June 2012

Are Offence-Focused Correctional Rehabilitation Programs Affecting Inmates' Executive Cognitive Functions?

Erin H. Ross  
*The University of Western Ontario*

Supervisor  
Peter N.S. Hoaken  
*The University of Western Ontario*

Graduate Program in Psychology

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

© Erin H. Ross 2012

Follow this and additional works at: [https://ir.lib.uwo.ca/etd](https://ir.lib.uwo.ca/etd)

Part of the [Clinical Psychology Commons](https://ir.lib.uwo.ca/etd)

Recommended Citation  
[https://ir.lib.uwo.ca/etd/574](https://ir.lib.uwo.ca/etd/574)

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlswadmin@uwo.ca.
ARE OFFENCE-FOCUSED CORRECTIONAL REHABILITATION PROGRAMS AFFECTING INMATES’ EXECUTIVE COGNITIVE FUNCTIONS?

(Spine title: Prison Rehabilitation and Executive Cognitive Functions)

(Thesis format: Monograph)

by

Erin H. Ross

Graduate Program in Psychology

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Erin H. Ross 2012
The thesis by

Erin Haley Ross

entitled:

Are Offence-Focused Correctional Rehabilitation Programs Affecting Inmates’ Executive Cognitive Functions?

is accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy

Date ____________________________ ________________________________

Chair of the Thesis Examination Board
Abstract

There is a vast amount of data linking antisocial behaviours and deficits in executive cognitive functions (ECFs); however, there is a dearth of empirical research to address whether ECF abilities are being affected by correctional rehabilitation programming. Using a pre/post design, ECF performance of male inmates who completed a violence, family violence, or substance abuse rehabilitation program, was compared to ECFs of controls. Results indicated that across the six measures of ECF, including behavioural and self-report measures, only 2 of 10 variables showed a change across groups over time. Therefore, strong support for the notion that correctional programming is improving the ECF of offenders was not evident. However, the current sample displayed relatively unimpaired ECFs during the pre assessment, indicating fewer impairments compared with other forensic samples. In addition to measuring ECFs, changes in motivation for change and criminal attitudes were also examined. After completing programming, inmates displayed an increase in motivation for change, and a general decrease in criminal attitudes. Implications, recommendations, and areas of future research are discussed.

*Keywords*: executive cognitive functions, offenders, rehabilitation, motivation to change, criminal attitudes
Acknowledgments

First and foremost, I would like to acknowledge my supervisor Peter Hoaken. This, along with my many other overly zealous projects, would not have been imaginable without his support. It was his vision that allowed me the opportunity to do what I love best: spend endless hours hanging out in prisons in rural Canada. He is also commended for his continual efforts to advance my articulation of ideas; an onerous task indeed.

My prison adventures were greatly enhanced by the outstanding efforts of my dedicated research assistants, Kayla Truswell and Shawn Maronets. These two proved to be equally enthusiastic about prisoners, and were all around fabulous travelling companions. My Hoaken lab mates, Vanessa Pedden, Megan Hancock, and Jen Tapscott, provided continuous support by way of brainstorming, solution generation, and the not-to-be-underestimated, moral support. Special appreciation is extended to Jen who regularly dedicated her time to keeping me on the straight and narrow with my statistical analyses. In this regard, Pamela Seeds was also instrumental with helping to implement a statistical plan of attack. My Ph. D. committee, consisting of Greg Moran and Alan Leschied, was also invaluable throughout this process, particularly during times of setbacks and uncertainty.

Finally, I would like to recognize my beloved husband and partner, Quinn Ross. He has donned many hats during this process: sounding board, editor, motivational speaker, shoulder to cry on, colluder, and sober second thought. He has celebrated the highs and comforted during the lows. In the end, he deserves the fruit of this labour as much as I.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certificate of Examination</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Appendices</td>
<td>x</td>
</tr>
<tr>
<td>The Theory of Risk, Need, and Responsivity</td>
<td>2</td>
</tr>
<tr>
<td>Executive Cognitive Functions</td>
<td>4</td>
</tr>
<tr>
<td>Efforts to Ameliorate ECFs among Non-Forensic Populations</td>
<td>9</td>
</tr>
<tr>
<td>The Landscape of Current Correctional Rehabilitation Programming</td>
<td>13</td>
</tr>
<tr>
<td>Limitations of Forensic Rehabilitation Research</td>
<td>20</td>
</tr>
<tr>
<td>Measuring ECFs</td>
<td>23</td>
</tr>
<tr>
<td>Current efforts to measure ECFs of offenders</td>
<td>27</td>
</tr>
<tr>
<td>The Current Study</td>
<td>28</td>
</tr>
<tr>
<td>Method</td>
<td>32</td>
</tr>
<tr>
<td>Participants</td>
<td>32</td>
</tr>
<tr>
<td>Measures</td>
<td>38</td>
</tr>
<tr>
<td>Measures of ECFs</td>
<td>38</td>
</tr>
<tr>
<td>Impairment measures</td>
<td>38</td>
</tr>
<tr>
<td>Participation measures</td>
<td>41</td>
</tr>
<tr>
<td>Disability measure</td>
<td>43</td>
</tr>
</tbody>
</table>
List of Tables

