006)**

The primary rational for reframing Fishbein and Sheppard (2006) focused on the fact that this work largely ignored the importance of age in relation to the development of executive functioning and potential that age may be an important factor in regards to the efficacy of remediation efforts targeting executive functions. For all analyses conducted by Fishbein and Sheppard (2006), “adjustments were made for age” (p. 17); however, such analyses did not assess the contribution of age in potential effects or consider the developmental differences in executive functioning that exist among the age ranges captured in the sample.

Moreover, Fishbein and Sheppard (2006) did test whether change in cognitive functions predicted treatment response as indicated by treatment responsivity, gain, completion, and by segregations and infractions and studied relations between treatment gain and responsivity scores
and changes in scores on the Novaco Provocation Inventory. However, they did not assess whether changes in executive functioning were predictive of either changes in Novaco Anger Scale or Provocation Inventory scores. Therefore, they only considered how changes in executive functioning may affect recidivism as measured by institutional behavior. In a confined setting, such as a prison, behavior is largely constrained. Thus, institutional behaviors may not fully translate to how one would act without external constraints.

Furthermore, Fishbein and Sheppard (2006) assessed changes in infractions and segregations using inmates’ prior number of segregations and infractions and the frequency of such occurrences during treatment. This is problematic for three critical reasons. First, the time frame during which pre-treatment infractions and segregations occurred is unknown and not accounted for and therefore does not provide an equivalent comparison to the number of infractions committed during treatment. Second, observing the change in infractions from pre-treatment to during treatment is problematic as the observed outcome measure does not evaluate changes derived from the program as it does not assess behavior following program completion.

Lastly, using the number of during-treatment infractions and segregations as a proxy for change in antisocial behavior or recidivism is confounded by inmates’ enrollment in treatment. One cannot parse out effects receiving treatment versus being enrolled in treatment. It is possible that being enrolled in the program led to reductions in institutional infractions as inmates were largely occupied by time spent in treatment sessions or completing treatment-related work. By using during treatment infractions as a proxy of recidivism, rather than post-treatment infractions, it is unclear whether changes in number of infractions are a product of reduced
opportunities for making infractions or true behavioral change stemming from treatment-related gains.

This idea is similar to the time-course framework proposed by Linning, Bowers and Eck (2019), which states that although interventions can reap desired results overall, more subtle effects can occur at different times, namely before, during and after implementation. The time-work framework was proposed to guide the design and evaluation of interventions in criminal justice settings including correctional programming (Linning, Silver, & Papp, 2022).

Interestingly, directions of proposed effects of correctional rehabilitation on institutional misconduct vary. As noted by Linning et al. (2022), though correctional programs may seek to reduce problematic behaviors (e.g., rehabilitative programs) and increase employable skills (e.g., vocational and educational programs), these programs may simultaneously influence instances of prison misconduct which undermines the safety and order goals of these institutions. For example, Wooldredge (1998) noted that prions may observe increases in misconduct as inmates are spending more time with correctional staff. Spending more time participating in programming increases the amount of time inmates interact with staff which could increase the likelihood an inmate is caught for infractions.

A systematic review by Steiner, Butler and Ellison (2014) of prison misconduct research revealed mixed results regarding the influence of institutional routines and experiences. Specifically, though inmates involved in prison work assignments and religion-based programs committed less misconduct, those involved in other forms of prison programming engaged in misconduct more often. Other work by Linning et al. (2022) found that institutional misconduct decreased during prison programming, however, increases in misconduct were observed prior to and following exposure to programming. As noted by Vuk and Doležal (2019) reductions in
misconduct during programming may stem from structured time spent engaging in prosocial activities during programming which reduces idle time that can facilitate misconduct. Ideally, post-treatment infractions would be assessed and the length of time served post-treatment completion would have been recorded, allowing me to assess misconduct following treatment while controlling for time.

Taken together, the recidivism-related outcome measures used by Fishbein and Sheppard (2006) were not ideal. Utilizing change in Novaco Anger Scale and Provocation Inventory scores can be viewed as providing better indices of recidivism as these measures speak to an individual’s self-reported feelings of anger and reactivity. As these measures were provided pre- and post-treatment, they also provide indicators of change related to recidivism or antisocial behavior and/or feelings following participation in treatment, when such treatment has been able to take effect.

Finally, another motivation for reanalysis of Sheppard and Fishbein (2006) stems from the fact that they did not assess the mechanisms by which baseline executive functions may influence recidivism. Is baseline executive functioning inherently a predictor of recidivism, or is baseline executive functioning predictive of recidivism via its effect on an individual’s capacity for engaging in and benefiting from treatment? It has been well established that executive function deficits are associated with antisocial behavior and recidivism. In aiming to reduce recidivism among correctional populations, however, an understanding of the mechanisms that underlie this relation is critical.

**Contribution of this Reanalysis**

As outlined in the previous section, there are several gaps in the analyses conducted by Fishbein and Sheppard (2006). This rich data set provided a unique opportunity to explore
several important questions pertaining to the relations among age, executive functioning, antisocial behavior and recidivism among a correctional population.

Specifically, my reanalysis addressed the gaps in Fishbein & Sheppard (2006) by: 1) investigating the effects of age in relation to executive function remediation among a correctional sample, 2) investigating the mechanisms by which baseline executive functions predicted future recidivism, specifically whether such effects were exerted via offenders’ ability to complete and engage in treatment, and 3) exploring relations between changes in executive functioning and recidivism with more robust (e.g., comprehensive) and potentially more reliable measures.

Thus, this research is unique as it stands at the crux of developmental and forensic research. This research is developmental as the effects of age and developmental stages were considered in exploring potential relations among variables of interest. Moreover, it is developmental as I considered the process through which change happens, rather than simply seeking to identify which factors predict various outcomes. It is useful to take a developmental perspective within the field of criminal justice as this perspective can shed light on the factors that may affect why a person may positively respond to correctional treatment or fail to benefit from such interventions. Exploring potential mechanisms through which correctional treatment can lead to reductions in antisocial behavior is paramount, as such research can inform rehabilitation efforts in order to best serve this population.

**Current Study Hypotheses**

To address gaps in the literature regarding executive functioning and recidivism among correctional populations and in analyses previously conducted by Fishbein and Sheppard (2006) the current study sought to test the following hypotheses:
H1: Does baseline executive functioning predict recidivism? Is the relationship between baseline executive functioning and recidivism mediated by participation/engagement in treatment? It was predicted that baseline executive functioning would predict recidivism such that worse performance on executive function performance tasks at baseline would be associated with higher rates of recidivism (e.g., less change in number of infractions from baseline to follow-up). Moreover, it was predicted that the relation between baseline executive function and recidivism would be mediated by participation and engagement in treatment, such that those who had better baseline executive functions would demonstrate greater treatment engagement and subsequently demonstrate lower rates of recidivism.

H2: Does change in executive functioning vary by age? It was predicted that changes in executive functioning would vary by age such that younger inmates would show greater improvement in executive functioning following treatment than older inmates.

H3: Does engagement in treatment predict change in EF? It was predicted that engagement in treatment would predict change in executive functioning such that inmates with higher levels of treatment engagement (e.g., greater treatment gains) would demonstrate greater improvements in executive functioning.

H4: Does change in executive functioning predict recidivism? It was predicted that changes in executive functioning would predict recidivism such that inmates with greater improvement in executive abilities would demonstrate less recidivism (e.g., greater reductions in number of infractions from baseline to follow-up, greater change on Novaco Anger Scale (NAS) and Provocation Inventory (PI) from baseline to follow-up).
CHAPTER THREE

METHODS

Study Design

The current study involved secondary analyses of the data collected in Fishbein and Sheppard’s (2006) study, which was funded by the National Institute of Justice (NCJ No. 216303). Originally collected with a sample of male inmates in Maryland state prisons (Roxbury Correctional Institution (RCI), Western Correctional Institution (WCI), and Maryland Correctional Training Center (MCTC)) between 2003 and 2005, the purpose of the study was to ascertain the role of neuropsychological functioning, including executive cognitive functions, and emotional perception and regulation in inmates’ responses to standard correctional therapy.

The present study examined these data to explore the potential for improving executive abilities in correctional populations and whether improvements in executive functioning are associated with improved outcomes, including responses to treatment and behavioral improvement (e.g., reduced recidivism and antisocial behavior). In addition, this study assessed whether executive gains vary by age. Moreover, this study tested whether the relationships between baseline executive abilities and recidivism are mediated by inmates’ performance and engagement in correctional treatment. The purpose of this analysis was to identify whether executive deficits are inherently related with recidivism (e.g., if these deficits are an underlying cause of continued antisocial behavior), or if executive deficits inhibit the efficacy of standard correctional treatment, and in turn lead to minimized behavioral improvement (e.g., greater
Participants/Participant Characteristics

The sample evaluated in the current study consisted of 203 subjects. This number represents the number of subjects following deletion of participants with no data \( n = 7 \) and subjects who provided consent, but did not enroll in the treatment program \( n = 22 \). (Note the total sample in Fishbein & Sheppard (2006)’s data set was 232). These subjects ranged in age from 21 to 45 years \( (M = 30.95, SD = 5.69) \). The sample was predominantly Black \( (75.9\%, n = 154) \) and included smaller numbers of White \( (18.2\%, n = 37) \), American Indian \( (2.0\%, n = 4) \), and Hispanic \( (1.0\%, n = 2) \) inmates. An additional six subjects \( (3.0\%) \) identified as Other.

Histories of alcohol and substance use were common among the sample. Nearly 86\% of subjects had reported using drugs in the past \( (n = 174, 85.7\%) \) and more than 40\% of subjects \( (n = 83, 40.9\%) \) had previously received treatment for drug abuse (e.g., had received treatment for drug abuse at least once). In addition, approximately 30\% of subjects \( (n = 59, 29.1\%) \) had previously received treatment for alcohol abuse (e.g., had received treatment for alcohol abuse at least once).

Among this sample, verbal IQ scores ranged from 70 to 117 \( (M = 88.09, SD = 11.56) \) and performance IQ scores ranged from 70 to 136 \( (M = 91.06, SD = 15.24) \). Full scale IQ scores ranged from 70 to 122 \( (M = 88.85, SD = 12.59) \). Subjects had an average of 11.39 years of education (range: 4-21 years, \( SD = 1.66 \) )

On average, at the time of baseline testing, subjects had served 36.68 months \( (3.06 \text{ years}) \) in prison for their current sentence (range: 1-251 months, \( SD = 38.54 \)). The total amount of time incarcerated (e.g., time spent incarcerated for the current sentence in addition to time served for previous offenses) was not reported. Among the sample, the average length for one’s current
sentence was 15.64 years (range: 1-80 years, \(SD = 11.77\)). On average, subjects reported having approximately three charges (\(M = 2.91, SD = 2.15\)). At the time of testing, the majority of subjects were incarcerated for violent offenses (62.6\%, \(n = 127\)) while one third (33.0\%, \(n = 67\)) were incarcerated for drug offenses. Only 3\% (\(n = 6\)) of subjects were incarcerated for property offenses. See Table 1 for descriptives on other relevant sample demographics including levels of psychopathy, baseline aggression and treatment readiness.

Table 1. Means, Standard Deviations, and Ranges for Demographic Variables Assessed at Baseline

<table>
<thead>
<tr>
<th>Variable</th>
<th>(N)</th>
<th>(M)</th>
<th>(SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Readiness (Treatment Readiness: Short Scale, Serin et al., 2002)</td>
<td>192</td>
<td>13.68</td>
<td>3.48</td>
<td>5-22</td>
</tr>
<tr>
<td>Primary Psychopathy (LSRP Levenson et al., 1995)</td>
<td>198</td>
<td>32.68</td>
<td>8.36</td>
<td>16-54</td>
</tr>
<tr>
<td>Secondary Psychopathy (LSRP Levenson et al., 1995)</td>
<td>200</td>
<td>23.32</td>
<td>5.52</td>
<td>11-37</td>
</tr>
<tr>
<td>Total Psychopathy (Primary + Secondary) (LSRP, Levenson et al., 1995)</td>
<td>197</td>
<td>56.11</td>
<td>11.64</td>
<td>30-85</td>
</tr>
<tr>
<td>Proactive Aggression (RPQ, Raine et al., 2006)</td>
<td>203</td>
<td>6.54</td>
<td>4.77</td>
<td>0-24</td>
</tr>
<tr>
<td>Reactive Aggression (RPQ, Rain et al., 2006)</td>
<td>202</td>
<td>11.22</td>
<td>4.53</td>
<td>0-22</td>
</tr>
</tbody>
</table>

Note. \(^a\)Treatment Readiness total scores range from 0 to 24, higher scores indicate greater treatment readiness. \(^b\)RPQ proactive aggression scores range from 0 to 24 and RPQ reactive aggression scores range from 0 to 22, with higher scores indicating greater levels of aggression. \(^c\)LSRP primary psychopathy scores range from 16 to 64, LSRP secondary psychopathy scores range from 10 to 40, with higher scores indicating higher levels of psychopathic traits. Total scores on LSRP range from 26 to 104.

\(^d\)Although there are no cutoff scores for the LSRP, Brinkley, Schmitt, Smith and Newman (2001), used a cutoff score of 58 (e.g., individuals with LSRP total scores \(\geq 58\) were considered “psychopathic”, individuals with LSRP total scores of 49-57 were considered “mixed” and individuals with total LSRP scores \(\leq 48\) were considered “nonpsychopathic”.
Measures and Covariates

Age Group

To explore the relationship between age and change in executive abilities post-treatment, I grouped age into five categories (e.g., 21-25 years, 26-30 years, 31-35 years, 36-40 years, and 41-45 years). My decision was based largely on research demonstrating variability in the development of executive functioning across age and variability in executive task performance through the lifespan.

Certain executive abilities have been shown to mature more quickly. For example, research on individuals between eight and 30 years old has shown that individuals reach mature-adult level performance of processing speed at 15 years, response inhibition at 14 years, and working memory at 19 years (Luna, Garver, Urban, Lazar, & Sweeney, 2004). However, other executive abilities mature throughout late adolescence and into adulthood. For example, Huizinga, Dolan and van Der Molen (2006) found that certain aspects of inhibitory control continue to improve until 21 years old, and performance on more complex executive tasks (e.g., Wisconsin Sorting Task (a measure of cognitive flexibility), Tower of London (a measure of planning ability)) improves into early adulthood. Similarly, Romine and Reynolds (2005) found executive functions show divergent developmental trajectories with planning and verbal fluency continuing to develop linearly between late adolescence and early adulthood, suggesting that these capacities are not fully established until early adulthood. Other research has shown that cognitive flexibility continues to improve between the ages of 20 and 29 (Zelazo et al., 2014).

Overall, evidence suggests that more complex executive abilities, such as planning and decision-making mature later (Huizinga et al., 2006), and that hot EFs develop more slowly than cool EFs (Prencipe et al., 2011). As discussed, Hot EFs can be understood as goal-directed,
future-oriented cognitive processes elicited in contexts that generate emotion, motivation and tension between instant gratification and long-term rewards (Zelazo & Müller, 2002; Zelazo et al., 2005). Hot EFs mirror real life decision-making in contexts that generate affective responses. In contrast, cool EFs are cognitive abilities separate from affective processes (Poon, 2018). As such, cool EFs are often seen as the capacity for executive abilities in neutral situations. Therefore, evidence suggests people develop executive abilities before they are able to implement these capacities in more emotionally charged situations that typically define real-world experiences.

Mirroring these trends, capacities for self-regulation increase steadily from preadolescence into early adulthood, plateauing in the mid-20s (Steinberg et al., 2018). Improvements in executive task performance and self-regulation during early adulthood coincide with a period of steep maturation of frontal pathways that occur around ages 17 to 25 (Barker, Andrade, Morton, Romanowski, & Bowles, 2010).

Research on developmental trajectories has shown variability across different executive abilities, with some executive abilities showing linear or cubic trends. For example, in a study exploring the developmental trajectories of executive functioning from adolescence to old age, Ferguson, Brunsdon and Bradford (2021) found that performance on the Tower of Hanoi task (a measure of problem-solving and planning abilities) increases from 10 to 30 years old and then declines from 30 to 70 years old, followed by a small, but variable increase from 70 years onward. Between ages 10 to 30, scores on an operation span task increased (modified version of Unsworth et al.’s 2005 automated operation span task (OSpan)), indicating improvement in working memory capacity. This increase was followed by a decrease in scores from the age of 30, indicating declines in working memory capacity (Ferguson et al., 2021). In addition,
Ferguson and colleagues (2021) found that Stroop task performance (measure of one’s ability to inhibit cognitive interference/inhibitory control) improves from 10 to 35 years old and then declines from the age of 36 to 86.

Beyond the research cited above, which demonstrates age-related changes in executive abilities across different periods (e.g., emergent, early and later/middle adulthood), my choice to explore differences in executive improvement across more narrow age ranges was also based on concerns that broader age ranges could decrease sensitivity (De Luca et al., 2003) and mask the non-linear nature of development observed among certain executive abilities (Taylor, Barker, Heavey, & McHale, 2015).

Overall, in light of research suggesting rapid neurological changes during the early 20s in conjunction with research demonstrating improvements in executive ability task performance during this period, I decided to include individuals ages 21 to 25 as an age group. As I previously discussed, the early 20s is a period in which significant cognitive development occurs; therefore, I expected that these individuals were likely distinct from other inmates. Similarly, research demonstrates that certain EFs improve until the ages of 30 and 35 while others begin to decline after the age of 30 (Ferguson et al., 2021). Research suggests that certain brain regions fail to reach plateaus in structure growth curves by the age of 30 (Tamnes et al., 2010) and that the late 20s (e.g., 26-30 years) and early 30s (e.g., 31-35 years) are also distinct periods of executive ability development. The late 20s and early 30s are likely periods in which EFs might still be improving, but at a protracted rate compared with the early 20s. Hence, the late 20s/late emerging adulthood (e.g., 26-30 years) and early 30s/early mid-adulthood (e.g., 31-35 years) were included as distinct age groups.
Finally, I did not anticipate significant differences in potential for executive gains across middle adulthood (e.g., among participants 36-49 years in the sample). However, I chose to include different age categories (e.g., 36-40 years, 41-45 years) to encompass this segment of the sample to account for any age-related differences that could potentially be overlooked. Note, however, only three participants were 41 years of age or older. Due to inadequate sample size, age-group related analyses only considered participants between 21 and 40 years of age.

**General Neuropsychological Function**

**Multidimensional Aptitude Battery (MAB) (Jackson, 1984, 1998).** The Multidimensional Aptitude Battery is a group-administered abilities test based on the Wechsler Adult Intelligence Scales-Revised (WAIS-R) (Wechsler, 1981). Designed as an alternative to the WAIS-R, the MAB tests general intelligence in group settings for individuals aged 16 years and older (Carless, 2000). As Krieshok and Harrington (1985) note, the primary advantage of MAB is that it was designed to measure in an “objective, group-format the same underlying factors measured by the WAIS-R” (p. 87). The MAB consists of five verbal subtests (information, comprehension, arithmetic, similarities and vocabulary) and five performance subtests (digit symbol, picture completion, spatial, picture arrangement, and object assembly). Scores on these subtests are used to create a Verbal score and a Performance score, and a combined Full Scale score. The time limit for each subtest is seven minutes and administration of the entire measure takes 100 minutes.

Due to time constraints associated with administering assessments in a prison setting, the short form of the Multidimensional Aptitude Battery was administered (Sheppard & Fishbein, 2006). The short form version of the MAB outlined by Hill & Jackson (1984) is comprised of four subtests (comprehension, similarities, digit symbol, and spatial); the short form is a strong
predictor of full scale IQ (Hill & Jackson, 1984). Completion time for the short form version of the MAB is approximately 35 minutes.

**Questionnaires**

**Background Inventory (Fishbein & Sheppard, 2006).** A revised version of the Addiction Severity Index (ASI) (McLellan et al., 1992) was used to measure the nature and extent of participants’ prior drug use in addition to other background factors, such as gender, marital status, socioeconomic status (Hollingshead rating), religious preference, race/ethnicity, and family history of drug and alcohol use and other psychiatric disorders. Other factors assessed in the Background Inventory included information on participants’ medical status (e.g., number of previous hospitalizations for medical problems, chronic medical problems that interfere with life, current prescriptions for medication to address physical problems, history of head injury, etc.), education and/or employment status, and interviewer impressions of participant (e.g., “obviously depressed/withdrawn,” “obviously hostile,” “obviously anxious/nervous,” or “having trouble with reality testing, thought disorders, paranoid thinking” etc.).

**The Levenson Self-Report Psychopathy Scale (LSRP) (Levenson, Kiehl, & Fitzpatrick, 1995).** The Levenson Self-Report Psychopathy Scale was originally designed to assess psychopathy in individuals who do not demonstrate extreme levels of this trait (Psederska et al., 2020). This measure was designed to reflect the dual-factor model of psychopathy, which distinguishes two subtypes of psychopathy: primary and secondary psychopathy (Karpman, 1941; Blackburn, 1975; Vassileva, Kosson, Abramowitz, & Conrod, 2005).

Primary psychopathy focuses on interpersonal and affective factors including traits such as coldness and lack of feelings of guilt or remorse (Blackburn, 1975). In contrast, secondary psychopathy measures chronic risky and impulsive behaviors and antisocial lifestyles; thus,
secondary psychopathy focuses on behavioral manifestations of psychopathy (Hare, 2003). Corresponding with this dual-factor model of psychopathy, the LSRP consists of a primary psychopathy factor that measures a selfish, uncaring, manipulative posture towards others and a secondary factor that measures impulsivity and a self-defeating lifestyle (Levenson et al., 1995).