Table 1. Mean (and Standard Error) of Demographic Information…………………… 34

Table 2. Number of Participants of each Ethnicity and Marital Status………………… 35

Table 3. Correlations between Pre and Post ECF variables………………………… 53

Table 4. Results of the Five Repeated-Measures (RM) Analyses Comparing Pre and Post Scores by Group (Programmers vs. Controls)…………………………………… 58

Table 5. Results of the Five Repeated-Measures (RM) Analyses Comparing Pre and Post Scores by Group (NSAP-M vs. VPP-M vs. Controls)………………………………. 63

Table 6. Mean (and Standard Deviation) of ECF Variables .............................. 72

Table 7. Pearson Correlation Coefficients between ECF Change Scores and Pre URICA Composite Score …………………………………………………………………… 79

Table 8. Pearson Correlation Coefficients between ECF Scores and Pre CSS-M Total Score ………………………………………………………………………………………… 83
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.</td>
<td>IGT Net 4 $t$ scores for Programmers and Controls pre and post rehabilitation</td>
<td>59</td>
</tr>
<tr>
<td>Figure 2.</td>
<td>BRIEF Composite $t$ scores for Programmers and Controls pre and post rehabilitation</td>
<td>61</td>
</tr>
<tr>
<td>Figure 3.</td>
<td>IGT Net 4 $t$ scores for NSAP-M, VPP-M, and Controls pre and post rehabilitation</td>
<td>65</td>
</tr>
<tr>
<td>Figure 4.</td>
<td>BRIEF Composite $t$ scores for NSAP-M, VPP-M, and Controls pre and post rehabilitation</td>
<td>67</td>
</tr>
<tr>
<td>Figure 5.</td>
<td>URICA Composite scores for Programmers and Controls pre and post rehabilitation</td>
<td>78</td>
</tr>
<tr>
<td>Figure 6.</td>
<td>CSS-M Total scores for Programmers and Controls pre and post rehabilitation</td>
<td>81</td>
</tr>
</tbody>
</table>
## List of Appendices

<table>
<thead>
<tr>
<th>Appendix A. Ethics Approval Notice</th>
<th>129</th>
</tr>
</thead>
</table>
Are Offence-Focused Correctional Rehabilitation Programs Affecting Inmates’ Executive Cognitive Functions?

Incarceration continues to be the primary means by which western society punishes those who have committed criminal acts. The premise is that incarceration may deter individuals from future criminal behaviour. However, empirical evidence suggests that the characteristics of incarceration, features such as loss of freedom as a result of long prison sentences, reduced living and social conditions, or stigma associated with convict status, are insufficient deterrents for many offenders (Andrews, Bonta, Gendreau, & Cullen, 1990; Gendreau, Groggin, & Cullen, 1999; McGuire, 2002). In fact, a meta-analysis illustrated that longer incarceration sentences may actually result in up to 3% higher recidivism rates (Gendreau et al., 1999). Others have remarked that incarceration periods that focus primarily on punishment, and are void of rehabilitation interventions also appear to have a negative impact on the successful community reintegation of offenders (Andrews et al., 1990; Lipsey & Cullen, 2007). Many countries have come to recognize that a pure punishment model, relying exclusively on the aversive aspects of lengthy prison sentences and lack of rehabilitation, is insufficient for reducing recidivism or supporting successful reintegration. Thus, these countries have moved away from subscribing exclusively to a punishment model; instead, the past four decades have seen various types of correctional rehabilitation programs implemented.

The purpose of the subsequent sections is to present the theory most commonly used in the development of Canadian correctional rehabilitation programs. Within the description of this theory, the concept of executive cognitive functions (ECFs) will be introduced, as well as their relevancy to offenders, with the goal of demonstrating that deficits in ECFs relate to criminal behaviour. Rehabilitation research from non-offender
populations will then be reviewed to illustrate innovations in ameliorating ECF deficits. Within these studies special attention will be paid to the methods used to improve deficits and their effect on behaviour, as well as the methods used to evaluate efficacy of rehabilitation efforts. Later, the current landscape of Canadian correctional rehabilitation will be presented, along with research documenting the reported efficacy of these programs. Lastly, limitations within the body of correctional rehabilitation literature will be discussed with the purpose of emphasizing areas where improvement is warranted. These improvements are then incorporated into the description of the current dissertation which examines ECFs and correctional programming among Canadian inmates. As stated, the discussion will begin with an introduction to the theory that has influenced most North American and European correctional rehabilitation programs.