The LSRP consists of 26 items measured on a 4-point Likert scale, ranging from 1 (completely disagree) to 4 (completely agree). Sixteen items measure primary psychopathy (e.g., psychopathic emotional affect and manipulative and selfish behavior) and ten items measure secondary psychopathy (e.g., psychopathic lifestyle including impulsivity and antisocial behavior). Primary psychopathy scores range from 16 to 64 and secondary psychopathy scores range from 10 to 40. In order to control for response sets, some items are reverse scored. Higher scores indicated higher levels of psychopathic traits.

The measure has sound psychometric properties, including internal validity (Cronbach’s alpha of .84, Sellbom, 2011) and adequate convergent and discriminant validity (see Brinkley, Schmitt, Smith, & Newman, 2001; Sellbom, 2011; Shou, Sellbom, & Han, 2017; Garofalo, Noteborn, Sellborn, & Bogaerts, 2018). Moreover, the measure has been validated with both prisoner and non-prisoner samples (Sellbom, 2011) in addition to many other populations, including university undergraduate (Salekin, Chen, Sellborn, Lester & MacDougall, 2014) and general community samples (Jones, Miller & Reynolds, 2012). Note three participants had invalid primary psychopathy scores (e.g., scores of 12) and one participant had an invalid secondary psychopathy score (e.g., score of 9). These participants were recoded as “missing data” for these variables.

The Reactive-Proactive Questionnaire (RPQ) (Raine et al., 2006). The Reactive-Proactive Questionnaire is a 23-item self-report measure designed to evaluate two forms of
aggressive behavior: reactive and proactive aggression. Reactive or hostile aggression can be understood as combative responses to perceived threats; therefore, reactive aggression is in response to a provocation. In contrast, proactive or instrumental aggression involves predatory attacks motivated by external rewards (Card & Little, 2007; Dodge, 1991). This behavior is goal-directed to achieve an objective beyond inflicting physical violence. Consequently, proactive aggression is instrumental, organized, and “cold-blooded” and not typically accompanied by autonomic arousal (Dodge, 1991; Meloy, 1998; Mirsky & Siegel, 1994).

The RPQ is appropriate for use with children, adolescents, and young adults (Pechorro, Ayala-Nunes, Kahn, & Nunes, 2018), and the measure has demonstrated internal reliability in various samples, including youth (Seals, Sharp, Ha, & Michonski, 2012) and adolescents (Raine et al., 2006; Borroni, Somma, Andershed, Maffei, & Fossati, 2014) as well as male prisoners (Cima & Raine, 2009) and other adult populations (Dinić & Raine, 2020). RPQ items are rated on a 3-point ordinal scale (0 = Never, 1 = Sometimes, 2 = Often); eleven items measure reactive aggression (e.g., “How often have you yelled at others when they have annoyed you?”) and 12 items measure proactive aggression (e.g., “How often have you had fights with others to show who was on top?”). Summed totals yield measures for proactive and reactive aggression as well as a total aggression score. Scores on the reactive aggression scale range from 0 to 22, and scores on the proactive aggression scale range from 0 to 24. Higher scores indicate higher levels of aggression. One participant was listed as having an invalid score on the reactive aggression scale (e.g., a score of 24) and was recoded as missing for this item.

The Novaco Anger Scale and Provocation Inventory (NAS-PI) (Novaco, 1994). Initially developed in conjunction with the MacArthur Foundation Network on Mental Health and Law, the NAS-PI is a two-part test (the NAS and the PI) designed to assess anger as a
problem of psychological functioning and physical health and also to assess change stemming from therapeutic treatment (Novaco, 2003). The NAS-PI was developed in order to address the concern that existing measures of anger had not been grounded in a theoretical foundation (Novaco & Welsh, 1989).

According to Jones and colleagues (1999), the measure’s subscales closely reflected Novaco’s information processing model of anger (Novaco, 1975); the measure also represents a clear progression from the original Novaco Anger Inventory (NAI), also known as the Provocation Inventory (Novaco, 1975). The NAS-PI was developed and standardized with community and clinical populations and was intended to be used as a measurement tool for research purposes as well as for individual assessment and outcome evaluation.

In the original version of the NAS-PI (1994 edition), Part A, the Novaco Anger Scale (NAS), consisted of 48 items that were divided equally into three domains that corresponded with Novaco’s information processing model of anger (Novaco, 1975) (e.g., Cognitive, Arousal, Behavioral), which were then further divided into four subscales. Part A focused on how an individual experiences anger (Novaco, 1994). The Cognitive domain (16 items) measured cognitive aspects of anger including: attentional focus (e.g., “I notice annoying things right away”), rumination (e.g., “Once something makes me angry, I keep thinking about it”), hostile attitudes (e.g., “Every week I meet someone I dislike”), and suspicion (e.g., “I know that people are talking about me behind my back”).

The Arousal domain (16 items) included subscales of intensity (e.g., “When I get angry, I get really angry”), duration (e.g., “When I get angry, I stay angry for hours”), somatic activation/tension (e.g., “My head aches when people annoy me”), and irritability (e.g., “I get annoyed when someone interrupts me”).
The Behavioral domain (16 items) consisted of subscales for: impulsive reaction (e.g., “My temper is quick and hot”), verbal aggression (e.g., “When someone yells at me, I yell back at them”), physical confrontation (e.g., “When I get mad, I can easily hit someone”), and indirect expression (e.g., “I feel like smashing things”). NAS items were measured on a 3-point Likert scale (1 = Never True, 2 = Sometimes True, 3 = Always True). The Novaco Anger Scale (NAS) yielded five scores: Cognitive, Arousal, Behavioral, and Anger Regulation subscale scores as well as a NAS total score.

Part B of the NAS-PI (e.g., the Provocation Inventory (PI)) consisted of 25 items that described situations that might induce anger in the respondent. These items are subdivided into five subscales that summarized the nature of the provocation: disrespectful treatment, unfairness/injustice, frustration/interruptions, annoying traits, and irritations. For each item, respondents were asked to report how angry they would be in response to various situations. Item responses were measured on a 4-point Likert scale from 1 (Not at all angry) to 4 (Very angry). Responses to PI items were summed to yield a PI Total score.

In the current data set, a small number of participants had “invalid” scores for NAS subscales. As I previously mentioned, for each of the NAS domain subscales, scores ranged from 4 to 12. For baseline measurements, in the NAS Cognitive Domain, one participant had a score of “13” on the Rumination subscale. In addition, a small number of participants had scores of “3” for three NAS Arousal domain subscales (e.g., Duration (1 participant), Somatic Activation/Tension (1 participant), and Irritability (1 participant)). In addition, in two of the subscales for the NAS Behavioral domain (e.g., Impulsive Reaction and Indirect Expression subscales), one participant had a score of “3”. In the NAS Behavioral domain Indirect Expression subscale, one participant had a score of “13” and another participant had a score of
An additional participant had a score of “13” on the NAS Behavioral domain Physical Confrontation subscale.

On the Provocation Inventory, subscale scores ranged from 5-20. However, again, a small number of participants had scores that were invalid at baseline administration. For example, on the Disrespectful Treatment subscale, two participants had scores of “0,” one participant had a score of “3” and one participant had a score of “4.” On the Unfairness subscale, two participants had scores of “0” and one participant had a score of “24.” For the Frustration subscale, two participants had scores of “0” and an additional two participant had scores of “2” while one participant had a score of “151”. For the Annoying Traits subscale, two participants had scores of “0” and one participant had a score of “2”. Lastly, for the Irritations subscale, two participants had scores of “0”, one participant had a score of “1” and one participant had a score of “2”.

At Follow-Up 4 (final follow-up, post-treatment), I also found several invalid scores across the NAS-PI subscales. Within the NAS Arousal domain, two participants had scores of “3” on the Intensity subscale. Within the Behavioral Domain, two participants had scores of “3” for the Impulsive Reaction subscale and one participant had a score of “14” for the Indirect Expression subscale. For the Provocation Inventory, for the Annoying Traits subscale, one participant had a score of “4”.

To address these invalid values, I computed new variables for NAS Cognitive Domain, NAS Arousal Domain, NAS Behavioral Domain and Provocation Inventory-Annoying Traits subscale. In addition, I also created a total NAS score (e.g., sum of scores on Cognitive, Arousal

---

1 In Fishbein and Sheppard (2006), subjects were administered the NAS-PI four times: 1) at baseline, prior to treatment, 2) after completion of first treatment group, 3) after completion of second treatment group, and 4) after completion of third treatment group/completion of entire treatment program. In this analysis, I only considered NAS-PI scores at baseline and follow-up 4 (e.g., post-treatment).
and Behavioral Domains) and total PI score (e.g., sum of scores on disrespectful treatment, unfairness/injustice, frustration/interruptions, annoying traits, and irritations subscales) that were the sum of the subscale scores in each domain including only participants with valid scores on the subscales that constitute the broader domain. Therefore, individuals who had invalid scores on any subscale within a domain were considered “missing” for the domain. These new variables excluded invalid values and were used to measure changes in NAS-PI scores from baseline to follow-up (e.g., post-treatment).

Since its original publication, the NAS-PI has been modified to include an intermediate 1998 version (Lindqvist, Dåderman, & Hellström, 2005) as cited by Moeller, Novaco, Heinola-Nielsen & Hougaard, 2016); the most recent edition was formally published in 2003 (Novaco, 2003). In the most recent edition, the NAS and PI sections were designated as separate, but complementary instruments and an additional Anger Regulation subscale was added to the Novaco Anger Scale (Part A). This version of the NAS includes 60 items (as compared to 48 in the 1994 version) and covers four domains, yielding four subscale scores in addition to a total NAS score. In addition to these changes, the 2003 version includes a subset of four “attentional focus” items within the Cognitive subscale, which replaced four items concerning “justification”, and other changes in wording were made in the instrument (Novaco, 2003).

Executive Cognitive Function (ECF) Tasks

The executive cognitive function tasks employed in the original study were “computerized and non-intrusive”; these tasks did not induce physical or emotional discomfort among participants, and the original researchers reported that most subjects enjoyed participating in these tasks (Fishbein & Sheppard, 2006, p. 19). Administration of cognitive tasks did not interfere with standard prison operations such as times scheduled for meals or counts or during
lock downs (Fishbein & Sheppard, 2006, p. 19). Each cognitive task took 10-20 minutes to complete.

**The Dice Game (Fishbein & Sheppard, 2006).** Designed to assess decision-making, a Dice Game was created to measure different aspects of decision-making, including risky/rational choices, betting behavior, reaction time, and risk adjustment. The measure assessed individuals’ willingness to take risks and sensitivity to consequences (Fishbein & Sheppard, 2006, p. 19).

In the Dice Game, subjects were presented with three virtual dice. In condition one, subjects were presented with a die with one yellow side and five blue sides. In condition two, subjects were presented with a die with two yellow sides and four blue sides, and in condition three, subjects were presented with a die with three yellow faces and three blue faces. Subjects were given opportunities to roll these dice and guess the color of the face following the roll. This decision involved gambling a certain number of points associated with each choice (odds are 10 vs 90, 20 vs 80, 30 vs 70, 40 vs 60, and 50 vs 50). If the subject selected the correct color, the points associated with that choice were added to the total points score; however, if the subject chose the wrong color, the same number of points were subtracted from the total points score.

Participants started each condition of the game with 100 points and were told to “use a plan and try to win as many points as [they] can” (Fishbein & Sheppard, n.d., p. 53). For this game, the experimenter conducted three practice trials to explain the game and 45 “real” trials.

The ratio of colored sides (e.g., 5:1 = highest risk, 4:2 = moderate risk, 3:3 = low risk) and the balance between the associated rewards varied independently between trials according a fixed pseudorandom sequence. This sequence ensured that each balance of reward and each ratio of yellow and blue side boxes co-occurred equally, with the restriction that on all trials—with an
unequal ratio of yellow and blue sides—the larger reward was associated with the least likely outcome, thereby capturing the conflict inherent in risk-taking scenarios.

Subjects with ECF impairment have been shown to take greater risks in pursuit of a large reward and to tolerate a higher probability of a large loss (Fishbein & Sheppard, 2006). As Fishbein and Sheppard (2006) indicate, this tendency often describes inmates who engage in impulsive offenses. Performance scores generated by this task included percentage of choice of the most likely outcome (e.g., percentage of safe choices), number of highest risk choices (e.g., number of times participants chose yellow with one yellow side and bet 90), and number of risky choices (e.g., number of times participants chose yellow with one yellow side with no neutral bets) and mean deliberation times as a function of the balance of rewards.

Change in executive functioning was evaluated by comparing baseline and post-treatment performance on the Dice Game. Specifically, to assess change in decision-making, difference scores (e.g., baseline scores minus post-treatment scores) were calculated for the percentage of safe choices made throughout entire task and the number of risky choices made. In addition, change in processing speed was assessed by comparing baseline and post-treatment reaction times across all Dice Game trials (e.g., reaction time for all trials at baseline minus reaction time for all trials post-treatment).

**Stop-Signal Task.** The stop-signal task is a neurocognitive task designed to provide a sensitive measure of the time needed to inhibit or suppress inappropriate motor responses (Morein-Zamir & Sahakian, 2010). The stop-signal paradigm was originally developed by Gordon Logan in the 1980s, based on a cognitive task first used by Lappin and Erikson in 1966 (Morein-Zamir & Sahakian, 2010). The Stop-Signal Task used by Fishbein and Sheppard (2006) began with the computer displaying an asterisk or a circle, alternatively. When the asterisk was
displayed, participants were instructed to quickly press the left mouse button and when the circle was displayed, they were instructed to quickly press the right mouse button.

In this task, incorrect responses (e.g., pressing the wrong button) elicited a beep and the subject had to correct their response. This correction slowed their reaction time. If the subject responded slowly, they were presented with a message of “too slow” on the screen and if they responded prior to the presentation of the stimulus, they were presented with a message reading “too fast” on the screen.

In the next section of the task, participants were to press the corresponding buttons for the asterisk and the circle, but if a tone sounded after the stimulus was presented, they had to alter their response and quickly press the middle mouse button. If participants pressed the right or left mouse buttons before the tone was presented, they received notification that they responded incorrectly and the task continued. The speed with which the tone was presented varied throughout the task and was affected by the speed of participant’s initial reaction time, as recorded in the first portion of the task (e.g., in baseline training condition). The Logan Stop-Signal task “requires deep concentration, impulse control, and timing ability” (Fishbein, & Sheppard, 2006, p. 20).

There were three conditions of this task. First, was the baseline condition in which no tone was presented; 36 trials were included in the baseline condition. The second and third conditions (e.g., Blocks 1 and 2) both included presentation of the tone distracter. Blocks 1 and 2 each consisted of 48 trials. The tone trials consisted of the presentation of “easy” and “hard” tone distracters. Easy tone distracters were when the tone was displayed immediately after the visual stimulus was presented. Hard tone distracters were when the tone was displayed a few milliseconds after the visual stimulus was presented.
Relevant measures included the percentage correct on various trials, the number of early, late and non-presses on Tone Blocks and Non-Tone Blocks, and reaction times for all conditions. The percentage correct on Tone trials, number of early, late and non-presses on Tone Block trials indicate the ability to suppress unwanted or inappropriate actions and impulses (e.g., ‘response inhibition’). Higher numbers of early, late and non-presses on Tone Block trials indicate reduced capacity for response inhibition whereas a higher percentage of correct items on Tone trials indicates greater capacity for response inhibition. Reaction times for trials also provide an indication of one’s ability for response inhibition. Generally, lower reaction times for the Tone Blocks indicate greater response inhibition. The Stop Signal task took participants approximately 15 minutes to complete.

Traditionally, performance in the Stop-Signal paradigm is conceptualized as a race between a “go process”, which is triggered by the presentation of a go stimulus, and a “stop process,” which is triggered by the presentation of a stop signal (Logan & Cowan, 1984). When the stop process finishes before the go process, a response is inhibited. When the go processes finish before the stop process, a response is expressed/emitted (Verbruggen & Logan, 2008). The latency of the stop-process is covert (Verbruggen & Logan, 2008); response-inhibition latency cannot be observed directly as successful response inhibition results in the absence of an observable response (e.g., no response) in contrast to the latency of go responses (Verbruggen et al., 2019).

Different methods are employed to estimate the time it takes to implement response inhibition (e.g., Stop-Signal Reaction Time, SSRT) each requiring different assumptions and different input information (see Band, van der Molen, & Logan, 2003). However, the general idea is that the Stop Signal Reaction Time (SSRT) consists of the time between the start and
finish of the stop processes (Logan, 1981). The fixed or variable presentation of the stop signal, or the timing of the presentation of the stop signal (e.g., the stimulus-onset asynchrony (SOA)) indicates the start, and the finish time can be derived from the response rate on signal trials and the reaction time distribution on non-signal trials (or no-go trials).

As I discussed above, in the version of the Stop Signal task presented to participants by Fishbein and Sheppard (2006) during Tone Blocks, participants were instructed to press the middle key on a mouse when a tone was presented rather than pressing either the left or right mouse key corresponding with presentation of an asterisk or circle, respectively, as they were told to do during the baseline condition. Hence, there was no true “no go” response thereby excluding calculation of a Stop-Signal Reaction Time.

To approximate response inhibition, I calculated the difference in percentage of correct responses during baseline trials and across all tone block trials. It was hypothesized that larger scores would indicate worse response inhibition as larger scores would indicate greater impairment in performance following presentation of distracter tone presentations and therefore greater difficulty inhibiting prepotent response. Values on this item were as expected (range: .04-.68, $M = .29$), indicating that all participants performed better on the Stop Signal task during baseline trials than during tone block trials.

Note I originally planned to approximate response inhibition at baseline by comparing the difference in average reaction times for baseline trials (e.g., no tone trials) with average reaction times for tone block trials. Under this model, an approximation of response inhibition would be created by subtracting average reaction time during tone trials from average reaction time during baseline trials (e.g., average reaction time for baseline trials – average reaction time for tone block trials). Under this model, it was thought that larger, negative difference scores would
indicate worse response inhibition by indicating performance was more impaired with presentation of the distracter tone sounds and accompanying change in task responses.

However, when I ran descriptives for this variable, I found all participants had positive values for this item (range: 5.45–15.23, $M = 9.73$) indicating that all participants’ performance improved during tone block trials from the baseline trials; all participants had faster average reaction times for tone block trials than baseline trials. This trend was unexpected as I predicted that the presentation of the tone distracters would slow response times and subsequently, that average reaction times would be slower during the tone block trials than during the baseline trials. It is possible that participants grew familiar with the task during baseline trials, which may account for the observed universal trend of faster reaction times during the tone block trials. In light of this finding, I decided that the difference in reaction time from baseline to tone block trials was not a good approximation of response inhibition.

The Stroop Interference Task. The Stroop Interference Task is a widely researched measure of executive control (Naber, Vedder, Brown, & Nieuwenhuis, 2016). The Stroop Color and Word Test (SCWT) is a neuropsychological test used extensively for both experimental and clinical purposes. It assesses the ability to inhibit cognitive interference, which occurs when the processing of a stimulus feature affects the simultaneous processing of another attribute of the same stimulus (Stroop, 1935), a process known as the Stroop Effect (Scarpina & Tagini, 2017). This task uses previously learned information to assess three attributes of executive frontal lobe function: complexity, a “nonroutine” nature, and the novel use of old information (Fishbein & Sheppard, 2006, p. 20). The Stroop Task includes presentation of three forms.

The first presentation displays the words “red,” “green,” and “blue” as black letters on a white background. The subject is asked to read the words as quickly and accurately as possible
(e.g., Word ‘W’ condition). Next, the second presentation presents four X’s in three randomly repeating colors, and the subject must quickly say aloud the color that corresponds with the color of the X’s (e.g., Color ‘C’ condition). In the third presentation (e.g., the Color-Word ‘CW’ condition), the words “red,” “green,” and “blue” are displayed in colors that do not correspond with the meaning of the word. This is an interference task and requires that the subject say the color of the lettering ink, rather than the color spelled out by the word. In the incongruent condition, subjects are required to name the color of the instead of reading the word listed (Scarpina & Tagini, 2017). In this condition, participants are required to perform a less automated task (e.g., naming the color of the ink used) while inhibiting interference arising from a more automated task (e.g., reading the word listed (MacLeod & Dunba, 1998; Ivnik, Malec, Smith, Tangalos, & Petersen, 1996).

The difficulty in inhibiting the more automated process is what is known as the Stroop effect (Stroop, 1935). The Stroop Color-Word Test is widely used as a measure of individual’s ability to inhibit cognitive interference (Scaprina & Tagini, 2017); however, previous research also demonstrates the task’s ability to measure other cognitive functions such as attention, processing speed, cognitive flexibility (Jensen & Rohwer, 1966) and working memory (Kane & Engle, 2003).

The measures for this task include: Word Score, Color Score, Predicted Color-Word Score, and the Interference Score. Each of the first three scores reflect the total number of responses completed in 45 seconds for each form. Different methods have been proposed for calculating this interference score (see Scarpina & Tagini, 2017 for review); however, one of the ones most commonly used and the one employed in this study was that proposed by Golden (1978). In this method, the number of items correctly named in 45 seconds in each condition is
calculated, producing three number of correct items scores (e.g., W, C, and CW). These scores are then used to calculate a predicted CW score (Pcw) using the following format:

\[ P_{cw} = \frac{45}{((45 \times W) + (45 \times C))/(W \times C)} \]

which is equivalent to:

\[ P_{cw} = \frac{(W \times C)}{(W+C)} \]

Then, the Pcw value (e.g., predicted number of correct items on Color-Word condition) is subtracted from the actual number of items correctly named in the incongruous condition (i.e., Interference score = CW – Pcw) to create an interference score that is based on the performance in both of the congruous conditions (e.g., Word and Color conditions). Lower interference scores indicate greater difficulty inhibiting interference, whereas negative interference scores indicate “a pathological inability to inhibit interference” (Scarpina & Tagini, 2017, p. 2).

Reaction times for all conditions were also recorded by Sheppard and Fishbein (2006). Reaction times for congruous conditions (Color and Word) conditions provide an indication of individual’s overall processing speed and cognitive efficiency. Comparison of reaction times for the congruous conditions to the Color-Word condition further indicates the ability to inhibit cognitive interference, with longer reaction times in Color-Word condition indicating reduced ability in this domain. As noted by Fishbein and Sheppard (2006), this task is well documented as a reliable measure of prefrontal functioning, with high discriminant ability (see Audenaert et al., 2001; Baxter & Liddle, 1998; Nathan, Wilkinson, Stammers, & Low, 2001).

Treatment Readiness, Responsivity, and Gain Scale: Short Version (TRRG: SV) (Serin, Kennedy, & Mailloux, 2002)

The purpose of this scale is to assist correctional staff to systematically assess an offender's readiness for and responsivity to treatment and to measure the degree to which gains
have been made following treatment. Items in this scale were selected based on a review of the literature on questionnaire construction and based on discussions with clinicians and program staff. TRRG:SV items were selected based on a factor analysis with the original scale of 50 items in a sample of 265 male offenders entering a cognitive skills program.

Principle components analysis revealed one underlying factor for both the treatment readiness and responsivity domain (Serin et al., 2002). A reduction in items from 22 to 8 lowered the overall consistency for both domains from .90 to .82, which is still within the range for excellent internal consistency (Serin et al., 2002). Each item has a specific behavioral anchor and description to assist in scoring. Questions for each item are provided simply as a guide for use in the format of a semi-structured interview. The protocol recommends that questions be incorporated into interview-based assessment strategies (Serin et al., 2002). Individual items are summed in order to provide a total score that represents an individual’s readiness and responsivity for treatment. A higher score on this scale reflects greater readiness or responsivity for treatment whereas a lower score reflects less readiness or responsivity for treatment.

The construct of treatment readiness and responsivity is not as trait-specific but rather a process that reflects change (Serin et al., 2002). For this reason, the ratings should be completed prior to treatment and at the conclusion of the treatment program (Serin et al., 2002). These ratings, to be completed by program staff, can then be used as a measure of change. Although it is difficult to determine how much contact is required before the pre-treatment assessment can be completed, up to 3 sessions will be necessary (Serin et al., 2002). Treatment readiness and responsivity are critical in the process of determining treatment needs and placement (Serin et al., 2002). For instance, offenders who score low on readiness and responsivity might benefit
from a treatment primer session in order to prepare them for a treatment program, maximizing the potential treatment gains.

**Treatment Readiness: Short Scale.** The Treatment Readiness scale is self-administered and assesses eight factors related to an individual’s willingness and readiness to engage in the treatment process including: a) problem recognition, b) macro treatment benefits, c) micro treatment benefits, d) treatment distress, e) treatment goals, f) treatment behaviors, g) behavioral congruency, and h) treatment support. Offenders “complete this test by indicating how treatment ready they view themselves” (Fishbein & Sheppard, 2006, p. 25). The domain of treatment readiness was intended to operationalize the continuum of an offender’s readiness to engage in treatment as a way “to assist clinicians in determining treatment placement” (Serin et al., 2002, p. 3). This scale was shown to have “excellent internal consistency producing an alpha of .83 in a sample of 265 male offenders entering a cognitive skills program. The items produced factor loadings in the very good to excellent range (.60 to .77) with a mean of .67.” (Serin et al., 2003 p. 3).

Problem Recognition assesses an offenders’ appraisal of their current situation, specifically their understanding and ownership of their problems. This item is scored from 0 to 3. A score of “0” indicates that a person denies responsibility and accepts no ownership (e.g., “views the problem is solely the result of other circumstances”), a score of “1” indicates that a person accepts marginal ownership (e.g., “views the problem as mainly the result of others or circumstances”), a score of “2” indicates that a person accepts some ownership (e.g., “views self as a part of the problem”), and a score of “3” indicates that a person accepts full responsibility without rationalization and accepts ownership (e.g., “views self as the major part of the problem”).
Macro Treatment Benefits measures an offender’s views regarding the overall benefits of their participation in treatment. Offenders who see no benefits of treatment or who are unable to generate any benefits from participating in treatment receive a score of “0”. Those who are able to identify at least one long term and short-term benefit of treatment receive a score of “1”. Individuals who “consider limited long-term and short-term benefits of treatment” receive a score of “2” and those who “accurately consider long term and short-term benefits of treatment” receive a score of “3” (Serin et al., 2002, p. 6).

Micro Treatment Benefits measures offenders’ general views regarding treatment. Individuals who cannot identify any benefits or are unable to perceive benefits of treatment receive a score of “0”. Those who perceived treatment as beneficial only for themselves receive a score of “1”. Individuals who perceived treatment as beneficial for themselves or others receive a score of “2” and those who perceived treatment as beneficial for themselves and others receive a score of “3”.

Treatment Distress measures offenders’ state of emotional distress regarding treatment. Individuals who are indifferent (e.g., lack emotional distress) and see no need for treatment receive a score “0”. Individuals who are distressed, but such distress does not motivate them to consider change receive a score of “1”. Individuals who have distress that “motivates them to consider changing” receive a score of “2” (Serin et al., 2002, p. 7) and lastly, individuals who demonstrate “evidence of emotional distress” and “want to participate in treatment” receive a score of “3” (Serin et al, 2002, p. 7).

Treatment Goals assesses the ability to identify and realistically create treatment goals and “considers the knowledge and skills necessary for treatment gain” (Serin et al., 2002, p. 8). A score of “0” is given to those who are “unable to set realistic treatment goals,” a score of “1” is
assigned to those who are “unaware of skills and knowledge required for treatment gain,” a score of “2” is assigned for those who are “somewhat able to set realistic treatment goals,” and lastly, a score of “3” is assigned to those who are “able to set realistic treatment goals” (Serin et al., 2002, p. 8).

Treatment Behaviors assesses offenders’ motivation for treatment through behavioral indicators such as “timely attendance at interviews and/or groups; homework completion; compliance with prior treatment” or positive comments about treatment as a process not an outcome” where applicable. A score of “0” is given if offenders demonstrate “consistent behavioral indication of poor motivation,” a score of “1” is given when there is “inconsistent indication of good motivation,” a score of “2” is given when there is “somewhat inconsistent indication of good motivation,” and a score of “3” is given when there is “consistent behavioral indication of good motivation” (Serin et al., 2002, p. 9).

Behavioral Congruency emphasizes “the importance of an offender’s verbal statements and their actions regarding treatment” (Serin et al., 2002, p. 9). A score of “0” is given when an offender’s “verbal and behavioral expressions of motivation are inconsistent.” A score of “1” is given when an offender is “often inconsistent between stated motivations and actions.” A score of “2” is given to those who are “somewhat inconsistent between stated motivation and actions” and lastly, a score of “3” is given to those who demonstrate “complete congruence between verbal and nonverbal expressions of good motivation” (Serin et al., 2002, p. 9).

Treatment Support evaluates the degree of support for change an offender receives from others who are “significant to them,” as designated by the offender, “preferably family, friends, employer[s] or clergy” (Serin et al., 2002, p. 9). Individuals who report “no external support for changing” are scored with a “0”, those who report “minimal external support for changing” are
scored with a “1”, those who report “moderate external support for changing” are scored a “2”, and those who report “strong external support for changing” are scored with a “3”. Individual scores for each of the eight categories are summed in order to provide a total score that represents an individual’s readiness for treatment; scores on this item range from 0 to 24. Higher scores on this scale reflect greater readiness for treatment while lower scores indicate less readiness for treatment. (Fishbein & Sheppard, 2006, p. 81).

Treatment Performance Indicators

As noted by Fishbein & Sheppard (2006), when evaluating efficacy of treatment programming provided to antisocial populations/offenders, the validity of using self-report measures, particularly among violent offenders, is questionable (cf., Novaco, 1994; Bellemare & McKay, 1992; Hughes, 1993). The original study primarily utilized “behavioral performance measures, taken at baseline and post-intervention, which [did] not rely solely on self-reports” (Fishbein & Sheppard, 2006, p. 23). In addition, the study also collected evaluations of offenders conducted by social workers.

Treatment Responsivity: Short Scale. The Treatment Responsivity: Short Scale (along with the Treatment Gain Scale) was developed by Ralph Serin (former director of Programs Research, Correctional Service of Canada) as an indicator of offender’s treatment performance (Fishbein & Sheppard, 2006, p. 23). This instrument was later adopted by the Wisconsin Department of Corrections for use as an evaluation of their cognitive program. Both the Treatment Responsivity and Treatment Gain Scales were designed to assess CBT performance (Kennedy & Serin, 1997, 1999; Serin, 1998) by a variety of staff and across a range of programs. The Treatment Responsivity scale includes eight items “selected to represent potential
responsivity factors in offender’s compliance with, and response to, therapeutic intervention and treatment programs in general” (Fishbein & Sheppard, 2006, p. 82).

Items were chosen to tap into offender’s general interpersonal style “and are not necessarily specific to treatment” (Fishbein & Sheppard, 2006, p. 82). Specifically, this scale measures: callousness, denial, procrastination, intimidation, power and control, rigidity, victim stance, and procriminal views.

Callousness assesses whether offenders appreciate the impact their antisocial conduct has had on others. This item deems callous offenders as those “who have no concept of the injury they have caused others” and “lack concern for others except when it can serve them” (Fishbein & Sheppard, 2006, p. 89). Individuals who “us[e] people to meet [their] own needs” receive a “0” on this item. Those who are “indifferent about the needs of others” receive a score of “1”. Individuals who are “willing to “consider the needs of family or close friends” receive a score of “2” and lastly, those who take “others’ needs into consideration” receive a score of “3” (Fishbein & Sheppard, 2006, p. 89).

Denial measures the extent to which offenders rationalize their criminal behavior. Individuals who deny their problems (e.g., “it’s everyone else’s fault”) and provide excuses for their behavior whether they are external reasons (e.g., drugs, alcohol, social pressure) or internal reasons (e.g., difficult childhood, previous victimization, mental illness) score a “0” on this item. Those who “refus[e] to accept they have a problem” receive a “1” for this item. Individuals who “accep[t] they have a problem, with reservations” score a “2” on item. A score of “3” is assigned to those who “assum[e] responsibility.”

Procrastination measures offenders’ ability to set and meet goals in general. However, this item notes that an individual’s level of resistance, unwillingness to do homework, and
pattern of making excuses for failing to meet obligations should also be considered (Fishbein & Sheppard, 2006, p. 89). Individuals who demonstrate a “lack of effort, inability to follow through on plans” and lack of goals receive a score of “0”. A score of “1” is given to those who “rarely” follow through on plans while a score of “2” is given to those who “occasionally” follow through on plans. A score of “3” is given to those who are “very task oriented.”

Intimidation measures the level of intensity and expression of anger offenders express in interpersonal situations; often, offender’s emotional expressions “of anger [are] excessive for the situation” demonstrating “both an inability to evaluate the situation and poor self-control” (Fishbein & Sheppard, 2006, p. 90). Individuals who use “over expression[s] of anger to control and manipulate others” receive a score of “0” on this item. Those who are “willing to let anger help them meet their goals” receive a score of “1”. Individuals who are “are and concerned about negative impact of [their] anger on others” receive a score of “2” and those who do not “intimate others” receive a score of “3”.

Power and Control addresses the degree to which offenders express a sense of entitlement when dealing with others. This item considers whether an offender’s concept of fairness is “solely egocentric” and if “they respond poor to criticism” or believe “they must win at all costs.” (Fishbein & Sheppard, 2006, p. 90). Those who perceive “life as unfair and feel they own others” score a “0” on this item. Those who view demonstrate the view “life is unfair, so take what you can” receive a score of “1”. Those who demonstrate the view that “life is unfair, look out for yourself” receive a score of “2”. Finally, those who try “to be fair in resolving disputes” receive a score of “3” for this item.

Rigidity measures offenders’ ability to problem solve effectively. Those who are “rigid” and “stic[k] with a solution, even when it doesn’t work” are scored “0” on this item. Those who
start “with an old solution, but can evaluate” receive a score of “1”. A score of “2” is given to those who consider “new solution[s], but fall[1] back on old ways” and a score of “3” is assigned to those who are “flexible” and demonstrate they are “willing to try other things” (Fishbein & Sheppard, 2006, p. 91).

Victim Stance identifies offenders who “are characterized by self-pity and present as being victims” (Fishbein & Sheppard, 2006, p. 91). For this item, individuals who “appear unwilling to accept their culpability and look to others for support and to improve their situation” receive a score of “0”. A score of “1” is assigned to those who “just want things to be better” and a score of “2” is assigned to those who are “willing to accept consequences for prior behaviour” (Fishbein & Sheppard, 2006, p. 91). A score of “3” is assigned to offender’s who “accep[t] consequences and learn from them.” (Fishbein & Sheppard, 2006, p. 91).

Procriminal Views was designed “to distinguish those offenders whose investment in crime is high from those who are essentially prosocial but who have infrequently committed a crime” (Fishbein & Sheppard, 2006, p. 91). A score of “0” is given to those who demonstrate “pride in criminal views.” A score of “1” is assigned to those for whom “criminal views [are] present, but mainly due to lifestyle.” Individuals for whom “some prosocial views are noted” receive a score of “2” and a score of “3” is given to those who present “mainly prosocial views.” (Fishbein & Sheppard, 2006, p. 91). For this scale, scores on individual items are summed in order to provide a total score that represents an offender’s responsivity to treatment. Scores on this scale range from 0 to 24 with lower scores indicating less responsivity to treatment and higher scores indicating greater responsivity to treatment. This scale has demonstrated excellent internal consistency (alpha of .82) in a sample of male offenders entering a cognitive skills program (Fishbein & Sheppard, 2006, p. 82).
**Treatment Gain: Short Scale.** The Treatment gain includes items that “represent a combination of knowledge, participation and competencies” (Fishbein and Sheppard, 2006, p. 82). Although Sheppard and Fishbein (2006) note that many programs include specific measures to evaluate potential gains for “particular program targets, the purpose of this domain [was] to provide an overall estimate of an offender’s performance in a correctional program” (Fishbein & Sheppard, 2006, p. 82). At the time of data collection, no current data existed “to support the predictive validity of this domain” . . . however, the items included were believed to “provide a useful and defensive overview to utilize in a post-treatment report” (Fishbein & Sheppard, 2006, p. 82).

Evidence of Increase Skills From Program measures the specific skills a treatment program is intended to impart to an offender. Scores on this item range from 0 (none) to 3 (shows/reports successful skill use in other situations), with higher scores indicating an “offender can do more than simply repeat in group that which has been demonstrated in previous group sessions” (Fishbein & Sheppard, 2006, p. 94).

Disclosure in Program measures the extent to which offenders share information in the context of treatment. Acknowledging that offenders differ with respect to their level of comfort or willingness to share in group versus individual sessions, disclosures in both contexts are considered when scoring this item. Individuals who are “resistant” (e.g., deny, refuse to participate, are obstructionist) receive a score of “0”. Those who are deemed “marginal” (e.g., “uncommunicative”) receive a score of “1”, while those who are “satisfactory” (e.g., open up in group sessions) receive a score of “2”. A score of “3” is reserved for individuals who demonstrate “full” disclosure (e.g., those who are candid, reveal extra information in group sessions) (Fishbein & Sheppard, 2006, p. 94).
Application of Knowledge “considers the extent to which an offender is able to consider and apply knowledge from [a] program to his or her own situation,” and not only to other group members (Fishbein & Sheppard, 2006, p. 94). Scores on this item range from 0 (e.g., *poor application of program knowledge, is unable to apply program knowledge*) to 3 (e.g., *able to be reflective and problem-solve in many situations, is insightful*). A score of “3” is “reserved for those offenders who [are] able to apply the information [in] a reflective and systematic manner” (Fishbein & Sheppard, 2006, p. 94).

Application of Skills measures the range of skills an offender has gained through group participation. Such skills to be evaluated “need not be restricted to role-play situations,” however, such situations could prove to be most convenient for evaluators to consider (Fishbein & Sheppard, 2006, p. 94). Offenders who were unable to apply skills receive a score of “0”. Individuals who were able to role play only as a confederate receive a score of “1”. A score of “2” is assigned to those who can participate as themselves during role play and a score of “3” is assigned to those who apply role-playing skills in other situations (Fishbein & Sheppard, 2006, p. 94).

Depth of Emotional Understanding of Program Content measures the extent to which an offender is emotionally connected to a program’s content and “treatment change requirement” (Fishbein & Sheppard, 2006, p. 95). Scores on this item range from 0 to 3, with a score of “0” assigned to those who “[say] the right words without emotional connectedness” and a score of “3” assigned to those who demonstrate a “very good” level of emotional understanding of program content (e.g., “those who find treatment to be “an emotional but rewarding challenge””) (Fishbein & Sheppard, 2006, p. 95).
Appropriateness of Behavior in Group measures an offender’s ability (or lack thereof) to abide by treatment/program group rules and maintain “personal boundaries of staff and other group members” when such issues/concerns have been brought to the offender’s attention. In context of evaluating this item, it is necessary to differentiate those offenders for whom inappropriateness is due to skill deficits versus those for whom such behavior is due malicious intent (Fishbein & Sheppard, 2006, p. 95).

As with the other items in this scale, this item is scored from 0 to 3, with the highest scores reserved for those who “would be effective role models by challenging peers whose behavior exceeds boundaries” (Fishbein & Sheppard, 2006, p. 95). Specifically, a score of “0” is assigned to those who demonstrate a “poor” level of appropriate behavior in group (e.g., ask intrusive, personal questions of staff, are verbally abusive). A score of “1” is given to those who demonstrate a “marginal” level of appropriate behavior in group (e.g., make sarcastic remarks directed at staff and offenders). A score of “2” is assigned to those who demonstrate a “satisfactory” level of appropriate group behavior (e.g., demonstrate no problems within the group). Lastly, a score of “3” is assigned to those who demonstrate “very good” group behavior (e.g., rebuts inappropriate behaviors of others) (Fishbein & Sheppard, 2006, p. 95).

Participation estimates offenders’ participation over the course of the treatment program, including the offender’s “group behavior, attendance, timeliness of homework completion, and quality of work done” (Fishbein & Sheppard, 2006, p. 95). A score of “0” is assigned to those who demonstrate “poor” participation (e.g., those who demonstrate minimal effort, begrudgingly attended program sessions/meetings etc.) and a score of “1” is assigned to those who demonstrate “marginal” participation (e.g., those who did not actively participate in the program in a positive manner). A score of “2” is assigned to those who demonstrate “satisfactory”
Therapeutic Alliance assesses relationships between offenders and program staff. Fishbein and Sheppard (2006) note that evaluators should not simply rate offenders based on their response to their treatment reports. Rather, this item should evaluate whether “there was any kind of connection or engagement between the offender and therapist” (Fishbein & Sheppard, 2006, p. 95). Scores on this item range from 0 to 3, with high scores being those offenders who “demonstrate an attachment to the therapist, regardless of the nature of challenges throughout the treatment program or the recommendations contained in the final report” (Fishbein & Sheppard, 2006, p. 95). For this item, individuals who demonstrate a “poor” relationship with program staff (e.g., who are confrontation, resistant) receive a score of “0”. Those who demonstrate “marginal” relationship with program staff (e.g., those who demonstrated perfunctory disclosure and still demonstrated us versus them view) receive a score of “1”. Individuals who demonstrate “satisfactory” relationship with program staff (e.g., demonstrated good disclosure and a sense of cooperation with a therapist) receive a score of “2” and those who demonstrate a “very good” relationship with program staff (e.g., those who provided evidence of emotional attachment to the therapist) receive a score of “3” (Fishbein & Sheppard, 2006, p. 96).

As with the other scales included in the Treatment Readiness, Responsivity, and Gain Scale: Short Version, scores on the treatment gain items were summed in order to provide a total score that serves as an indicator of offenders’ performance gains following participation in the
treatment program. Like the other scales in this measure, scores on this scale range from 0 to 24 (eight items on a 4-point scale (e.g., 0 to 3), with higher scores indicating greater gains following treatment and lower scores indicating fewer gains. A score of “0” was assigned to all individuals who did not complete the treatment program.

**Indicators of Treatment Efficacy**

**Institutional Conduct.**

As a proxy for recidivism, changes in institutional misconduct were considered. Although some early research indicated little to no association between prison misconduct and recidivism (see reviews by O’Leary & Glaser, 1972; Glueck & Glueck, 1930; Ohlin 1951), other research does demonstrate evidence that institutional misconduct is associated with recidivism. A review by Hill (1985) found that several studies reported a modest association between institutional misconduct and recidivism. Gottfredson and Adams (2008) found that misconduct was significantly related to parole infractions. In addition, Lattimore, Visher, and Linster (1995) found juveniles who were aggressive during confinement were more likely to recidivate for a violent offense. Moreover, additional studies also provide evidence that institutional misconduct predicts reoffending (see Lattimore, MacDonald, Piquero, Linster, & Vishner, 2004; Trulson, Marquart, Mullings, & Caeti, 2005; Trulson, DeLisi, Caudill, Belshaw, & Marquart, 2010; Spivak & Damphousse, 2006 etc.). More recent research also found that institutional conduct predicts future recidivism among both male and female offenders, however such effects are small (Olson, Stalans, & Escobar, 2016).

**Change in Number of Segregations.** Change in number of segregations was measured by subtracting the number of segregations an offender received while in treatment from the number of segregations they received prior to treatment. Higher scores on this item indicate greater
change in behavior/more improvement in behavior whereas lower scores indicate less change in institutional behavior/conduct.