**The Theory of Risk, Need, and Responsivity**

In the 1970s, critics of the rehabilitative approach argued that little or nothing worked to affect re-offending rates, regardless of program type, or issue targeted. The consequence of this criticism was the production, over the next two decades, of a large number of studies demonstrating ‘what works’ to rehabilitate offenders (e.g., Andrews et al., 1990; Lipton, 1995; Wells-Parker, Bangert-Drowns, McMillen, & Williams, 1995). Out of this large body of literature came the theory widely known as the “principles of effective correctional intervention” set forth by Andrews, Bonta, Gendreau, and their colleagues (Andrews, 1995; Andrews & Bonta, 2006; Andrews, Bonta, & Wormith, 2011; Gendreau, 1996; Gendreau, Smith, & French, 2006). The theory, which acts as a blueprint for implementing successful treatment programming, has evolved into eighteen detailed principles that characterize effective treatment programs. The principles relate to intensity of services, program approach, facilitator style, techniques offered, and content
targeted. In essence, the theory states that the principles of risk, need, and responsivity, often abridged to RNR, must be addressed within correctional programming (Dowden & Andrew, 2000).

In brief, the risk principle states that the level of treatment must be appropriately matched to the risk level of the offender. Risk can be determined based on both static risk factors, such as age, gender, and offending history, and dynamic risk factors, such as housing and substance abuse history. The need principle speaks to both the criminogenic needs and non-criminogenic needs of the offender. Criminogenic needs refer to offender characteristics that, when changed, decrease criminal behaviour (Andrews & Bonta, 1998), such as antisocial attitudes or beliefs, and substance dependency. Changes in non-criminogenic needs, such as self-esteem, do not appear to impact criminal behaviour; therefore, are not deemed relevant factors to reducing recidivism.

The responsivity principle is directly concerned with the characteristics of program delivery. The principle states that the styles and modes of service used within a treatment program should be matched to the learning styles of offenders (Andrews & Bonta, 1998). Both general and specific responsivity considerations are encompassed within the responsivity principle. General responsivity states that cognitive behavioural and social learning approaches are the most effective methods of bringing about positive behavioural change. Included are techniques such as modeling, graduated practice, rehearsal, role playing, reinforcement, and cognitive restructuring (Andrews et al, 1990). Specific responsivity focuses on the individual characteristics of offenders, such as motivation, and cognitive abilities. This principle argues that program outcome will be most successful when steps are taken to address these types of offender difficulties. Dowden and Andrews (2000) highlighted the importance of some such factors, stating
that “offenders who have limited cognitive and problem-solving skills require more structured and concrete programs than offenders who have high levels of interpersonal, self-reflective, and verbal skills” (p. 453). These authors allude to the reality that a number of offenders experience deficits in their cognitive and problem-solving abilities, and they advocate for program variations which attend to these issues.

The concept of ‘problem-solving’, also often referred to as ‘cognitive skills’, is a phrase common within the forensic rehabilitation literature. Typically these phrases are interpreted as referring to the ability to generate prosocial thoughts, attitudes and actions. As the reader will see, current correctional programming attempts to target these ‘cognitive skills’ with the belief that amelioration will result in decreased recidivism. In support of this notion is a body of literature outlining the presence of cognitive deficits among offenders. However, this literature incorporates a more complex definition of ‘cognitive skills’, referring to a group of ECFs. What follows is a brief description of specific ECFs, as well as a review of the research linking them to aggressive, antisocial, and criminal behaviour.

**Executive Cognitive Functions**

Forensic rehabilitation literature often uses the phrase ‘cognitive skills’ to refer to prosocial thoughts, attitudes and actions. However, within the neuropsychological literature the term ‘cognitive’ is used to describe how the brain perceives, processes, and uses information from the environment, without the influence of beliefs or attitudes. Further exploration of the neuropsychological concept of cognitive skills is necessary to appreciate the relevancy to offender rehabilitation. Psychological researchers, many based in neuropsychology, have identified a specific constellation of cognitive skills or abilities which neuropsychologists refer to as ECFs. ECFs can be thought of as a group
of abilities involved in the planning, initiation, and regulation (i.e., maintaining or altering) of goal-directed behaviour (Luria, 1980; Roberts, Robbins & Weiskrantz, 1998).

Among the specific skills identified as ECFs are the skills of cognitive flexibility (also relating to set-shifting and set maintenance), strategy formation, attention, working memory, response monitoring, and inhibition (Blud, Travers, Nugent & Thornton, 2003; Lezak, Howieson & Loring, 2004; Olvera, Semrud-Clikeman, Pliszka, & O’Donnell, 2005; Pennington & Ozonoff, 1996; Petrides, 1990). Among many conceptualizations of ECFs, these abilities often represent the ‘cold’ component of ECFs because their corresponding cognitive processes typically do not involve emotional arousal and can be considered more mechanistic or logically based (Grafman & Litvan, 1999). However, when the operation of ECFs relate more to emotional experiences and involve the regulation of affect and motivation, such as when the experience of reward and punishment and other consequences are employed, they are said to represent the ‘hot’ component of ECFs (Buelow & Suhr, 2009; Seguin & Zelazo, 2005). When applied together, both the ‘cold’ and ‘hot’ components are thought to support the more complex tasks of planning, decision-making, and problem-solving. These latter terms have often been used to oversimplify the concept of ECFs. It is not uncommon to encounter the use of a single term or phrase such as ‘problem-solving’ or ‘executive functioning’ to represent the entire constellation of ECFs, failing to recognize the independent components (Suchy, 2009). For the most part, this represents the outmoded theory of a unified executive construct (Miyake et al., 2000). However, implications of this manner of conceptualizing ECFs continue to prevail. As the reader will find in the next paragraph, and throughout this dissertation, much of the current knowledge acquired about specific ECFs originated from research examining these more general constructs. This dilemma is
present within older neuropsychological literature, and the forensic and rehabilitation
literature, as well as, but less commonly, within current neuropsychological literature.
Throughout the current dissertation these general terms have been replaced with specific
ECFs whenever possible; however, when no description was provided within the cited
literature, general descriptors were used.

ECFs are said to involve several brain regions, but most notably the prefrontal cortex
(Paschall & Fishbein, 2002), and a number of sub-cortical pathways (Koechlin, Corrado,
Pietrini, & Grafman, 2000; Monchi, Petrides, Strafella, Worsley, & Doyon, 2006). Much
of what is known about the manifestations of ECF deficits came out of the
neuropsychological literature exploring brain injury or abnormality (e.g., Bechara,
Damasio, Tranel, & Anderson, 1998; Koski & Petrides, 2001; Marsh & Martinovich,
2006). Many individuals with lesions to the prefrontal cortex exhibit difficulties in the
general areas of planning and decision making, as well as deficits in social behaviours,
impulsivity, and emotional dysregulation (Masterman & Cummings, 1997; Tateno, Jorge,
electroencephalography, and neuroimaging studies involving brain injury or lesion
suggest an association between aggression and frontal lobe dysfunction (Brower & Price,
2001). Although offenders are up to 60 times more likely than non-offenders to have
experienced a traumatic head injury (Mullin & Simpson, 2007), brain injury or
abnormality status is not necessary to link ECF impairments and antisocial behaviour.

There is evidence to support the notion that antisocial behaviour among adolescents
and adults is associated with deficits in ECFs. Populations of male and female
incarcerated young offenders (Enns, Reddon, Das, & Boukos, 2007; Olvera et al., 2005),
as well as non-incarcerated aggressive adolescents (Giancola, Mezzich & Tarter, 1998;
Seguin, Pihl, Tremblay, Boulerice & Harden, 1995), and aggressive children (Raaijmakers et al., 2008) have demonstrated impairments on various ECFs. Incarcerated male and female adults have also displayed a range of ECF deficits (Hoaken, Allaby, & Earle, 2007; Marceau, Meghani, & Reddon, 2008; Yechian et al., 2008), including deficits of inhibition, cognitive flexibility, and the ability to predict future consequences (Broomhall, 2005), compared to non-incarcerated control groups. In addition, poor performance on measures of impulsivity and cognitive flexibility predicted offenders who were more likely to have committed a larger number of violent offences (Hancock, Tapscott, & Hoaken, 2010). Unfortunately, methodological variability within this literature, such as inconsistent conceptualization of ECFs, or the use of divergent ECF measures, including those lacking normative data, or the use of inconsistent scoring methodologies, has made it difficult to make reliable conclusions regarding the degree of impairment present among offenders across the various ECF domains. Thus, prevalence rates of impairment among offenders vary widely, with estimates as high as 80% (see Muscatello et al., 2011; Reiss, Miczek & Roth, 1994). Despite these obstacles, there appears to be a relationship between ECFs and antisocial, or criminal, behaviour.