**Change in Number of Infractions.** Change in number of infractions was measured by subtracting the number of infractions and offender received while in treatment from the number of infractions they committed prior to treatment. For this item, higher scores indicate greater improvement in institutional behavioral/conduct, whereas lower scores indicate less change in an offender’s institutional behavior.

**Change in Aggressive Behaviors and Attitudes.** The Novaco Anger Scale and Provocation Inventory (NAS-PI) was administered to offenders at baseline and again after each of the three treatment groups. As previously discussed, this inventory consists of two parts. Part A, the Novaco Anger Scale assesses cognitive, arousal and behavioral domains and provides information on how one experiences anger. Part B, the Provocation Inventory, provides information on the kinds of situations that induce anger for a particular individual. In the version administered, Part A, the Novaco Anger Scale, consisted of 48-items that described things that “people think, feel and do” and offenders were asked indicate the extent to which these statements were true for them “right now (not in the past)” (Fishbein & Sheppard, 2006; Fishbein & Sheppard, n.d., p. 36). All items in this section were rated on a 3-point scale from 1 (*never true*) to 3 (*always true*). Higher scores indicated greater tendency towards anger reactions.

Part B, the Provocation Inventory is a 25-item self-report instrument that measures anger intensity. This component describes situations that could potentially elicit anger, and then asks respondents to indicate “the amount of anger that [they] would feel if [the event] actually happened to [them].” Respondents rated anger intensity on a 4-point scale of 1 (*not at all angry*), 2 (*a little angry*), 3 (*fairly angry*) and 4 (*very angry*). The Provocation Inventory consists of
provocations concerning five domains: disrespectful treatment, unfairness/injustice, frustration/interruption, annoying traits of others, and irritations. Higher scores indicated higher levels of elicited anger/provocation. The Provocation Inventory is “sensitive to behavioral change particularly in response to an intervention” (Fishbein & Sheppard, 2006, p. 23).

To evaluate changes in offender’s aggressive behaviors and attitudes, scores on the Novaco Anger Scale (NAS) and Provocation Inventory (PI) from baseline to post-treatment were considered. Specifically, scores on the Novaco Anger Scale at follow-up were subtracted from pre-treatment scores (e.g., NAS total score baseline – NAS total score at post-treatment follow-up) to create a change in NAS score. Similarly, Provocation Inventory scores at follow-up were subtracted from pre-treatment scores (e.g., PI total score at baseline – PI total score at post-treatment follow-up) to create a change in PI score.

Table 2. Baseline Test Battery and Measures

<table>
<thead>
<tr>
<th>Variables</th>
<th>Measurement Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>General neuropsychological function (verbal, performance &amp; full scale IQ)</td>
<td>Multidimensional Aptitude Battery (MAB)</td>
</tr>
<tr>
<td>Demographics (age, race/ethnicity, prior drug use etc.)</td>
<td>Background Inventory (adapted from Addiction Severity Index (ASI))</td>
</tr>
<tr>
<td>Psychopathy</td>
<td>Levenson Self-Report Psychopathy Scale (LSRP)</td>
</tr>
<tr>
<td>Executive cognitive performance</td>
<td>Dice Game, Stop-Signal Task, Stroop Interference Task</td>
</tr>
<tr>
<td>Type of Aggression/Aggressive Behavior</td>
<td>Reactive-Proactive Aggression Questionnaire (RPQ)</td>
</tr>
<tr>
<td>Behavioral Control (tendency to be provoked to anger/aggression)</td>
<td>Novaco Anger Scale and Provocation Inventory (NAS-PI)</td>
</tr>
<tr>
<td>Treatment Readiness</td>
<td>Treatment Readiness Scale (Serin, Kennedy &amp; Mailoux, 2002)</td>
</tr>
</tbody>
</table>

Table 3. Baseline Test Battery and Measures

<table>
<thead>
<tr>
<th>Outcome Variables (pre and post-test)</th>
<th>Measurement Instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Behavioral Control</td>
<td>Novaco Anger Scale and Provocation Inventory (NAS-PI)</td>
</tr>
<tr>
<td>Treatment Response and Treatment Gain*</td>
<td>Treatment Responsivity and Gain Scales (scores from Social Workers)</td>
</tr>
<tr>
<td>Institutional Behavior</td>
<td>Offender-Based State Corrections Information System (OBSCIS) data on segregations and infractions</td>
</tr>
<tr>
<td>Executive Decision Making</td>
<td>Dice Game (Sheppard &amp; Fishbein, 2006)</td>
</tr>
<tr>
<td>Processing Speed</td>
<td>Dice Game (Sheppard &amp; Fishbein, 2006)</td>
</tr>
</tbody>
</table>

Note. * = Only administered at post-test (e.g., following completion of treatment program)
Data Processing and Diagnostics

Prior to conducting any hypothesis testing, including inferential statistical tests and path analyses, data were thoroughly examined using IBM SPSS Version 27. As a first step in exploratory data analysis, descriptive statistics were run to assess the overall integrity of the data and to check for distributions of all variables to be analyzed. Descriptive results were used to identify invalid scores (e.g., scores that were below the minimum and above the maximum value for a given item, measure or subscale). Due to the nature of secondary data analyses, I did not have access to the original data and therefore was unable to correct data entry issues, such as the entry of invalid scores. To address this issue, individuals with invalid scores were recoded as having missing data for a given response because it was impossible obtain their “true” score for that item.

After addressing instances of cases with invalid scores, I then created new composite variables by summing the reconstituted versions of variables (e.g., those excluding invalid cases) to create total scores for measures I wanted to assess collectively. This included total scores for: Treatment Responsivity (e.g., sum of scores on eight subscales), Treatment Gain (e.g., sum of scores on eight subscales), NAS-Cognitive Domain (e.g., sum of scores on four subscales), NAS-Arousal Domain (e.g., sum of scores on four subscales), NAS-Behavioral Domain (e.g., sum of scores on four subscales), NAS Total Score (e.g., sum of scores on Cognitive, Arousal, and Behavioral Domains), Provocation Inventory (e.g., sum of scores on five Provocation Inventory subscales).

After addressing these data issues, I then utilized SPSS Explore procedure to assess for the presence of univariable outliers among my variables of interest. Variables with extreme outliers (e.g., outliers that were greater than the 3\textsuperscript{rd} quartile plus 3*interquartile range \textit{or} less than
1st quartile minus 3*interquartile range) we recoded into new variables which excluded extreme outlier cases. Thus, the original variable with all cases was retained.

Among the variables of interest, I found the change for baseline administration of the Dice Game, there were two extreme outliers for reaction time for all trials (e.g., RTAVG), one extreme outlier for percentage of safe choices for entire task (e.g., PERSAFE) and one extreme outlier for chose y with one y face, no neutral bets (e.g., CYWY1NB). For follow-up administration of the Dice Game, there was one extreme outlier for reaction time for all trials (e.g., BRTAVG) and five extreme outliers for chose y with one y face, no neutral bets (e.g., BCYWY1NB).

To create change in executive function scores, I used versions of Dice Game variables that did contain extreme outliers. Individuals who had extreme scores on either baseline or follow-up variables pertaining to a particular Dice Game task element were considered missing for created change scores. In addition, I found three extreme outliers for number of early presses on baseline and tone block trials in the Stop Signal task administered at baseline (e.g., ALLERLY). Scores on this item ranged from 0 to 5, with any score above two (e.g., two early presses) designated as outliers. Due to this restricted range and to retain variation in scores for this item, I chose not to exclude extreme outliers for this variable.

**Evaluation of Missing Data & Data Imputation**

Prior to hypothesis testing, I also conducted an evaluation of missing data because of the large quantity of missing data. First, I omitted seven cases originally included in the data file that contained no data, leaving a remaining sample of 225 participants. Of these 225 participants, 22 were coded as having provided consent only, indicating they had never enrolled in the CBT treatment program.
To assess for patterns in missing data, using SPSS version 27, I conducted a missing value analysis for the split file (e.g., file split by those who had only completed consent vs. those who completed additional assessments), allowing me to compare trends in missing data among individuals who had provided consent only and did not enroll in the CBT program and those who did enroll in the treatment program.

The results of this analysis identified subjects who had provided consent but did not enroll in the CBT program and did complete some baseline measures, such as the NAS-PI. Variables associated with this measure indicated no missing cases (e.g., all 22 participants had values for NAS-PI related variables). However, these individuals generally appear to not have been administered baseline executive function tasks. The percentage of missing data for executive function task related measures, such as the percentage of safe choices for the Dice Game, the median correct Word condition reaction time (Stroop task) etc., ranged from 90.9% missing to 95.5% missing. In addition, among this group, 86.4% had missing data for the Treatment Readiness scale. Due to the high level of missing baseline data among this group, subsequent analyses were conducted using only subjects who had enrolled in the CBT treatment program.

Next, I performed a missing values analysis of baseline measures on remaining cases (e.g., those who had enrolled in the treatment program). Data was not found to be missing completely at random (Little MCAR test $\chi^2 (448, 203) = 545.593, p < .05$. Findings indicated that approximately 30% of subjects did not complete the Stroop Task. For Stroop-related measures, the percentage of cases with missing data ranged from 28.6 to 29.6%. In addition, approximately 15% of subjects did not complete the Stop Signal Task (15.3% of respondents were missing data for Stop Signal Task-related measures, including reaction time for all trials and reaction time for
tone block trials). Moreover, approximately 8% of subjects appeared to not have completed the Dice Game. For Dice Game related-measures percentage of missing values were 8.4%. Recall a small number of cases with extreme outliers were removed from baseline measures, including Dice Game-related measures.

After assessing percentages of missing values across baseline variables, I assessed the data for patterns in missingness. Again, this was done using SPSS’s Missing Value Analysis procedure in SPSS Version 27. The most common pattern was no missing data across all variables. The next most common pattern was for individuals to having missing data for all Stroop Task-related variables (interference score, color condition median reaction time, word condition median reaction time).

To assess whether there were differences among participants who had completed the Stroop task compared to those who had not been administered this task, I created a dummy code variable using the interference score. Individuals who had a score for the Interference variable were assigned a score of “1” (e.g., they had completed the Stroop Task) and a score of “0” was assigned to individuals who were missing interference scores, indicating that they had not completed the Stroop Task. Using this dummy coded variable, I conducted an independent samples t-test to determine whether individuals who completed the Stroop Task differed from those who had not completed the Stroop Task across other baseline measures of interest.

These analyses indicated that there was a significant difference in number of riskiest decisions made between those who completed the Stroop Task versus those who did not complete the Stroop Task. The Levene’s Test for Equality of Variances indicated that the variance in number of riskiest choices of those who had completed the Stroop Task was significantly different from those who had not completed the Stroop Task ($F = 9.595, p = .002$).
Moreover, it appears that there was a significant difference in number of riskiest choices made by those who completed the Stroop Task \((M = .63)\) versus those who had not completed the Stroop Task \((M = .36)\), \(t(112.691) = -2.081, p = .020\), such that those who completed the Stroop Task made more of the riskiest selections during the Dice Game.

Results also indicated that individuals who had completed the Stroop Task differed from those who had not completed the Stroop Task in regards to the extent to which they were impacted by early experiences of physical abuse. Again, the Levene’s Test for Equality of Variances indicated that the variance in functional impairment stemming from childhood physical abuse among those who had completed the Stroop Task was significantly different from those who had not completed the Stroop Task \((F = 18.701, p < .001)\). Results from the independent-samples \(t\)-test indicate that there was a significant difference in the impact of early physical abuse on current functioning among those who completed the Stroop Task \((M = 2.45)\) compared to those who did not complete the Stroop Task \((M = 1.73)\), \(t(85.005) = -3.487, p < .001\). Individuals who completed the Stroop Task demonstrated more impairment stemming from early physical abuse (e.g., impairment across more domains of functioning) than those who did not complete the Stroop Task.

Those who completed the Stroop Task versus those who did not complete the Stroop Task did not seem to differ in scores for all other baseline measures of interest including: age, Stop-Signal Task-related measures, other Dice Game-related measures (e.g., percentage of safe choices, number of risky choices), number of types of childhood abuse experienced, level of impairment stemming from childhood emotional and sexual abuse, NAS-PI scores or RPQ scores.
Due to the high level of respondents who did not complete the Stop Signal Task (15.3% among \( N = 203 \)), the same procedure outlined for the Stroop Task was repeated for the Stop Signal Task. First, I created a dummy code variable to indicate participants who had completed the Stop Signal Task and those who had not completed the Stop Signal Task. A score of “0” was assigned to those who had not completed the Stop Signal Task (e.g., those who were coded as missing for the variable – difference in percentage correct during non-tone(baseline) and tone trials. A score of “1” was assigned to those who had completed the Stop Signal Task (e.g., those who had a score for the aforementioned Stop-Signal measure).

Results indicated that individuals who completed the Stop Signal Task differed from those who did not complete the task in regards to the number of risky decisions made during the baseline administration of Dice Game. The extent of variance in the number of risky decisions made differed among those who had completed the Stop Signal Task as compared to those who had not completed this task (Levene’s Test for Equality of Variances \( F = 5.566, p = .011 \)). Individuals who completed the Stroop Task \( (M = 1.1953) \) made more risky decisions compared to those who did not complete the Stroop Task \( (M = .5582) \), \( t(29.012) = -2.289, p = .030 \).

Similarly, those who completed the Stop Signal Task also differed from those who did not complete the Stop Signal Tasks in number of riskiest decisions made during baseline administration of the Dice Game. The amount of variance in the number of riskiest decisions made differed among those who completed and those who did not complete the Stop Signal Task (Levene’s Test for Equality of Variances \( F = 7.315, p = .007 \)). Again, individuals who completed the Stop Signal Task made more of the riskiest decisions during the Dice Game \( (M = .60) \) than those who did not complete this task \( (M = .30) \), \( t(32.408) = -2.223, p = .033 \).
Finally, those who completed the Stop Signal Task also appeared to differ from those did not complete the Stop Signal Task in the extent to which they experienced impairment associated with early physical abuse. The extent of variance in the degree of functional impairment reported due to early physical abuse differed among those who had completed the Stop Signal Task and those who did not complete this task (Levene’s Test for Equality of Variances, $F = 6.197, p = .014$). Again, as observed among those who had and had not completed the Stroop Task, I found that individuals who completed the Stop Signal Task appeared to experience more functional impairment stemming from childhood physical abuse ($M = 2.34$) than those who did not complete the Stop Signal Task ($M = 1.73$), $t(37.030) = -2.262, p = .030$).

Overall, my analyses indicated the data were missing at random. Missing at random refers when missing data are systematically related to observed data (Mack, Su, & Westreich, 2018). In my sample, I did observe that there were some trends in missing data as individuals who did not complete baseline executive function tasks were more likely to make risky choices during Dice Game and to report more impairment stemming from childhood abuse. However, missing data for these items did not appear to be associated with other baseline measures.

**Data Imputation.** Following my analysis of missing data, I conducted multiple imputation with expectation-maximization (EM) imputation using SPSS version 27 to impute values for missing baseline data. The EM method “estimates the means, the covariance matrix, and the correlation of quantitative (scale) variables with missing values” (IBM, 2021). This method involves an expectation step and a maximization step, which are repeated several times until maximum likelihood estimates are obtained (Soley-Bori, 2013). As outlined by Kang (2013), during the Expectation phase, the parameters (e.g., variances, covariances, means) are estimated. These estimates are then used to create a regression equation that is used to predict the
missing data (Kang, 2013). During the maximization stage, these equations are used to fill in missing data (Kang, 2013). Next, the expectation phase is repeated with new parameters and new regression equations are determined to replace, or fill in, missing data. These steps are repeated until the system has stabilized (e.g., when the covariance matrix for the subsequent iteration is virtually identical to the preceding iteration) (Kang, 2013). EM estimation requires that data are missing at random (MAR) (Soley-Bori, 2013).

There are some drawbacks of this approach, for example, it has been noted that EM-imputed data sets are biased because error is not added to the imputed data set (Graham, Cumsille, & Elek-Fisk, 2003, as cited by Tabachnick & Fidell, 2007). As a result, “analyses based on imputed data sets have inappropriate standard errors for testing hypotheses” (Tabachnick & Fidell, 2007, p. 68). However, it has been suggested that analyses with EM imputed data sets may serve to provide useful insights when the amount of missing data is small as long as inferential statistics are interpreted with caution (Tabachnick & Fidell, 2007).

Regarding what amount of missing data is acceptable, there is no clear consensus. As noted by Dong and Peng (2013), there is no established cutoff from the literature regarding an acceptable percentage of missing data to produce valid statistical inferences. Schafer (1999) argued that a missing rate of five percent or less was inconsequential. More recently, Bennett (2001), suggested that statistical analyses are likely to be biased when more than 10% of the data are missing. However, others have suggested the mechanisms producing missing data and patterns of missing data exert greater influence on results than does the proportion of missing data (Tabachnick & Fidell, 2012, as cited in Dong & Peng, 2013).

In the current analyses, I found that less than 9.6% of the data was missing among baseline variables of interest. This relatively low missing data rate in conjunction with minimal
patterns of missing data prompted my decision to utilize EM imputation to impute data for baseline variables.

This decision was further informed by the fact that the Multiple Imputation module in LISREL “implements the EM algorithm and the Markov Chain Monte Carlo (MCMC) method for imputing missing values in multivariate datasets” (du Toit & du Toit, 2001). As both SPSS and LISREL were used in my analyses and EM estimation is the imputation method employed by LISREL, I chose to complete EM data imputation in SPSS to obtain a final data set with imputed data that would be used for hypothesis testing. This way, the imputed values would be equivalent across program analyses. Note multiple imputation was not used to provide estimates for follow-up data due to the extensive amount of missing data. Only 78 subjects completed follow-up administration of the Dice Game.
CHAPTER FOUR

RESULTS

Hypothesis 1: Baseline EF and Recidivism and Hypothesis 1a: Mediating Effects of Treatment on Recidivism

To test the first hypothesis, does baseline executive functioning predict recidivism, I ran two path analyses. The first path (Path 1a) tested whether the relationship between baseline executive functioning and recidivism was mediated by the number of treatment groups completed. The second path (Path 1b) tested whether the relationship between baseline executive functioning and recidivism was mediated by treatment gains. The number of treatment groups completed and treatment gains (e.g., gains following treatment program as assessed by social workers) were included as mediators in these analyses as it was thought that the ability to benefit from treatment, as evidenced by completion of all treatment groups and by higher treatment gain scores, would be influenced by baseline executive functioning. It was predicted that those with higher baseline executive functioning would be better able to participate in and benefit from the treatment program and in turn, would demonstrate lower recidivism rates compared to individuals with poorer baseline executive functioning.
In these analyses, baseline executive functioning was assessed with four measures:

- response inhibition (e.g., difference in percent correct on non-tone block vs. tone block trials of the Stop Signal Task)
- risky decision-making (e.g., number of risky decisions made during the Dice Game)
- inhibitory control (e.g., Stroop interference score)
- impulsivity (e.g., number of early presses on baseline and tone block trials of Stop Signal Task)

Recidivism was assessed by change in number of infractions following treatment to (e.g., number of infractions before treatment minus number of infractions while in treatment). Note originally I planned to consider changes in both number of infractions and segregations, however, these variables were highly correlated ($r$s ranging from .805 to .819 depending on the sample used- $r = .805$ for sample included in Path 1a and 1b analyses, $r = .819$ for total sample ($N = 225$)).

**Path 1a: Baseline EF, Recidivism and Treatment Group Mediation**

*Descriptive, Correlational, and Collinearity Analyses of Path 1a*

Means, standard deviations, ranges of scores, and correlations among Path 1a variables are presented in Table 5. Note that variables in Path 1a include five continuous variables (e.g., $x_1$, $x_2$, $x_3$, $x_4$, and $y_2$) and one ordinal variable ($y_1$). Thus, the correlations reported include Pearson product-moment correlation coefficients for pairs of continuous variables. For correlations between the ordinal variable ($y_1$ number of treatment groups completed) and the other variables in model 1a, polyserial correlations are reported. Polyserial correlations are used when
estimating correlations between continuous and ordinal variables (see Olsson, Drasgow, & Dorans, 1982). A listing of which correlation coefficient type was used for reported correlations among Path 1a variables is presented in Table 4.

Table 4. Correlation Coefficient Matrix for Path 1a Variables

<table>
<thead>
<tr>
<th>Study variables</th>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( x_3 )</th>
<th>( x_4 )</th>
<th>( y_1 )</th>
<th>( y_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) ( x_1 )</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) ( x_2 )</td>
<td>PE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) ( x_3 )</td>
<td>PE</td>
<td>PE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) ( x_4 )</td>
<td>PE</td>
<td>PE</td>
<td>PE</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of treatment groups completed ( y_1 )</td>
<td>PS</td>
<td>PS</td>
<td>PS</td>
<td>PS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Infractions (CInfract) ( y_2 )</td>
<td>PE</td>
<td>PE</td>
<td>PE</td>
<td>PS</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Note. PE = Pearson Product Moment (two continuous variables), PS = Polyserial (one continuous variable, one ordinal variable), PC = Polychoric (two ordinal variables)

Response inhibition \((x_1)\), risky decision making \((x_2)\), inhibitory control \((x_3)\) and impulsivity \((x_4)\) were all significantly intercorrelated in the expected directions with the exception of response inhibition \((x_1)\) and impulsivity \((x_4)\) which were not significantly correlated \((r = .044, p = .534)\). Two of the predictors were significantly correlated with the mediator \((y_1)\) number of treatment groups completed)-response inhibition \((x_1)\), \( r = -.163, p < .05 \) and impulsivity \((x_4)\), \( r = -.143, p < .05 \). None of the baseline executive functions were found to be significantly correlated with the outcome variable \((y_2)\) change in number of infractions).

It is notable that correlations were slightly stronger among predictor variables and the mediator \((y_1)\) number of treatment groups completed) compared to the correlations between predictor variables and outcome variable \((y_2)\) change in number of infractions) with the exception of risky decision making, which was more strongly correlated with change in number of infractions \((y_2)\), \( r = .129 \). The number of treatment groups completed \((y_1)\) was also found not to
be significantly correlated with change in number of infractions ($y_2$ outcome variable) (see Table 5 for correlations among Path 1a variables).

Table 5. Pearson Correlations, Means and Standard Deviations for Path 1a Variables ($N=198$)

<table>
<thead>
<tr>
<th>Study variables</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$y_1$</th>
<th>$y_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) $x_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $x_2$</td>
<td>.143</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) $x_3$</td>
<td>-.301</td>
<td>-.158</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $x_4$</td>
<td>.044</td>
<td>.151</td>
<td>-.313</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Number of treatment groups completed $y_1$</td>
<td>-.163</td>
<td>.034</td>
<td>.125</td>
<td>-.143</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Infractions (CInfract) $y_2$</td>
<td>-.102</td>
<td>.129</td>
<td>.013</td>
<td>.005</td>
<td>.066</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>.29</td>
<td>1.13</td>
<td>19.61</td>
<td>.62</td>
<td>2.59</td>
<td>2.41</td>
</tr>
<tr>
<td>$SD$</td>
<td>.11</td>
<td>1.66</td>
<td>2.59</td>
<td>1.01</td>
<td>1.35</td>
<td>2.83</td>
</tr>
<tr>
<td>Range</td>
<td>.04-.68</td>
<td>.00-.700</td>
<td>11-26</td>
<td>0-5</td>
<td>1-4</td>
<td>0-17</td>
</tr>
</tbody>
</table>

Note. $SD$ = standard deviation. For $N = 198$, $|r| > .1395$ is statistically significant at two-tailed $p < .05$, $|r| > .1827$ is statistically significant at two-tailed $p < .01$, $|r| > .2322$ is statistically significant at two-tailed $p < .001$.

Before conducting path analysis, I examined the degree of multicollinearity among the four predictor variables ($x_1$-$x_4$) when using them to predict number of treatment groups completed ($y_1$ mediator) and change in number of infractions ($y_2$ outcome) in multiple regressions using IBM SPSS Statistics, Version 27. As evidence that the predictors are distinct and unique from one another, collinearity diagnostics revealed relatively large variable tolerances for: response inhibition $x_1$ (variance inflation factor (VIF) = 1.116, TI = .896), risky decision making $x_2$ (VIF = 1.050, TI = .952), inhibitory control $x_3$ (VIF = 1.228, TI = .814), and impulsivity $x_4$ (VIF = 1.127, TI = .888).

As noted by Marcoulides and Raykov (2019) “informal, rough “rules-of-thumb” suggest that a predictor, $X_j$, with $V_j > 10$ or $T_j < 0.10$ values, may well be a cause of serious (near) multicollinearity” (p. 876). However, others have proposed more conservative informal threshold criteria (Marcoulides & Raykov, 2019). For example, predictors with VIF values greater than
five or tolerance indexes (TI) less than 0.20 could also contribute considerably to
collinearity (Chatterjee & Simonoff, 2013; O’Brien, 2007) and generally deserve close
inspection (Marcoulides & Raykov, 2019). (Note tolerance index (TI), which is the inverse of
VIF (e.g., equal to 1 - $R^2$), is another indicator of collinearity in which values range from 0
to one with high tolerance values closer to 1 indicating lower collinearity (Daoud, 2017;
Hair, Black, Babin, & Anderson, 2010). In this case, the low VIF values and high TIs observed
among predictor variables indicate sizable proportions of variance in each predictor that cannot
be explained by other predictors included in the model (Cohen, Cohen, West, & Aiken, 2003).

**Path Analysis (Path 1a)**

I tested the hypothesis that number of treatment groups completed ($y_1$) mediates the
impact of baseline executive abilities (e.g., response inhibition, risky decision making, inhibitory
control and impulsivity) on recidivism (e.g., change in number of infractions). As I previously
discussed, various executive deficits have been associated with higher rates of recidivism (see
Langevin & Curnoe, 2011; Ross & Hoaken, 2011; Sánchez de Ribera, Trajtenberg, & Cook,
2021; Hancock, Tapscott, & Hoaken, 2010) and with poorer treatment-related outcomes among
at-risk populations, including individuals with substance use disorders (Worley, Tate, Granholm,
& Brown, 2014). Therefore, I controlled for the effects of baseline executive abilities by
including these variables (e.g., $x_1$-$x_4$) as covariates in my path analyses and also by specifying
each of these baseline executive abilities as predictors of both the number of treatment groups
completed (e.g. $y_1$ mediator) and the outcome variable, change in number of infractions (e.g., $y_2$
outcome), in mediational analyses.

As I did not expect the indirect effects in the model to fully mediate the impact of
baseline executive abilities on recidivism (note prior research has demonstrated evidence of
several other individual and community-level risk factors associated with recidivism (see Stahler et al., 2013; Håkansson & Berglund, 2012; Molina-Coloma, Salaberría, & Pérez, 2021), the path model employed included the direct effects of each of the baseline executive functions on both the number of treatment groups completed ($y_1$ mediator) and change in number of infractions ($y_2$ outcome). The residual variances for endogenous $y$ variables in the path model (e.g., $y_1$ and $y_2$) were not allowed to correlate and thus this model produced independent residual variances for each both $y_1$ and $y_2$. As previously reported, the correlation between $y_1$ and $y_2$ was quite small ($r = .066$). Therefore, I decided that specifying independent residual variances, rather than correlated residual variances, among $y$ variables was most appropriate.

To estimate the hypothesized path model, I conducted structural equational modeling analyses using LISREL 8.8 (Jöreskog & Sörbom, 1996) with diagonally weighted least squares (DWLS) estimation based on the asymptotic covariance matrix derived from the final sample retained post-data cleaning and processing ($N = 198$). In assessing the goodness-of-fit of this proposed path model, the model is exactly identified (e.g., the number of free parameters exactly equals the number of known values, the number of knowns exactly equals the number of unknowns (Kenny, 2011) and therefore produces a perfect fit to the data (Bollen, 1989; Kline, 2011). Consequently, goodness-of-fit statistics for this “saturated” model are irrelevant; by definition, $\chi^2 = 0$, $df = 0$ for such models. Regardless, the model provides a viable means of testing research hypotheses by estimating: a) standardized path coefficients ($\beta$s), standard errors ($SE$s), and their $p$ values, b) total, direct, and indirect effects for exogenous predictor variables, as well as tests of mediation, and c) the proportion of variance explained in each endogenous outcome variable. (Note DWLS estimated a standardized, polychoric matrix, therefore
standardized and unstandardized values are equivalent when using DWLS estimation as the input matrix analyzed was already in standardized form).

To evaluate potential mediation in greater depth, I also assessed the strength and statistical significance of indirect effects. To compare the strength of indirect effects, I used Hayes’ PROCESS macro Version 4.0 (Hayes, 2013) to run mediational analyses and conduct bootstrap tests of indirect effects to examine number of treatment groups completed as a potential mediator of the effects of baseline executive functioning on recidivism. Specifically, utilizing the Hayes macro, mediational analyses were conducted to evaluate the indirect effects of each of the four baseline executive function variables: response inhibition, risky decision making, inhibitory control (Stroop interference score), and impulsivity on recidivism through number of treatment groups.

To do this, four multiple regression models were run, each one including one baseline executive function with the remaining three baseline executive functions included as covariates. For each model, the baseline executive function being evaluated and covariates (e.g., remaining baseline executive functions) were modeled as having direct effects on both the mediator (e.g., number of treatment groups) and the dependent variable in the regression model (e.g., recidivism).

The bootstrap bias-corrected 95% confidence interval (CI) (CI based on 10,000 random samples with replacement) was used to assess the significance of mediation as this approach adjusts for skewness in the sampling distribution of indirect effects and provides greater statistical power to detect indirect effects as compared to alternative methods for assessing statistical significance of mediation (Fritz & MacKinnon, 2007; Shrout & Bolger, 2002, see Bryant, Osowski, & Smith, 2021). In the analyses, indirect effects whose 95% bootstrap CI
excluded zero were considered statistically significantly different from 0 at two-tailed $p < .05$. Indirect effects whose 95% bootstrap CI included zero were deemed non-statistically significantly different from zero (Hayes, 2013).

The Sobel test (Sobel, 1982) has been used to assess the significance of indirect, or mediated, effects. However, mediation experts have advised against this procedure (e.g., Hayes, 2009; Preacher & Hayes, 2008) as this test assumes the sampling distribution of indirect effects is normal, which is unlikely (Shrout & Bolger, 2002). Again, the Sobel test is based on faulty assumptions regarding $p$ values. Therefore, the Sobel test produces biased results that are prone to Type II errors (Fritz & McKinnon, 2007). To address these concerns, bootstrapping is advised, with this approach being recommended as a more accurate alternative strategy that provides greater power for assessing the statistical significance of mediation (MacKinnon, Lockwood, & Williams, 2004; Mallinckrodt, Abraham, Wei, & Russell, 2006).

The sample size in the present study was relatively small ($N=198$). To achieve sufficient (i.e., 80% power) assuming the effects of baseline factors on mediators (number of treatment groups completed and treatment gain) and the effects of mediators on recidivism are small/weak (i.e., $\beta = .14$), the Sobel test would require a sample size of 667, whereas the bias-corrected bootstrap procedure would require a smaller sample size of 462. Assuming the effects of baseline factors on mediators and the effects of mediators on recidivism are moderate, the Sobel test would require a sample size of 196, whereas the bias-correct bootstrap procedure would require a sample size of only 148 to achieve sufficient power to detect effects (see Fritz & MacKinnon, 2007, Table 3 p. 237). As our greatest sample size among all mediation analyses tested is $N = 198$ (e.g., Models 1a and 1b), the bias-corrected bootstrap procedure was further preferred over
the Sobel test as it provides greater power to detect effects. (Note despite recommendations against using the Sobel test, Sobel test results are reported in Table 7 for readers’ interest).

Figure 1. Path Diagram of the Conceptual Model Tested in the Present Research (Model 1a)

**Direct Effects in Path 1a.** Figure 1 displays the proposed path for Model 1a, which again predicted that the relationship between baseline executive abilities and recidivism would be mediated by the number of treatment groups’ subjects completed. The structural relationships in this model were predicted a priori on the basis of existing evidence as discussed in the review of the literature.

Overall, the baseline executive abilities included in Path 1a were not found to be predictive of the number of treatment groups completed ($y_1$). Among the four included executive functions, only response inhibition ($x_1$) was found to be a significant predictor of number of treatment groups completed ($\beta = -.15$, $SE = .07$, $p = .0238$). Thus, when controlling for other baseline executive abilities (e.g., $x_2 - x_4$), a 1 $SD$ increase in response inhibition (e.g., greater response inhibition score, indicating worse capacity for response inhibition) will result in a .15 $SD$ decrease in number of treatment groups completed ($y_1$). Impulsivity (# of early presses, $x_4$)
was not found to be a significant predictor of number of treatment groups completed; however, this effect approached significance ($\beta = -0.13$, $SE = 0.07$, $p = 0.0549$). Risky decision making ($x_2$) and inhibitory control ($x_3$) were not found to predict the number of treatment groups completed. (See Table 6 for complete listing of standardized path coefficients for direct effects in Model 1a).

Overall, baseline executive abilities (e.g., $x_1 - x_4$) were found to explain little variance in the number of treatment groups completed ($R^2 = 0.05338$). Taken together, response inhibition, risky decision making, inhibitory control and impulsivity were found to account for only 5.34% of the variation in number of treatment groups completed.

Baseline executive abilities included in Model 1a were also not found to be predictive of change in number of infractions ($y_2$). (See Table 6 for complete list of parameter estimates for direct effects in Model 1a). Together, baseline executive abilities accounted for only about three percent of the variation in change in number of infractions ($R^2 = 0.03319$), indicating these variables were less predictive of recidivism than of completion of treatment groups. Lastly, number of treatment groups completed ($y_1$) was also not found to predict change in number of infractions ($y_2$) ($\beta = 0.04$, $SE = 0.07$, $p = 0.5755$).

Table 6. Direct Effects in Model 1a (LISREL parameter estimates)

<table>
<thead>
<tr>
<th>Paths</th>
<th>$\beta$</th>
<th>$SE$</th>
<th>Z-value</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) $x_1$ $\rightarrow$ Number of treatment groups completed $y_1$</td>
<td>-0.15</td>
<td>0.07</td>
<td>-2.26</td>
<td>0.0238*</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $x_2$ $\rightarrow$ Number of treatment groups completed $y_1$</td>
<td>0.08</td>
<td>0.06</td>
<td>1.31</td>
<td>0.1902</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) $x_3$ $\rightarrow$ Number of treatment groups completed $y_1$</td>
<td>0.05</td>
<td>0.08</td>
<td>0.67</td>
<td>0.5029</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $x_4$ $\rightarrow$ Number of treatment groups completed $y_1$</td>
<td>-0.13</td>
<td>0.07</td>
<td>-1.92</td>
<td>0.0549</td>
</tr>
<tr>
<td>Response Inhibition (DPerCor) $x_1$ $\rightarrow$ Change in Number of Infractions $y_2$</td>
<td>-0.12</td>
<td>0.18</td>
<td>-0.67</td>
<td>0.5029</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $x_2$ $\rightarrow$ Change in Number of Infractions $y_2$</td>
<td>0.14</td>
<td>0.40</td>
<td>0.36</td>
<td>0.7188</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) $x_3$ $\rightarrow$ Change in Number of Infractions $y_2$</td>
<td>-0.01</td>
<td>0.64</td>
<td>-0.01</td>
<td>0.9920</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $x_4$ $\rightarrow$ Change in Number of Infractions $y_2$</td>
<td>-0.01</td>
<td>0.22</td>
<td>-0.04</td>
<td>0.9681</td>
</tr>
<tr>
<td>Number of treatment groups completed $y_1$ $\rightarrow$ Change in Number of Infractions $y_2$</td>
<td>0.04</td>
<td>0.07</td>
<td>0.56</td>
<td>0.5755</td>
</tr>
</tbody>
</table>

Note. $\beta =$ standardized path coefficients from LISREL solution. $SE =$ standard error.
**Indirect Effects in Path 1a/Test of Mediation.** The PROCESS macro for SPSS, version 4 (Hayes, 2013) was used to generate estimates for path coefficients as well as bootstrap CIs for total and specific indirect effects of exogenous predictor variables on the endogenous outcome variable (Preacher & Hayes, 2008).

Tests of mediation indicated that the number of treatment groups completed did not mediate the relationships between any baseline executive ability (e.g., $x_1 - x_4$) and change in number of infractions ($y_2$). Bootstrap analyses revealed the indirect effects of baseline executive abilities via number of treatment groups on change in number of infractions were not significant; the 95% bias-corrected bootstrap CIs obtained for each predictor (e.g., $x_1 - x_4$) included zero. Therefore, none of the four tested indirect effects were found to be statistically significant (Tryon, 2001). (See Table 7 for standardized path coefficients for indirect effects in Model 1a.)

In addition to bootstrap mediation analyses, potential mediation was also assessed using the Sobel test (Sobel, 1982). Sobel test results matched the previous findings that indicated that the number of treatment groups completed does not mediate the relationships between baseline executive abilities (e.g., $x_1 - x_4$) and change in number of infractions ($y_2$). (See Table 7 for Sobel test results).
Table 7. Indirect Effects in Model 1a (LISREL parameter estimates)

<table>
<thead>
<tr>
<th>Paths</th>
<th>Sobel Test</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) $$x_1 \rightarrow$$ Change in Number of Infractions $$y_2$$</td>
<td>$$-0.01$$</td>
<td>0.01</td>
<td>-0.56</td>
<td>0.5755</td>
<td>-0.6080, 0.4069</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $$x_2 \rightarrow$$ Change in Number of Infractions $$y_2$$</td>
<td>0.00</td>
<td>0.01</td>
<td>0.57</td>
<td>0.5687</td>
<td>-0.0211, 0.0288</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) $$x_3 \rightarrow$$ Change in Number of Infractions $$y_2$$</td>
<td>0.00</td>
<td>0.00</td>
<td>0.43</td>
<td>0.6672</td>
<td>-0.0122, 0.0197</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $$x_4 \rightarrow$$ Change in Number of Infractions $$y_2$$</td>
<td>$$-0.01$$</td>
<td>0.01</td>
<td>-0.54</td>
<td>0.5892</td>
<td>-0.0775, 0.0468</td>
</tr>
</tbody>
</table>

Note. $$\beta$$ = standardized path coefficients from LISREL solution. SE = standard error. CI = bias-corrected bootstrap confidence interval, 95% two-tailed obtained via Hayes’ PROCESS Macro for SPSS, version 4 (Hayes, 2013) with 10,000 bootstrap samples.

Path 1b: Baseline EF Predicting Recidivism Mediated by Treatment Gain

Descriptive, Correlational, Collinearity Analyses of Path 1b

As I previously discussed, response inhibition ($$x_1$$), risky decision making ($$x_2$$), inhibitory control ($$x_3$$), and impulsivity ($$x_4$$) were all significantly intercorrelated in the expected directions with the exception of response inhibition ($$x_1$$) and impulsivity ($$x_4$$) which were not significantly correlated ($$r = 0.044$$, $$p = .534$$). Moreover, as I previously reported, none of the baseline executive abilities were found to be correlated with change in number of infractions ($$y_2$$).

Only one baseline executive ability—response inhibition ($$x_1$$)—was significantly correlated with the mediator ($$y_1$$ treatment gain score) ($$r = -0.192$$, $$p = .007$$). The correlation between response inhibition and treatment gain score was in the expected direction, such that better response inhibition was associated with greater treatment gains. Furthermore, I found that treatment gain scores were not significantly correlated with change in number of infractions ($$r = 0.056$$, $$p = .434$$). (See Table 8 for correlations among Path 1b variables).
As outlined in the discussion of Path 1a, prior to conducting path analysis for Path 1b, I examined the degree of multicollinearity among the four predictor variables ($x_1$-$x_4$) when using them to predict treatment gain score ($y_1$ mediator) and change in number of infractions ($y_2$ outcome) in multiple regression using IBM SPSS Statistics, version 27. These values are identical to those found in Path 1a. Again, the low VIF values and high TIs observed among predictor variables indicate sizable proportions of variance in each predictor that cannot be explained by other predictors included in the model (Cohen et al., 2003).

Path Analysis (Path 1b)

To test the hypothesis that treatment gain score ($y_1$) mediates the impact of baseline executive abilities (e.g., response inhibition, risky decision making, inhibitory control and impulsivity) on recidivism (e.g., change in number of infractions), the steps outlined for analysis of Path 1a were once again employed. However, instead of using DWLS estimation as was done in Path 1a, Path 1b analyses were completed using maximum likelihood (ML) estimation based on the asymptotic covariance matrix derived from the final sample retained post-data cleaning.

Table 8. Pearson Correlations, Means and Standard Deviations for Path 1b Variables ($N = 198$)

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$y_1$</th>
<th>$y_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) $x_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $x_2$</td>
<td>.143</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) $x_3$</td>
<td>-.301</td>
<td>-.158</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $x_4$</td>
<td>.044</td>
<td>.151</td>
<td>-.313</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Treatment Gain Score $y_1$</td>
<td>-.192</td>
<td>.058</td>
<td>.116</td>
<td>-.108</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Infractions $y_2$</td>
<td>-.102</td>
<td>.129</td>
<td>.013</td>
<td>.005</td>
<td>.056</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>.29</td>
<td>1.13</td>
<td>19.61</td>
<td>.62</td>
<td>11.35</td>
<td>2.41</td>
</tr>
<tr>
<td>SD</td>
<td>.11</td>
<td>1.66</td>
<td>2.59</td>
<td>1.01</td>
<td>8.67</td>
<td>2.83</td>
</tr>
<tr>
<td>Range</td>
<td>.04-.68</td>
<td>.001-7.00</td>
<td>11-26</td>
<td>0-5</td>
<td>0-24</td>
<td>0-17</td>
</tr>
</tbody>
</table>

Note. SD = standard deviation. For $N = 198$, $|r| > .1395$ is statistically significant at two-tailed $p < .05$, $|r| > .1827$ is statistically significant at two-tailed $p < .01$, $|r| > .2322$ is statistically significant at two-tailed $p < .001$. 


and processing \((N = 198)\). The switch to ML estimation was due to the fact that Path 1b contained no ordinal variables.

Once again, structural equational modeling analyses were completed using LISREL 8.8 (Jöreskog & Sörbom, 1996). Like Model 1a, Model 1b was exactly identified and thus produced a perfect fit to the data (Bollen, 1989; Kline, 2011). Therefore, goodness-of-fit statistics for Model 1b are not relevant. Structural equation modeling did produce estimates of: 1) standardized \((\beta)\) and unstandardized \((b)\) path coefficients, standard errors \((SEs)\), and their \(p\) values, 2) total, direct and indirect effects for exogenous predictor variables, in addition to tests of mediation, and 3) the proportion of variance explained in each endogenous outcome variable. (See Figure 2 for depiction of conceptual model of Path 1b).

**Figure 2. Path Diagram of the Conceptual Model Tested in the Present Research (Model 1b)**

\begin{center}
\begin{tabular}{c}
\textbf{\(x_1\) Response Inhibition (DPerCor)} \\
\textbf{\(x_2\) Risky Decision Making (CYY1NNB)} \\
\textbf{\(x_3\) Inhibitory Control (Interference Score)} \\
\textbf{\(x_4\) Impulsivity (# of early presses)} \\
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{c}
\textbf{\(y_1\) Treatment Gain Score (TxGain)} \\
\textbf{\(y_2\) Change in Number of Infractions (Cinfract)} \\
\end{tabular}
\end{center}

To test this hypothesis, I controlled for the effects of baseline executive abilities by including these variables as covariates in path analyses. In addition, baseline executive functions were included as predictors of both the mediator (e.g., \(y_1\) treatment gain score) and the outcome variable (\(y_2\) change in number of infractions) in both the path analysis and mediational analysis of Model 1b.
To assess for mediation in Model 1b, the Hayes PROCESS macro Version 4.0 (Hayes, 2013) was used to run mediational analyses and conduct bootstrap tests of indirect effects. Thereby, allowing me to examine treatment gain score as a potential mediator of the effects of baseline executive functioning on recidivism. Specifically, these mediational analyses evaluated for potential indirect effects of each of the four baseline executive functions (e.g., response inhibition, risky decision making, inhibitory control, and impulsivity) on recidivism through treatment gain scores.

As in Model 1a mediational testing, four multiple regression models were run, each one including a single baseline executive function with the remaining executive functions included as covariates. For each multiple regression, the baseline executive function being evaluated and the covariates (e.g., remaining three baseline executive functions) were modeled as having direct effects on both the mediator (e.g., treatment gain score $y_1$) and the dependent variable, recidivism (e.g., $y_2$ change in number of infractions). Again bootstrap, bias-corrected 95% confidence intervals (based on 10,000 random samples with replacement) were used to assess the significance of mediation. Indirect effects whose 95% bootstrap CIs included zero were deemed non-statistically significantly different from zero at two-tailed $p < .05$ (Hayes, 2013).

**Direct Effects in Path 1b.** Mirroring results obtained for Path 1a, baseline executive abilities were generally not found to be predictive of treatment gain scores ($y_1$). Among the four executive functions included as predictors, only response inhibition ($x_1$) was found to significantly predict treatment gain score ($y_1$), ($\beta = -0.19$, $SE = 5.98$, $p = .0164$). Therefore, when controlling for the other baseline executive functions (e.g., $x_2 - x_4$), a 1 $SD$ increase in response inhibition will produce a .19 $SD$ decrease in treatment gain score. Recall, greater response inhibition scores indicated worse response inhibition. Therefore, this significant direct effect
suggests a 1 SD increase in response inhibition (e.g., a 1 SD reduction in capacity for response inhibition) will result in a .19 SD decrease in treatment gain score. Overall, these four baseline executive abilities (e.g., \( x_1 - x_4 \)) were found to explain little variance in treatment gain scores (\( R^2 = .05889 \)); taken together, response inhibition (\( x_1 \)), risky decision making (\( x_2 \)), inhibitory control (\( x_3 \)), and impulsivity (\( x_4 \)) were found to account for only about six percent (5.89%) of the variation in treatment gain scores.

As observed in Model 1a, baseline executive abilities were generally not found to be predictive of change in number of infractions (\( y_2 \)). The effect of risky decision making (\( x_2 \)) approached significance (\( \beta = .15, SE = .13, p = .0643 \)); however, no baseline executive ability was found to significantly predict change in number of infractions (\( y_2 \)). Furthermore, treatment gain score (\( y_1 \) mediator) was not found to be a significant predictor of change in number of infractions (\( y_2 \), (\( \beta = .02, SE = .02, p = .7188 \)). (See Table 9 for complete list of standardized path coefficients of direct effects in Model 1b).

Together, baseline executive abilities accounted for only about three percent of the variation in change in number of infractions. Note results obtained in Model 1b closely resembled results obtained in Model 1a (Model 1a \( R^2 = .03319 \), Model 1b \( R^2 = .03213 \)). As I previously discussed, these results suggest that executive function capacities might be marginally more predictive of treatment-related outcomes (e.g., number of treatment groups completed, treatment gain score) compared to recidivism (e.g., change in number of infractions, \( y_2 \)).
Table 9. Direct Effects in Model 1b (LISREL parameter estimates)

<table>
<thead>
<tr>
<th>Paths</th>
<th>$b$</th>
<th>SE</th>
<th>$\beta$</th>
<th>Z-value</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) $x_1 \rightarrow$ Treatment Gain Score $y_1$</td>
<td>-14.36</td>
<td>5.98</td>
<td>-1.19</td>
<td>-2.40</td>
<td>.0164*</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $x_2 \rightarrow$ Treatment Gain Score $y_1$</td>
<td>.56</td>
<td>.34</td>
<td>.11</td>
<td>1.67</td>
<td>.0949</td>
</tr>
<tr>
<td>Inhibitory Control (Interference score) $x_3 \rightarrow$ Treatment Gain Score $y_1$</td>
<td>.15</td>
<td>.25</td>
<td>.04</td>
<td>.60</td>
<td>.5485</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $x_4 \rightarrow$ Treatment Gain Score $y_1$</td>
<td>-.88</td>
<td>.68</td>
<td>-.12</td>
<td>-1.53</td>
<td>.1260</td>
</tr>
<tr>
<td>Response Inhibition (DPerCor) $x_1 \rightarrow$ Change in Number of Infractions $y_2$</td>
<td>-2.95</td>
<td>1.92</td>
<td>-.12</td>
<td>-1.53</td>
<td>.1260</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) $x_2 \rightarrow$ Change in Number of Infractions $y_2$</td>
<td>.25</td>
<td>.13</td>
<td>.15</td>
<td>1.85</td>
<td>.0643</td>
</tr>
<tr>
<td>Inhibitory Control (Interference score) $x_3 \rightarrow$ Change in Number of Infractions $y_2$</td>
<td>-.01</td>
<td>.09</td>
<td>-.01</td>
<td>-.07</td>
<td>.9442</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) $x_4 \rightarrow$ Change in Number of Infractions $y_2$</td>
<td>-.03</td>
<td>.18</td>
<td>-.01</td>
<td>-.18</td>
<td>.8572</td>
</tr>
<tr>
<td>Treatment Gain Score $y_1 \rightarrow$ Change in Number of Infractions $y_2$</td>
<td>.01</td>
<td>.02</td>
<td>.02</td>
<td>.36</td>
<td>.7188</td>
</tr>
</tbody>
</table>

Note. $b$ = unstandardized regression coefficient from LISREL solution. SE = standard error. $\beta$ = standardized regression coefficient from LISREL solution.

Indirect Effects in Path 1b. Following the procedure outlined for Model 1a, Hayes’ (2013) PROCESS macro was used to conduct mediational analyses for Model 1b. Using this macro, I generated estimates for path coefficients as well as bootstrap CIs for total and specific indirect effects of exogenous, predictor variables (e.g., baseline executive functions $x_1 - x_4$) on the endogenous outcome variable, change in number of infractions ($y_2$).

Mediational analyses indicated that treatment gain score ($y_1$) did not mediate the relationships between any of the baseline executive abilities (e.g., $x_1 - x_4$) and change in number of infractions ($y_2$). Bootstrap analyses revealed the indirect effects of baseline executive abilities via treatment gain score on change in number of infractions were not significant. As seen in Table 10, the 95% bias-corrected bootstrap confidence intervals obtained for each predictor (e.g., $x_1 - x_4$), included zero. Consequently, none of the four tested indirect effects in Model 1b were found to be statistically significant at $p < .05$ (Tryon, 2001). (See Table 10 for complete listing of parameter estimates for indirect effects in Model 1b).

In addition to bootstrap analyses, results of the Sobel test (1982) are also reported (see Table 10). Sobel test results also indicated that treatment gain score ($y_1$) did not mediate the
relationship between any baseline executive ability \((x_1 - x_4)\) and change in number of infractions \((y_2)\); all Sobel test \(p\) values were greater than .74.

Table 10. Indirect Effects in Model 1b (LISREL parameter estimates)

<table>
<thead>
<tr>
<th>Paths</th>
<th>Sobel Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition (DPerCor) (x_1 \rightarrow \text{Change in Number of Infractions (Clnfract)} \ y_2)</td>
<td>(-.11)</td>
</tr>
<tr>
<td>Risky Decision Making (CYY1NNB) (x_2 \rightarrow \text{Change in Number of Infractions (Clnfract)} \ y_2)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Inhibitory Control (Interference Score) (x_3 \rightarrow \text{Change in Number of Infractions (Clnfract)} \ y_2)</td>
<td>(.00)</td>
</tr>
<tr>
<td>Impulsivity (# of early presses) (x_4 \rightarrow \text{Change in Number of Infractions (Clnfract)} \ y_2)</td>
<td>(-.01)</td>
</tr>
</tbody>
</table>

Note. \(\beta\) = standardized path coefficients from LISREL solution. \(SE\) = standard error. CI = bias-corrected bootstrap confidence interval, 95% two-tailed obtained via Hayes’ PROCESS Macro for SPSS, version 4 (Hayes, 2013) with 10,000 bootstrap samples

**Hypothesis 2: Does Age Predict Change in EF? Does Change in EF Vary by Age Group?**

To test the second hypothesis, I first ran a multivariate regression using LISREL 8.8 (Jöreskog & Sörbom, 1996) to test whether age (continuous variable) predicted changes in executive functioning. See Figure 3 for conceptual diagram of model being tested in this analysis.

This multivariate regression model was run with maximum likelihood (ML) estimation based on the asymptotic covariance matrix that was derived from the subsample of participants who had complete data for the four variables being tested in this model (e.g., age \((x_1)\), change in reaction time (Dice Game) \((y_1)\), change in number of risky decisions made (Dice Game) \((y_2)\), and change in percentage of choices made that were safe (Dice Game) \((y_3)\)), \(N = 65\). Using this method was advantageous as the use of the asymptotic covariance matrix adjusts for non-normality in the data and thereby provides more reliable estimates of effects/path coefficients.
than would be obtained by running this regression using another program, such as SPSS, which does not allow for corrections to adjust for non-normality.

Figure 3. Path Diagram of Conceptual Model Tested in the Present Research (Hypothesis 2/Model 2)

\[ x_1, \text{Age} \]

\[ y_1, \text{Change in RT for all Dice Game Trials} \]

\[ y_2, \text{Change in # of Risky Decisions} \]

\[ y_3, \text{Change in % Safe Choices} \]

**Descriptive and Correlational Analyses of Variables Tested in Hypothesis 2**

Means, standard deviations, ranges of scores, and correlations among Hypothesis 2 variables tested in this model (e.g., Model 2) are presented in Table 11. Among these variables, only change in number of risky decisions \( (y_2) \) and change in the percentage of choices that were safe \( (y_3) \) were found to be significantly correlated, \( r = -.621, p < .01 \), in the expected direction. Higher numbers of risky decisions was associated with lower percentage of safe choices. Changes in reaction time was not found to be correlated with either changes in number of risky decisions or changes in the percentage of choices that were safe. Moreover, I found that age \( (x_1) \) was not significantly correlated with any change in executive function performance score (e.g., scores on \( y_1, y_2, \) or \( y_3 \)).
Table 11. Pearson correlations, means and standard deviations for Hypothesis 2 Variables (N=65)

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>$x_1$</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age $x_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game) $y_1$</td>
<td>-1.20</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game) $y_2$</td>
<td>.085</td>
<td>.094</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game) $y_3$</td>
<td>.034</td>
<td>.005</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>30.400</td>
<td>-.101</td>
<td>.523</td>
<td>-.049</td>
</tr>
<tr>
<td>SD</td>
<td>5.393</td>
<td>1.115</td>
<td>1.415</td>
<td>.132</td>
</tr>
<tr>
<td>Range</td>
<td>21-41</td>
<td>-4.70-3.13</td>
<td>-4.00-4.00</td>
<td>-.38-.20</td>
</tr>
</tbody>
</table>

Note. SD = Standard Deviation. ** = statistically significant at two-tailed $p < 0.01$

Hypothesis 2: Multiple Regression Results

Regression analyses indicated that age did not significantly predict change in any executive function (e.g., $y_1$, $y_2$, $y_3$). (See Table 12 for complete list of unstandardized and standardized regression coefficients, SEs, and $p$ – values obtained for Model 2 variables). Age was found to explain approximately 1.4% of the variation in change in reaction time ($y_1$, $R^2 = 0.01435$) and less than one percent of the variation in both change in number of risky decisions ($y_2$, $R^2 = 0.00718$) and change in the percentage of choices made that were safe ($y_3$, $R^2 = 0.00117$).

Table 12. LISREL Parameter Estimates for Hypothesis 2: Does Age Predict Change in EF (Model 2)

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>$b$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>Z-value</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game) $y_1$</td>
<td>-.02</td>
<td>.02</td>
<td>-.12</td>
<td>-1.30</td>
<td>.1936</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game) $y_2$</td>
<td>.02</td>
<td>.04</td>
<td>.08</td>
<td>.62</td>
<td>.5353</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game) $y_3$</td>
<td>.00</td>
<td>.00</td>
<td>.03</td>
<td>.30</td>
<td>.7642</td>
</tr>
</tbody>
</table>

Note. $b$ = unstandardized regression coefficient from LISREL solution. $SE$ = standard error. $\beta$ = standardized regression coefficient from LISREL solution.

Hypothesis 2: Does Change in Executive Functioning Vary by Age Group

To further explore the relationship between age and change in executive functioning, additional analyses were performed. In light of research demonstrating executives’ abilities
continue to improve through the early 20s and continue to develop at a slightly protracted rate through the late 20s, as discussed in the review of the literature, I also wanted to test whether changes in executive functioning varied by age group. Specifically, younger inmates (e.g., those who were 21 to 25 years) were expected to exhibit greater changes in executive functioning as compared to older inmates (e.g., those 26-30 years, 31-35 years, and 36-40 years).

Before testing for differences in executive functioning across age groups, I evaluated whether Analysis of Variance (ANOVA) assumptions were met. Note each change in executive function score was considered separately in this analysis as correlations between these variables were found to be generally low as previously discussed (see Table 11). The lack of strong correlations across the change in executive functioning variables suggests these variables are not interrelated and do not capture a single underlying latent variable; therefore, Multivariate Analysis of Variance (MANOVA) was not appropriate in this case.

To test for homogeneity of variances, Levene’s Test was run using SPSS (Version 27). Results indicated that both change in reaction time \( (F(3,67) = 4.68, p = 0.005) \) and change in percentage of choices made that were safe \( (F(3,71) = 4.51, p = 0.006) \) failed this assumption. However, I failed to reject the null hypothesis for change in number of risky decisions \( (F(3,64) = 1.33, p = .272) \). Therefore, I assumed homogeneity of variances among age groups for the variable change in number of risky decisions.

In addition, I also assessed for normality using the Shapiro-Wilk Test (see Table 13). Results indicated that distributions in change in executive functions were not normally distributed for several age groups. Specifically, the distribution of scores for change in number of risky decisions were not normally distributed for inmates who were: 21 to 25 years \( (W(11) = .748, p = .002) \), 26 to 30 years \( (W(25) = .876, p = .006) \), or 31 to 35 years \( (W(16) = .814, p = .406) \).
In addition, the distribution in change in percentage of choices made that were safe was not found to be normally distributed among inmates who were 31 to 35 years of age ($W(16) = .852, p = .015$). These results indicate that for these variables, the assumption of normality across age groups was violated.

Table 13. Tests for Normality Among Hypothesis 2 Variables (Shapiro-Wilk)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Variable</th>
<th>W</th>
<th>Df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-25 years</td>
<td>Change in RT</td>
<td>.895</td>
<td>11</td>
<td>.158</td>
</tr>
<tr>
<td></td>
<td>Change in # of risky decisions</td>
<td>.748</td>
<td>11</td>
<td>.002*</td>
</tr>
<tr>
<td></td>
<td>Change in % safe choices</td>
<td>.886</td>
<td>11</td>
<td>.124</td>
</tr>
<tr>
<td>26-30 years</td>
<td>Change in RT</td>
<td>.977</td>
<td>25</td>
<td>.831</td>
</tr>
<tr>
<td></td>
<td>Change in # of risky decisions</td>
<td>.876</td>
<td>25</td>
<td>.006*</td>
</tr>
<tr>
<td></td>
<td>Change in % safe choices</td>
<td>.976</td>
<td>25</td>
<td>.802</td>
</tr>
<tr>
<td>31-35 years</td>
<td>Change in RT</td>
<td>.932</td>
<td>16</td>
<td>.261</td>
</tr>
<tr>
<td></td>
<td>Change in # of risky decisions</td>
<td>.814</td>
<td>16</td>
<td>.004*</td>
</tr>
<tr>
<td></td>
<td>Change in % safe choices</td>
<td>.852</td>
<td>16</td>
<td>.015*</td>
</tr>
<tr>
<td>36-40 years</td>
<td>Change in RT</td>
<td>.927</td>
<td>12</td>
<td>.347</td>
</tr>
<tr>
<td></td>
<td>Change in # of risky decisions</td>
<td>.915</td>
<td>12</td>
<td>.247</td>
</tr>
<tr>
<td></td>
<td>Change in % safe choices</td>
<td>.951</td>
<td>12</td>
<td>.649</td>
</tr>
</tbody>
</table>

Note. * = statistically significant at two-tailed $p < .05$. $p < .05$ indicates assumption of normality has been violated

**Hypothesis 2: Kruskal-Wallis One-Way ANOVA Results**

As multiple ANOVA assumptions were violated (e.g., assumption of homogeneity of variance and assumption of normality), the Kruskal-Wallis One-Way ANOVA was used to assess whether there were differences in any of our three change in executive function scores across age groups. The Kruskal-Wallis test is the non-parametric equivalent to one-way ANOVA (Kruskal & Wallis, 1952).

First, a Kruskal-Wallis test was performed to explore change in reaction times across the four age groups (e.g., 21-25 years, 26-30 years, 31-35 years and 36-40 years). Results indicated that there was no significant difference among changes in reaction time across age groups.
Thus, the overall distribution of change in reaction time scores across age groups was not significantly different.

Despite a non-significant Kruskal-Wallis result, when looking at the median scores across age groups (see Table 14), it is evident that the median scores for change in reaction time varies considerably across age groups. For example, among individuals who were 21 to 25 years, the median change in reaction time was .1120 as compared to -.4040 among individuals 36 to 40 years. Indeed, an Independent-Samples Median Test indicated that the median change reaction time was not the same across age groups ($X^2(3, 71) = 10.479, p = 0.015$. Post-hoc analyses indicated that there was a significant difference in median change in reaction time scores between 31 to 35-year-olds and 36 to 40-year-olds ($p = 0.044$, adjusted using the Bonferroni correction). There was also a significant difference in median change in reaction time scores between 21 to 25-year-olds and 36 to 40-year-olds; however, this difference did not maintain significance when adjusted with the Bonferroni correction (pre-correction $p = .047$, $p$ with Bonferroni correction = 0.283)

Table 14. Median Scores for Change in Reaction Time by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Median score</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-25 years</td>
<td>.1120</td>
</tr>
<tr>
<td>26-30 years</td>
<td>.3260</td>
</tr>
<tr>
<td>31-35 years</td>
<td>.2970</td>
</tr>
<tr>
<td>36-40 years</td>
<td>-.4040</td>
</tr>
</tbody>
</table>

To assess whether change in number of risky decisions varied by age group, I again ran a Kruskal-Wallis test. The results of this test indicated there was no difference in change in number of risky decisions across age groups ($H(3) = 1.887, p = 0.596$. The median change in number of risky decisions made was .0000 for all age groups. Unsurprisingly, an Independent
Samples Median Test indicated there was no significant difference in median change in number of risky decisions across age groups ($X^2(3, 68) = .820, p = 0.845$).

Finally, I also tested whether change in percentage of choices made that were safe varied by age group. A Kruskal-Wallis test indicated that the distribution in change in percentage of safe choices did not vary by age group ($H(3) = 4.794, p = 0.188$). The median score for change in percentage of safe choices was: -.0666 for those 21 to 25-years-old, -.0444 for those 26 to 30-years-old, .0000 for those 31 to 35-years-old, and .0444 for those 36 to 40- years-old. See Table 15 for median change in percentage of safe choices scores across age groups.

An Independent Samples Median Test indicated the median scores in change in percentage of safe choices were different across age groups ($X^2(3,75) = 8.001, p = .046$). Post-hoc analyses indicated that there was a significant difference in median scores for change in percentage of safe choices between individuals who were 21 to 25 years old compared to those who were 36 to 40 years old and between those who were 26 to 30 years old compared to those who were 36 to 40 years old; however these differences did not remain significant when adjusted using the Bonferroni correction (21-25 years vs. 36-40 years- $p = .014$, $p$ value with Bonferroni correction = .086; 26-30 years vs. 36-40 years- $p = .030$, $p$ value with Bonferroni correction = .179).

Table 15. Median Scores for Change in Percentage of Safe Choices by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Median score</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-25 years</td>
<td>-.0666</td>
</tr>
<tr>
<td>26-30 years</td>
<td>-.0444</td>
</tr>
<tr>
<td>31-35 years</td>
<td>.0000</td>
</tr>
<tr>
<td>36-40 years</td>
<td>.0444</td>
</tr>
</tbody>
</table>
Hypothesis 2: Additional Analyses – Baseline Executive Functioning by Age Group

After running the initially planned analyses exploring whether change in executive functioning differed across age groups, I also considered whether baseline executive functions varied by age group in this sample. Before testing for differences in baseline executive functioning across age groups, I assessed whether Analysis of Variance (ANOVA) assumptions were met. Again, to test for homogeneity of variances, Levene’s Test was run using SPSS (Version 27). Results indicated that number of risky decisions failed this assumption ($F(3, 196) = 6.52, p < .001$). However, I failed to reject the null hypothesis for: response inhibition ($F(3, 196) = .29, p = .833$), inhibitory control ($F(3, 196) = 1.53, p = .208$), and impulsivity ($F(3, 196) = 2.45, p = .065$). Therefore, I assume homogeneity of variances among age groups for baseline measures of response inhibition, inhibitory control, and impulsivity.

In addition, I also assessed for normality using the Shapiro-Wilk Test (see Table 16). Results indicate that distributions in baseline executive functions were not normally distributed for several age groups. Specifically, the distribution of scores for response inhibition was not normally distributed for inmates who were: 26 to 30 years ($W(51) = .956, p = .029$) and 36 to 40 years ($W(47) = .914, p = .002$). The distribution of scores for inhibitory control were also not found to be normally distributed among inmates who were 26 to 30 years of age ($W(61) = .945, p = .008$). Furthermore, the distribution of scores for impulsivity and risky decision making were not normally distributed across all age groups.
Table 16. Test for Normality Among Baseline Executive Functioning Across Age Group (Shapiro-Wilk)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Variable</th>
<th>W</th>
<th>Df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>21-25 years</td>
<td>Response Inhibition (Difference in % correct baseline vs tone trials Stop Signal Task)</td>
<td>.964</td>
<td>40</td>
<td>.222</td>
</tr>
<tr>
<td></td>
<td>Impulsivity (# early presses during all Stop Signal task trials)</td>
<td>.588</td>
<td>40</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Risky Decisions Made (# risky decisions in Dice Game)</td>
<td>.699</td>
<td>40</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Inhibitory Control (Stroop interference score)</td>
<td>.982</td>
<td>40</td>
<td>.759</td>
</tr>
<tr>
<td>26-30 years</td>
<td>Response Inhibition (Difference in % correct baseline vs tone trials Stop Signal Task)</td>
<td>.956</td>
<td>61</td>
<td>.029*</td>
</tr>
<tr>
<td></td>
<td>Impulsivity (# early presses during all Stop Signal task trials)</td>
<td>.737</td>
<td>61</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Risky Decisions Made (# risky decisions in Dice Game)</td>
<td>.591</td>
<td>61</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Inhibitory Control (Stroop interference score)</td>
<td>.945</td>
<td>61</td>
<td>.008*</td>
</tr>
<tr>
<td>31-35 years</td>
<td>Response Inhibition (Difference in % correct baseline vs tone trials Stop Signal Task)</td>
<td>.982</td>
<td>52</td>
<td>.626</td>
</tr>
<tr>
<td></td>
<td>Impulsivity (# early presses during all Stop Signal task trials)</td>
<td>.593</td>
<td>52</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Risky Decisions Made (# risky decisions in Dice Game)</td>
<td>.768</td>
<td>52</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Inhibitory Control (Stroop interference score)</td>
<td>.973</td>
<td>52</td>
<td>.278</td>
</tr>
<tr>
<td>36-40 years</td>
<td>Response Inhibition (Difference in % correct baseline vs tone trials Stop Signal Task)</td>
<td>.914</td>
<td>47</td>
<td>.002*</td>
</tr>
<tr>
<td></td>
<td>Impulsivity (# early presses during all Stop Signal task trials)</td>
<td>.721</td>
<td>47</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Risky Decisions Made (# risky decisions in Dice Game)</td>
<td>.823</td>
<td>47</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td></td>
<td>Inhibitory Control (Stroop interference score)</td>
<td>.975</td>
<td>47</td>
<td>.399</td>
</tr>
</tbody>
</table>

As several ANOVA assumptions were violated among baseline executive functions, the Kruskal-Wallis test was used to test for potential differences in baseline executive abilities across age-groups. Specifically, four Kruskal Wallis tests were run to explore for potential age-group related differences in baseline response inhibition, impulsivity, risky decision making, and inhibitory control.

The results of these tests indicated that there were no differences in impulsivity (e.g., number of early presses) across age groups ($H(3) = 6.207, p = .102$). However, significant differences were found across age groups in regards to: response inhibition ($H(3) = 10.902, p = .012$), risky decision making ($H(3) = 10.142, p = .017$), and inhibitory control ($H(3) = 27.934, p < .001$). (See Table 17 below for mean scores for baseline executive functions across age groups).

Post-hoc analyses indicated that there was a significant difference in baseline response inhibition among individuals 21-25 years old versus those who were 36-40 years old ($p$ value = .028, adjusted with Bonferroni correction) and among individuals who were 26-30 years old compared those who were 36-40 years old ($p = .036$, adjusted with Bonferroni correction). For
risky decision making, a significant difference was observed among those who were 26-30 years old and those who were 36-40 years old ($p = .013$, adjusted for Bonferroni correction). Lastly, for inhibitory control, significant differences were found between the oldest group (e.g., those 36 to 40 years) and all other age groups- 21-25 years old vs. 36 to 40 years old ($p = .011$, adjusted for Bonferroni correction), 26-30 years old vs. 36 to 40 years old ($p = .000$, adjusted for Bonferroni correction), and 31-35 years old vs. 36 to 40 years old ($p = .000$, adjusted for Bonferroni correction).

Table 17. Mean Scores in Baseline Executive Abilities by Age Group

<table>
<thead>
<tr>
<th></th>
<th>21-25 years</th>
<th>26-30 years</th>
<th>31-35 years</th>
<th>36-40 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Inhibition</td>
<td>.26</td>
<td>.27</td>
<td>.29</td>
<td>.33</td>
</tr>
<tr>
<td>(Difference in % correct during baseline &amp; tone trials in Stop Signal Task)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impulsivity (# early presses during all Stop Signal task trials)</td>
<td>.41</td>
<td>.49</td>
<td>.65</td>
<td>.87</td>
</tr>
<tr>
<td>Risky Decisions Made (# risky decisions in Dice Game)</td>
<td>1.02</td>
<td>.62</td>
<td>1.33</td>
<td>1.64</td>
</tr>
<tr>
<td>Safe Decision Making (% safe choices in Dice Game)</td>
<td>.67</td>
<td>.68</td>
<td>.66</td>
<td>.65</td>
</tr>
<tr>
<td>Inhibitory Control (Stroop interference score)</td>
<td>20.71</td>
<td>20.04</td>
<td>19.76</td>
<td>18.10</td>
</tr>
</tbody>
</table>

Hypothesis 3: Does Engagement in Treatment Predict Change in Executive Functioning?

Model 3: Does Program Treatment Gain Predict Change in Executive Functioning?

Model 3 considered the effects of treatment engagement on changes in executive functioning. As the CBT treatment program emphasized developing skills that closely align with executive abilities (e.g., addressing impulsivity, risky behaviors, poor decision making etc.), it was hypothesized that inmates who had greater engagement in treatment would demonstrate greater improvement in executive abilities as compared to other inmates.

For analyses testing Hypothesis 3, engagement in treatment/participation in treatment was assessed using treatment gain scores. Recall the Treatment Gain: Short Scale (Serin et al., 2002) includes items that “represent a combination of knowledge, participation and competencies” (Fishbein & Sheppard, 2006, p. 82) and was designed to “provide an overall
estimate of an offender’s performance in correctional programming” (Fishbein & Sheppard, 2006, p. 82). Thus, higher scores on this scale indicate better participation in treatment programming, engagement in programming, and implementation/internalization of program-targeted skills and/or content.

Originally, scores on the Treatment Responsivity: Short Scale (Serin et al., 2002) was also included as a predictor in this model as this scale, a companion to the Treatment Gain: Short Scale, was designed as an indicator of offender’s performance in treatment (Fishbein & Sheppard, 2006). As I previously discussed, this scale includes items that were selected to represent potential responsivity factors relating to inmates’ compliance with, and response to, therapeutic interventions and treatment programs in general. The focus of this measure differs slightly from the Treatment Gain: Short Scale as this scale more so taps into offenders’ general interpersonal styles and scale items are not necessarily treatment specific. However, scores on these scales were extremely correlated \( r = .959, p < .01, N = 198 \). Accordingly, the inclusion of both items was deemed redundant; to include both measures would violate assumptions of non-multicollinearity among predictors in a combined model.

To assess change in executive functioning following treatment, I considered: change in reaction times for all dice game trials \( (y_1) \), change in number of risky decisions made during the dice game \( (y_2) \), and change in percent of safe choices made in the dice game \( (y_3) \). Note the only measure of executive functioning administered at baseline and final follow-up was the Dice Game. Thus, it was not possible to explore changes in executive functioning evidenced by performance indicators derived from either the Stop Signal task or the Stroop interference task.

Changes in these Dice Game performance indicators were selected in order to capture various executive abilities to the extent possible. Specifically, reaction time (RT) is noted to be a
fundamental contributor to information processing and is considered to be an index for information processing speed (Ghuntla, Mehta, Gokhale, & Shah, 2014). Change in number of risky decisions made was chosen as it was thought to provide an indication of poor decision making and could potentially tap into short-sightedness or impulsiveness (e.g., focus on immediate benefits/gains over totality of circumstances). Conversely, change in percentage of safe choices was thought to provide an indication of better or more mature and rational decision-making.

**Descriptive and Correlational Analyses of Model 3/Hypothesis 3 Variables**

Means, standard deviations, ranges of scores, and correlations among Model 3 variables are presented in Table 18 below. Treatment gain scores were not found to be significantly correlated with any of outcome measures (e.g., change in reaction time for all trials (Dice Game) $y_1$, change in number of risky decisions (Dice Game) $y_2$, and change in percent of choices that were safe (Dice Game) $y_3$). Thus, treatment gain score did not appear to be correlated with changes in executive abilities from baseline to follow-up as was expected. However, the correlation between treatment gain score ($x_1$) and change in number of risky decisions ($y_2$) did approach significance ($r = .243, p = .053$).

Regarding intercorrelations among change scores for executive abilities (e.g., $y_1 - y_3$), I found that change in number of risky decisions ($y_2$) was significantly correlated with change in percentage of choices that were safe ($y_3$) in the expected direction ($r = -.623, p < .01$), such that those who demonstrated greater reductions in risky decision making from baseline to follow-up demonstrated greater improvements in regards to decision making/making safe choices. For change in number of risky decisions ($y_2$), larger positive scores were indicative of greater reductions in risky decision making from baseline to follow-up. Conversely, for change in
percentage of choices that were safe ($y_3$), larger negative scores were indicative of greater
increases in “good” decision making (e.g., more safe choices made during Dice Game) from
baseline to follow-up.

Change in reaction time for all trials ($y_1$) was not significantly correlated with either
change in number of risky decisions ($y_2$) or change in percentage of choices that were safe ($y_3$).
Thus, changes in reaction times did not correlate with changes to decision making, as evidenced
by changes to risky decision making or safe/responsible decision making from baseline to
follow-up.

Table 18. Pearson Correlations, Means and Standard Deviations for Hypothesis 3 Variables
($N=64$)

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>$x_1$</th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Gain Score</td>
<td>-</td>
<td>.171</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game)</td>
<td>.243</td>
<td>-</td>
<td>-.097</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game)</td>
<td>-.190</td>
<td>.002</td>
<td>-.623**</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>17.453</td>
<td>.093</td>
<td>.531</td>
<td>-.048</td>
</tr>
<tr>
<td>$SD$</td>
<td>3.651</td>
<td>1.122</td>
<td>1.425</td>
<td>.133</td>
</tr>
<tr>
<td>Range</td>
<td>8-24</td>
<td>-4.699–3.125</td>
<td>-4.4</td>
<td>-.378–.200</td>
</tr>
</tbody>
</table>

*Note. SD = standard deviation. * = statistically significant at two-tailed p < 0.05, ** = statistically significant at two-tailed p < 0.01

**Model 3 Hypothesis Testing**

To test the hypothesis that changes in executive abilities would be influenced by inmates’
engagement in treatment programming, I conducted three separate regression analyses using
LISREL 8.8 (Jöreskog & Sörbom, 1996) with Maximum Likelihood (ML) estimation based on
the asymptotic covariance matrix derived from the subsample of participants who enrolled in the
treatment program, were scored for treatment gain, and participated in baseline and follow-up
administration of the Dice Game (and had scores for each performance indicator assessed in this
model) \((N = 64)\). Again, this approach was taken as it allowed for correcting for non-normality in the data.

Each regression included treatment gain score as the sole predictor with each regression including a different change in executive function as the outcome. Thus, the first regression tested whether treatment gain score predicted change in reaction time. The second regression tested whether treatment gain score predicted change in number of risky decisions and the third regression tested whether treatment gain score predicted change in percentage of choices made that were safe. Exploring these relationships using separate regressions was appropriate as changes in executive functions were not highly correlated with one another and it was not the intention of this analysis to explore relationships between executive abilities, but rather to see if treatment gains were related to improvements in executive functioning. This method tested hypotheses by estimating: unstandardized \((b)\) and standardized and path coefficients \((\beta)s\), standard errors \((SEs)\), and their \(p\) values. In addition, this method produced estimates of the proportion of variance explained in each endogenous outcome variable \((e.g., y_1 - y_3)\).

As see in Table 19, treatment gain scores did not predict change in any executive ability \((all \ p-values \ greater \ than \ .05)\). Among all of the executive functions tested, treatment gain scores appeared to have the biggest effect on change in number of risky decisions \((\beta = .24, \ p = .0949)\); Treatment gain scores were found to explain nearly six percent of the variance in change in number of risky decisions \((R^2 = .05896)\). In contrast, treatment gain scores were found to account for less than four percent of the variance in change in percentage of safe choices made and less than three percent of the variance in change in reaction time.
Table 19. LISREL Parameter Estimates for H3: Do Treatment Gains Predict Change in EF

<table>
<thead>
<tr>
<th>Paths</th>
<th>b</th>
<th>SE</th>
<th>β</th>
<th>Z-value</th>
<th>p</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Gain Score $x_1 \rightarrow$ Change in Reaction Time for All Trials (Dice Game) $y_1$</td>
<td>.05</td>
<td>.04</td>
<td>.17</td>
<td>1.46</td>
<td>.1443</td>
<td>.02931</td>
</tr>
<tr>
<td>Treatment Gain Score $x_1 \rightarrow$ Change in Number of Risky Decisions (Dice Game) $y_2$</td>
<td>.09</td>
<td>.06</td>
<td>.24</td>
<td>1.67</td>
<td>.0949</td>
<td>.05896</td>
</tr>
<tr>
<td>Treatment Gain Score $x_1 \rightarrow$ Change in % of Choices that were Safe (Dice Game) $y_3$</td>
<td>-.01</td>
<td>.00</td>
<td>-.19</td>
<td>-1.50</td>
<td>.1336</td>
<td>.03617</td>
</tr>
</tbody>
</table>

*Note. b = unstandardized regression coefficient from LISREL solution. SE = standard error. β = standardized regression coefficient from LISREL solution.*

**Hypothesis 4: Does Change in Executive Functioning Predict Recidivism**

To test whether changes in executive functioning predicted recidivism, I ran three separate multivariable regressions. Each multivariable regression, included three independent variables: a) change in reaction time for all Dice Game trials, b) change in number of risky decisions made during the Dice Game and c) change in the percentage of safe choices made during the Dice Game. In addition, each regression included one dependent variable that was used to approximate recidivism. In Model 4, recidivism was assessed using change in number of infractions, which provided a behavioral indicator of recidivism (see Figure 4). In addition, I also assessed recidivism using self-reports of inmates’ feelings of anger and reactivity by testing changes in Novaco Anger Scale (NAS) total scores and Provocation Inventory (PI) total scores, respectively. The effects of changes in executive functions on changes in NAS score were explored in Model 5 (see Figure 5) while the effects of changes in executive functions on changes in PI score were explored in Model 6 (see Figure 6).

By considering both behavioral changes, as measured by institutional records, and self-reports of inmate’s feelings of anger and reactivity, or tendency towards being provoked into anger, these analyses were able to ascertain a more comprehensive picture of recidivism than would be captured by behavioral indicators alone.
The three aforementioned multivariable regression models were tested using LISREL 8.8 (Jöreskog & Sörbom, 1996) with maximum likelihood (ML) estimation based on the asymptotic covariance matrices that were derived from the subsample of participants who had complete data for the four variables being tested in a given model. Thus, the sample sizes across these three models varied slightly. (See Tables 15, 17 and 19 for sample sizes in Models 4, 5, and 6 respectively).

Again, using this method allowed me to adjust for non-normality in the data thereby providing more reliable estimates of effects/path coefficients. Moreover, by running these analyses separately, I retained the largest sample size possible for each analysis, which was a paramount concern as sample sizes were quite small and I sought to retain the maximum sample size possible for each analysis.

**Model 4: Does Change in EF Predict Recidivism (Change in Number of Infractions)**

Figure 4. Path Diagram of Conceptual Model Tested in the Present Research (Model 4)

Means, standard deviations, ranges of scores, and correlations among Model 4 variables are presented in Table 20 below. Among these variables, only change in number of risky decisions ($x_2$) and change in the percentage of choices that were safe ($x_3$) were found to be significantly correlated, $r = -.621, p < .01$. This correlation was in the expected direction. Higher
numbers of risky decisions was associated with a lower percentage of choices that were safe. Change in reaction time was not found to be correlated with either changes in number of risky decisions or changes in the percentage of choices that were safe.

I also found that changes in executive function performance scores (e.g., scores on $x_1$, $x_2$ and $x_3$) were not significantly correlated with change in number of infractions ($y_1$). However, correlations between change in executive function performance scores and change in number of infractions were quite similar with $r$ values ranging from $|0.200|$ to $|0.215|$. Thus, changes in various executive function performance scores were equally, but weakly associated with changes in number of infractions. (See Table 20 for complete list correlations among Model 4 variables).

Table 20. Pearson Correlations, Means and Standard Deviations for Model 4 Variables ($N=65$)

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$y_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game) $x_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game) $x_2$</td>
<td>-.094</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game) $x_3$</td>
<td>.005</td>
<td>-.621**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Infractions $y_1$</td>
<td>.200</td>
<td>.212</td>
<td>-.215</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>-.101</td>
<td>.523</td>
<td>-.049</td>
<td>2.277</td>
</tr>
<tr>
<td>SD</td>
<td>1.115</td>
<td>1.415</td>
<td>.132</td>
<td>2.837</td>
</tr>
<tr>
<td>Range</td>
<td>-4.70-3.13</td>
<td>-4.00-4.00</td>
<td>-.38-.20</td>
<td>0.00-13.00</td>
</tr>
</tbody>
</table>

*Note. SD = Standard Deviation. ** = statistically significant at two-tailed $p < 0.01$*

Regression analyses indicated that none of the change in executive function variables significantly predicted change in number of infractions (see Table 21). However, change in reaction time was found to be the strongest predictor; this effect approached significance ($p = .0672$). Together, change in reaction time, change in number of risky decisions and change in percentage of choices that were safe were found to account for approximately 10 percent of the variance in change in number of infractions ($R^2 = .10217$).
Table 21. LISREL Parameter Estimates for Regression Coefficients – Predicting Change in # of Infractions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>SE</th>
<th>$\beta$</th>
<th>Z-value</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game) $x_1$</td>
<td>.55</td>
<td>.30</td>
<td>.22</td>
<td>1.83</td>
<td>.0672</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game) $x_2$</td>
<td>.32</td>
<td>.43</td>
<td>.16</td>
<td>.75</td>
<td>.4532</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game) $x_3$</td>
<td>-2.51</td>
<td>2.95</td>
<td>-.12</td>
<td>-.85</td>
<td>.3953</td>
</tr>
</tbody>
</table>

Note. $b$ = unstandardized regression coefficient from LISREL solution. $SE$ = standard error. $\beta$ = standardized regression coefficient from LISREL solution.

**Model 5: Does Change in EF Predict Recidivism (Change in NAS total score)**

Means, standard deviations, ranges of scores, and correlations among Model 5 variables are presented in Table 22 below. Among the sub-sample of participants included in Model 5 analyses, I also found a significant correlation between change in number of risky decisions ($x_2$) and change in the percentage of choices that were safe ($x_3$); this correlation was slightly stronger than that observed in Model 4 analyses ($r = -.688$ vs. $r = -.621$ in Model 4). Change in reaction time ($x_1$) was found to be significantly correlated with change in NAS total score ($r = -.290, p < .05$). However, this association was not in expected direction. Initially, change in reaction time was conceptualized such that greater, positive scores were indicative of faster processing speed at follow-up as compared to baseline. Therefore, a positive relation between change in reaction time and change in NAS total score was expected, as improvements in processing speed were expected to be associated with greater reductions in anger. (Recall larger scores on change in NAS indicate greater decrease in anger across all NAS domains from baseline to follow-up).

The negative correlation observed between change in reaction time and change in NAS total score suggests that change in reaction time may be better interpreted such that negative scores represent slower processing and/or reduced impulsivity at follow-up compared to baseline. With this interpretation, it is plausible that such a change would correspond with
greater reductions in anger as people who are more thoughtful or less impulsive may be less prone to feelings of anger.

Table 22. Pearson Correlations, Means and Standard Deviations for Model 5 Variables (N=57)

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$y_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game) $x_1$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game) $x_2$</td>
<td>.106</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game) $x_3$</td>
<td>.031</td>
<td>-.688**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Novaco Anger Scale (NAS) Score $y_1$</td>
<td>-.290*</td>
<td>-.060</td>
<td>.031</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>-.108</td>
<td>.561</td>
<td>-.063</td>
<td>3.912</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.148</td>
<td>1.402</td>
<td>.130</td>
<td>13.656</td>
</tr>
<tr>
<td>Range</td>
<td>-.70-3.13</td>
<td>-4.00-4.00</td>
<td>-.38-.20</td>
<td>-37.00-35.00</td>
</tr>
</tbody>
</table>

* = statistically significant at two-tailed $p < 0.05$, ** = statistically significant at two-tailed $p < 0.01$

As shown in Table 23, regression analyses indicated that change in reaction time was the only significant predictor of change in NAS total score ($b = -3.58, p < .01$). Therefore, when controlling for change in number of risky decisions and change in the percentage of choices that were safe, every standard deviation increase (e.g., $+1\ SD$) in change in reaction time was associated with a .30 standard deviation decrease in change in NAS total score. In other words, for every standard deviation increase in change in reaction time, individuals showed .3 standard deviation less reduction in anger (as measured by change in NAS total score). Change in number of risky decisions and change in percentage of choices made that were safe were not found to predict change in NAS total score. In total, change in reaction time, change in number of risky decisions made, and change in percentage of choices made that were safe were found to account for approximately nine percent of the variation in change in NAS total score ($R^2 = .09319$).
Table 23. LISREL Parameter Estimates for Regression Coefficients – Predicting Change in NAS Score

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>SE</th>
<th>β</th>
<th>Z-value</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game)</td>
<td>-3.58</td>
<td>1.23</td>
<td>-.30</td>
<td>-2.91</td>
<td>.0036</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game)</td>
<td>-1.18</td>
<td>1.24</td>
<td>-.12</td>
<td>-.95</td>
<td>.3421</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game)</td>
<td>-4.53</td>
<td>15.56</td>
<td>-.04</td>
<td>-.29</td>
<td>.7718</td>
</tr>
</tbody>
</table>

Note. $b$ = unstandardized regression coefficient from LISREL solution. $SE =$ standard error. $\beta =$ standardized regression coefficient from LISREL solution.

**Model 6: Does Change in EF Predict Recidivism (Change in PI total score)**

Means, standard deviations, ranges of scores, and correlations among Model 6 variables are presented in Table 24 below. For the sub-sample of participants included in Model 6 analyses ($N=63$), I also found a significant correlation between change in number of risky decisions ($x_2$) and change in the percentage of choices that were safe ($x_3$), $r = -.641$. However, no other correlation between change in executive function variables was statistically significant. In addition, I found that change in Novaco Provocation Inventory (PI) total score (e.g., $y_1$) was not significantly correlated with change in any executive function (e.g., $x_1$, $x_2$ or $x_3$).

Table 24. Pearson Correlations, Means and Standard Deviations for Model 6 variables ($N=63$)

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$y_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game)</td>
<td>-.082</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game)</td>
<td>-.028</td>
<td>-.641**</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Change in Novaco Provocation Inventory (PI) Score</td>
<td>.027</td>
<td>-.049</td>
<td>-.110</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>-.111</td>
<td>.524</td>
<td>.055</td>
<td>10.365</td>
</tr>
<tr>
<td>$SD$</td>
<td>1.100</td>
<td>1.435</td>
<td>.129</td>
<td>11.986</td>
</tr>
<tr>
<td>Range</td>
<td>-4.70-3.13</td>
<td>-4.00-4.00</td>
<td>-.38-20</td>
<td>-14.00-42.00</td>
</tr>
</tbody>
</table>

Note. $SD =$ Standard Deviation. ** = statistically significant at two-tailed $p < 0.01$

In line with reported correlations among Model 6 variables, regression analyses indicated that neither change in reaction time ($x_1$), change in number of risky decisions made ($x_2$), nor
change in percentage of choices that were safe ($x_3$) were found to significantly predict change in total Provocation Inventory (PI) scores. (See Table 25 below for regression coefficients and corresponding $p$ values). In total, these three variables were found to account for less than four percent of the variance in change in total PI scores ($R^2 = .03507$).

Table 25. LISREL Parameter Estimates for Regression Coefficients – Predicting Change in PI Score

<table>
<thead>
<tr>
<th>Variable</th>
<th>$b$</th>
<th>$SE$</th>
<th>$\beta$</th>
<th>Z-value</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Reaction Time for All Trials (Dice Game) $x_1$</td>
<td>.04</td>
<td>1.45</td>
<td>.00</td>
<td>.03</td>
<td>.9761</td>
</tr>
<tr>
<td>Change in Number of Risky Decisions (Dice Game) $x_2$</td>
<td>-1.69</td>
<td>1.20</td>
<td>-.20</td>
<td>-1.40</td>
<td>.1615</td>
</tr>
<tr>
<td>Change in % of Choices that were Safe (Dice Game) $x_3$</td>
<td>-22.23</td>
<td>14.81</td>
<td>-.24</td>
<td>-1.50</td>
<td>.1336</td>
</tr>
</tbody>
</table>

Note. $b$ = unstandardized regression coefficient from LISREL solution. $SE$ = standard error. $\beta$ = standardized regression coefficient from LISREL solution.
CHAPTER FIVE
DISCUSSION

Summary of Results

The primary purpose of this study was to gain a better understanding of the relations between age and executive functioning among a correctional population. Specifically, the current study sought to examine how these variables may contribute to individuals’ capacity to benefit from correctional treatment and the likelihood of continued antisocial behavior (e.g., recidivism). It was hypothesized that the relation between offender’s baseline executive functioning and future recidivism would be mediated by engagement in correctional treatment (H1). In addition, it was predicted that changes in executive functioning following treatment would vary by age, such that younger inmates would demonstrate greater improvement in executive abilities compared to older inmates (H2). Moreover, it was hypothesized that engagement in treatment would predict changes in executive abilities, such that individuals who demonstrated higher levels of engagement in correctional treatment would evidence greater improvement in executive abilities (H3). Lastly, it was hypothesized that changes in executive functioning would predict future recidivism, such that individuals who demonstrated more improvement in executive abilities would have greater reductions in antisocial behavior (e.g., lower recidivism rates) (H4).

Hypothesis 1

Overall, results of the present study did not support the first hypothesis. Among the baseline executive functions tested, only two- response inhibition and impulsivity- were
found to be significantly correlated with the number of treatment groups completed, such that lower capacity for response inhibition and higher levels of impulsivity were associated with completion of fewer treatment groups. Response inhibition was found to be a significant predictor of treatment groups completed. The effect of impulsivity on number of treatment groups was not significant, however this effect approached significance. When considering the benefits inmates derived from correctional treatment as evaluated by social worker’s assessments, the effects of baseline executive functions were weaker. Only response inhibition was found to significantly correlate with reported treatment gains. This effect was found to be significant.

Counter to my hypotheses, baseline executive functioning was not found to be associated with recidivism. No baseline executive function was found to be significantly correlated with recidivism and accordingly, no baseline executive function was found to significantly predict recidivism.

In comparing correlations among these variables, it is notable completed than with recidivism except for risky decision making, which was more strongly correlated with recidivism. As previously noted, no baseline executive function was found to be a significant predictor of recidivism; however, the effect of risky decision making did approach significance.

Contrary to my hypotheses, treatment-related variables (e.g., the number of treatment groups completed, gains from correctional treatment) did not mediate relationships between baseline executive abilities and recidivism. Such effects would not be possible given that baseline executive abilities were not found to be related to future recidivism and therefore, there was no relationship to mediate.
Generally, obtained findings suggest that executive functioning is predictive of treatment-related outcomes, particularly whether an offender will complete treatment. Among executive capacities, response inhibition seems to be especially relevant in regards to offender rehabilitation as this capacity seems to be related to one’s ability to complete treatment and to derive benefits from treatment programming.

In contrast, executive abilities were not found to be associated with, or predictive of, recidivism, which is inconsistent with extensive body of literature demonstrating a clear relationship between executive deficits and recidivism (see Langevin & Curnoe, 2011; Ross & Hoaken, 2011; Sánchez et al., 2021; Hancock et al., 2010). However, such studies have generally evaluated relations between executive abilities and recidivism by assessing the capacity for executive deficits to predict frequency and/or severity of prior offending among currently incarcerated individuals (e.g., comparing executive capacities/deficits among first-time and returning inmates). Therefore, these studies are not prospective in nature thereby prohibiting evaluation of the utility of executive functions to predict future re-offending.

Regardless, it is possible that the measure of recidivism in the current study did not adequately capture this construct. In Fishbein and Sheppard (2006), the number of pre-treatment infractions were reported; however, the time frame during which these infractions was not specified. Therefore, comparing pre- and during-treatment numbers of infractions may not be appropriate as the time frame during which number of infractions were committed is unequal (e.g., pre-treatment infractions could be over the course of many months or even years as compared to the relatively short duration of the treatment program). Moreover, by considering the number of infractions committed while enrolled in treatment, I introduced enrollment in treatment as a possible confound. It is possible that simply being enrolled in the treatment
program influenced the number of infractions inmates committed simply as they were being occupied in treatment and therefore had less time to commit potential infractions.

Furthermore, it should be noted that infractions were generally rare, both prior to- and during- correctional treatment. Among the inmates enrolled in treatment, the average number of pre-treatment infractions was 2.58 and the average number of during-treatment infractions was only 0.16. Moreover, among my sample only 27 inmates (13.6%) committed any infraction while enrolled in treatment, with the vast majority of these individuals (22, 81.5%) committing a single infraction; the greatest number of infractions committed while enrolled in treatment was two. The restricted range of pre-and during-treatment infractions may explain why my results failed to capture relations between baseline executive abilities and recidivism.

**Hypothesis 2**

My second hypothesis predicted that age would be related to change in executive functioning and specifically, the change in executive abilities would vary by age-group. It was hypothesized that younger inmates (e.g., inmates 21 to 25 years of age) would demonstrate greater improvement in executive abilities following correctional treatment compared to older inmates.

Generally, results provided some evidence that changes in executive functioning vary by age. Although I did not find that age was a significant predictor of change in any of the executive abilities assessed, I did observe some differences in changes in executive functioning across age groups.

Results indicated that the median change in reaction time was significantly different across age groups. In particular, results indicate that there is a difference between individuals who were 31 to 35-years-old and those who were 36 to 40-years-old. When looking at the
distribution of change in reaction time scores among individuals 31 to 35 years of age in conjunction with the positive median change in reaction time score for individuals in this age group, I observed that the majority of individuals in this age group had improved reaction times from pre-test to post-test (e.g., reaction times at pre-test were slower than reaction times at post-test). In contrast, the median change in reaction time for individuals 36 to 40 years of age was negative. Moreover, in assessing the distribution of change in reaction time scores amongst this age group, I found that majority of participants (more than 75%) had negative scores. Therefore, most individuals in the oldest age group demonstrated slower reaction times from pre-test to post-test.

I also found that there was a significant difference in median change in percentage of safe choices across age groups. In particular, it seems that there is a significant difference in change in percentage of safe choices between 21 to 25-year-olds versus 36 to 40-year-olds. The negative median score for change in percentage of safe choices among 21 to 25-year-olds indicated that the majority of individuals in this age group had better decision-making (e.g., made safer and more rationale choices) at post-test than at baseline. In contrast, a majority of individuals 36 to 40 years of age made fewer safe choices at post-test as compared to baseline.

In looking at median scores across all four age groups, there appeared to be a trend of reduced change in decision making. Individuals 21 to 25 years of age showed the most improvement in regards to safe decision making, followed by individuals 26 to 30 years of age, who also generally showed improvement in safe decision making, however less consistently so than observed among 21 to 25-year-olds. In contrast, half of individuals 31 to 35-years-old demonstrated devolving capacities for safe decision-making evidenced by worse scores on safe decision making at post-test compared to pre-test. This trend continued among the oldest age
group. Among those 36 to 40 years old, 75% demonstrated less safe decision making at post-test than at pre-test.

In considering these results, it seems that younger inmates (e.g., those 21-25 years old) may more amenable than older inmates (particularly those above the age of 36 years) to rehabilitation programs that target safe decision making. Moreover, the results suggest that younger inmates may also have greater capacity for improving other cognitive functions, particularly processing speed. These results mirror that of prior research that has demonstrated that younger individuals demonstrate greater improvement following remediation targeting executive capacities compared to older individuals (see Wykes et al., 2009; Bowie et al., 2014; Corbera et al., 2017; Golas et al., 2015; Kontis et al., 2013; Thomas et al., 2017). Moreover, improvements in safe decision making observed among the youngest age group is also consistent with literature demonstrating improvement in decision-making during early adulthood (see Figner & Weber, 2011; Byrnes, 2002).

It is noteworthy, however, that I did not observe differences in changes in risky decision-making across age groups. Prior research has shown reductions in risky decisions on the Iowa Gambling Task (a task similar to the Dice Game employed by Fishbein and Sheppard (2006)), or what has been referred to as “avoidance behavior” (Christakou et al., 2013) with increasing age; performance on this task typically shows linear maturation as risky choices reduce as individuals progress in adulthood (see Cauffman et al., 2010; Crone & Van der Molen, 2004).

In exploring baseline executive functions across age groups, there was a clear pattern of greater executive deficits across older inmates, including in response inhibition, inhibitory control, and impulsivity. In my sample, older inmates demonstrated worse scores for response inhibition, inhibitory control, and impulsivity as compared to younger inmates. This trend may
be related to the amount of time spent incarcerated as it is possible that older inmates may have been incarcerated for longer periods compared to younger peers. As noted in the review of the literature, research has demonstrated that incarceration can lead to executive declines (Umbach et al., 2018). In the current sample, the total time spent incarcerated was not reported; therefore, speculations on potential differences in total time incarcerated are speculative.

**Hypothesis 3**

The results did not provide support for hypothesis 3 which predicted that engagement in treatment would predict change in executive abilities. There are several possible explanations for this observed result. For one, it is possible that the treatment program completed by inmates did not target executive abilities and subsequently did not lead to changes in these capacities. The CBT program inmates participated in was expected to address deficits commonly observed in correctional populations. This intervention, however, was not specifically designed to improved executive abilities, nor has it been proven to produce changes in these skills among this population. Therefore, it is plausible that the CBT program did not produce meaningful improvements in offenders’ executive abilities.

In addition, only one executive task was administered at pre-test, prior to participation in the treatment program, and again at post-test, following inmates’ completion of the treatment program. Therefore, comparisons on pre- and post-treatment executive functions were limited to changes in Dice Game performance scores, including the aforementioned changes to reaction time and changes to risky and safe decision making. It is possible that participation in the treatment program did produce changes in executive capacities that were not assessed in this task, for example response inhibition or inhibitory control.
Hypothesis 4

Overall, results did not provide evidence in support of hypothesis 4 which predicted that changes in executive functioning would predict recidivism. Changes in reaction time and decision-making (safe and risky) were not found to predict changes in infractions, nor were they found to predict changes in inmates’ reactivity, or tendency to be provoked into feelings of anger.

Results did indicate that changes in reaction time were predictive of changes in Novaco Anger Scale total scores, such that increases in reaction time from pre-test to post-test were associated with reductions in overall feelings of anger. This finding suggests that individuals who learned to react more slowly or become less impulsive may have gained skills in emotional regulation or in refraining of negative emotions, thereby leading to fewer and less intense feelings of anger. The observed changes make sense given that inmates were enrolled in a CBT program which was designed to help individuals reframe negative thinking patterns, develop coping skills to manage negative emotions, and to alter cognitions to change undesired behaviors.

The failure to observe relations between changes in executive functioning and markers of recidivism (both behavioral and self-report) may be due to the fact that the executive functions assessed are not particularly tied to antisocial behavior or recidivism. For example, perhaps other executive capacities, such as behavioral inhibition, may be more closely tied to recidivism.

Moreover, changes in executive functioning were not largely observed among my sample. Although some inmates did improve in executive capacities from pre-test to post-test, overall improvements in executive abilities were not observed. The general lack of improvement
in executive abilities among this sample may account for why I failed to observe relations between changes in these abilities and recidivism.

**Limitations**

There are at least three potential limitations concerning the results obtained in the current study. As previously discussed, a critical limitation is that participants in the current study did not complete an evidence-based cognitive remediation program. Thus, the extent to which executive capacities can be remediated among this population is still unknown. Moreover, it remains unclear whether improvements in executive functioning that could be derived from such remediation would lead to reductions in recidivism among correctional populations.

A second limitation can be observed in the limited assessment of executive functions post-treatment. In descriptions of the data provided by the Inter-university Consortium for Political and Social research (ICPSR), the consortium that currently stores and manages Fishbein and Sheppard’s (2006) data, it is indicated that baseline and follow-up assessments were completed for all executive performance tasks (e.g., the Dice Game, Stroop task, and Stop Signal task). However, the data set provided by ICPSR only included follow-up data for the Dice Game. This was likely a result of lack of resources as it is believed Fishbein and Sheppard (2006) intended to complete follow-up testing for all executive tasks measured at baseline.

Unfortunately, this resulted in only being able to assess for changes in executive functioning via performance on the Dice Game task. Therefore, it was not possible to assess whether subjects demonstrated improvement in other executive capacities. Nor was it possible to test for differences in potential changes among excluded executive abilities across designated age groups, as was originally intended.
A third limitation of the current study pertains to the measures used to evaluate recidivism. As noted previously, the current study was a secondary data analysis of data previously collected by Fishbein and Sheppard (2006). Therefore, the assessment of variables was limited by data that had been originally collected. In particular, the measure of recidivism was not ideal as it did not provide an equal comparison for infractions committed prior to and during treatment, as the time frame during which pre-treatment infractions were committed was not specified. Moreover, in studying the potential for any intervention to reduce recidivism, including one that seeks to target executive deficits, it is important that assessments of recidivism are based on behavior following the completion of the treatment program. It is not possible to assess whether an intervention produces desired outcomes if such outcomes are measured while treatment is on-going. To truly assess potential benefits of remediating executive deficits among correctional populations, a true experimental design should be employed wherein inmates are randomly assigned to an evidence-based program designed for cognitive remediation or to treatment as usual (e.g., standard correctional treatment). Then, comparisons of recidivism rates, institutional misconduct etc. following the culmination of treatment programs can be used to assess whether targeting executive deficits produces better behavioral outcomes than traditional correctional programming.

**Implications**

Despite these limitations, these results suggest several theoretical and practical implications. For one, the results provide evidence suggesting that certain executive abilities, namely response inhibition and to a lesser extent impulsivity, may be useful in identifying who will fare better in correctional treatment. These executive capacities may serve as useful predictors of individuals who may be more likely to drop out of treatment programs, or who may
be less likely to benefit from treatment program, and therefore may need additional resources or enhanced interventions beyond standard correctional programming.

In addition, the results provide interesting findings regarding developmental differences in executive functioning across age-groups of offenders. As previously noted, the results indicate that younger inmates may be more amenable to cognitive remediation as compared to older inmates, particularly in the areas of decision making and processing speed. These results may have practical implications for interventions targeted to antisocial and correctional populations. For example, intervention efforts aimed at addressing executive deficits among these populations may be most effective when provided during early adulthood.

Although not a primary interest of the current study, I also found striking results when comparing baseline executive functioning across age groups, noting that the oldest subjects appeared to demonstrate the worst performance in executive functioning, which counters trajectories observed among non-correctional populations. This unexpected finding underscores the need for additional research on executive functioning among correctional populations, including factors which may be associated with executive remediation or decline among this population.

**Directions for Future Research**

In total, the current study indicates the need for additional research on the relations between executive functioning and recidivism among correctional populations of diverse ages. In terms of future research, it would be useful to extend the current findings by examining the potential efficacy of an evidence-based executive remediation program within a correctional setting. Such efforts should include comprehensive baseline and follow-up assessment of executive functioning, ideally including both self-report and performance task measures of
executive functioning, which would allow for greater understanding of cognitive changes that may be possible within this population given targeted remediation.

Based on the current study, there is evidence of a potential sensitive period for executive remediation among correctional populations. Therefore, future studies evaluating the impacts of cognitive remediation with antisocial and/or correctional populations should target individuals of varying age in order to assess whether cognitive gains differ across developmental stages.

Finally, additional research is needed to assess the short- and long-term impacts of cognitive remediation on antisocial populations. Studies evaluating the impacts of cognitive remediation among this population should also consider the potential effects of such remediation on individuals’ general functioning and behavior over extended periods and during periods of incarceration and post-release. Potential findings regarding the long-term or sustained benefits of cognitive remediation within this population is critical as such efforts may be futile if potential gains are not sustained and/or do not correspond with meaningful behavioral improvement.
REFERENCE LIST


Luther, J. B., Reichert, E. S., Holloway, E. D., Roth, A. M., & Aalsma, M. C. (2011). An exploration of community reentry needs and services for prisoners: A focus on care to limit return to high-risk behavior. AIDS Patient Care STDS, 25(8), 475–481. doi: 10.1089/apc.2010.0372


VITA

Dr. Danielle Nesi was raised in Westchester, New York. Before attending Loyola University Chicago, she attended the University of Wisconsin-Madison, where she earned a Bachelor of Arts in Sociology and a Certificate in Criminal Justice in 2010. From 2010-2014, Dr. Nesi attended CUNY John Jay College of Criminal Justice in New York City, where she earned a Master of Arts in Criminal Justice.

While at Loyola, Dr. Nesi collaborated on research projects with faculty across university departments and with community stakeholders focused on research in criminal justice settings and justice-involved populations. In addition, while at Loyola, Dr. Nesi collaborated with researchers at other universities, including Rutgers University and the University of Southern California, on numerous projects focused on justice-involved youth and young adults. During her time at Loyola, Dr. Nesi received several academic awards including acceptance to the university’s Research Mentoring Program in 2021 and the Heckler Scholarship in 2019. In 2017, she was awarded funding to attend the Inter-university Consortium for Political and Social Research (ICPSR) Summer Program in Quantitative Methods of Social Research.

Currently, Dr. Nesi is an adjunct instructor of psychology at Loyola University Chicago. Outside of teaching, she also consults a forensic expert witness regarding developmental factors contributing to antisocial behavior among justice-involved youth and adults. Dr. Nesi resides in Chicago, Illinois.