Intuitively, it makes sense that having deficits in functions that support planning, initiation, and regulation of goal-directed behaviour would impact an individual’s actions. However, one may question how these deficits might result in criminal behaviour. If an individual is not equipped with appropriate problem-solving abilities, navigating through situations with potential criminal components (e.g., solving financial strain, or navigating a provocative interpersonal encounter) will inevitably be more problematic. For example, it is not difficult to imagine that if an individual’s general ability to formulate strategies is impaired, then his or her options may seem limited. In addition, if an individual is unable
to accurately monitor the outcome of previous decisions (or strategies) because of working memory deficits, then future decisions may be poorly constructed. Increased impulsivity or the inability to inhibit reward seeking behaviour may act to additionally impede strategy formation or response monitoring, resulting in a repetitive cycle of poorly thought out behaviours. Together, these interactions may result in continued criminal involvement. Clarke and Felson’s (1993) rational choice perspective model helps flush out how the aforementioned deficits may lead to repeated criminogenic activity. According to the model, an offender is a calculating decision maker who weighs positive and negative outcomes before committing an offence. However, the extent to which offenders, or individuals in general, are capable of gathering and utilizing all relevant information when making decisions is questionable; instead, other constraints, such as time, or as suggested here, ECFs, result in decision makers having limited information that can be employed (Wilson, Attrill & Nugent, 2003). As follows, deficits in ECFs could adversely influence the type, quantity, and/or quality of information used in the problem-solving process.

The above review of literature supporting the relationship between criminal behaviour and deficits in ECFs among offenders, as well as the review of the RNR framework for effective correctional rehabilitation programming, suggests that forensic researchers and practitioners have an empirically supported blueprint for creating appropriate and effective correctional programming that could ameliorate ECF deficits and reduce recidivism. However, the translation between empirical evidence and practice can be less than straightforward. Less is known about which empirically supported procedures are best suited to ameliorate those deficits that appear to be associated with crime and other undesirable behaviours. There appears to be two main barriers to this knowledge.
Firstly, literature documenting empirically supported methods of ameliorating ECF deficits is relatively new, and is predominately focused on non-forensic populations. Secondly, limitations of the correctional rehabilitation research, both in terms of methodological issues, as well as measurement issues, also confound the knowledge in this area. Although both of these barriers will be discussed within the current dissertation, the reader is first introduced to innovations in ameliorating ECFs among non-forensic populations, with the purpose of presenting an exemplar that may be used to support the notion that behavioural change is possible if ECF deficits of offenders are appropriately targeted.

**Efforts to Ameliorate ECFs among Non-Forensic Populations**

Currently, there is a dearth of forensic research empirically supporting the notion that ameliorating ECFs may directly improve an offender’s behaviour. However, examination of the rehabilitation efforts targeting ECFs among another population, specifically those diagnosed with schizophrenia, suggests that rehabilitation can improve ECFs, as well as functioning in other behavioural areas.

There are significant similarities between the neuropsychological deficits experienced by individuals diagnosed with schizophrenia (McGurk, Mueser, DeRosa, & Wolfe, 2009; Merriam, Medalia & Wyszynski, 1993), and those deficits observed in some offenders, particularly with regards to ECFs. Specific deficits which are common among individuals with schizophrenia include deficits of attention, memory, cognitive flexibility, strategy formation, and inhibition (Goldsamt, Barros, Schartz, Weinstein, & Iqbal, 1993; Green, 1998; Merriam et al., 1993). In schizophrenia, deficits in ECFs have increasingly been thought to be both a manifestation of the illness, and predictor of disability (Arciniegas, McAllister, & Kaufer, 2007; Flashman & McAllister, 2007). Although some
individuals with schizophrenia show no neurological impairments (Allen, Goldstein, & Warnick, 2003), ECF deficits are generally considered a prevalent comorbid condition of schizophrenia (McGurk et al., 2009). In addition, deficits in ECFs are considered a strong predictor of poorer psychosocial functioning (McGurk & Mueser, 2004).

Because of the role of ECF deficits within the expression of schizophrenia, significant efforts have been made to rehabilitate these deficits. Some encouraging rehabilitation studies are briefly reviewed below, highlighting how researchers have measured and targeted specific ECFs with the goal of improving the specific function, as well as more generalized behaviours. These studies are featured with the goal of providing a foundation for the notion that targeting offenders’ ECF deficits may also impact their criminal behaviours.

Several primary outcome studies have illustrated that specific ECFs can be improved with long lasting, as well as generalized, outcomes. Twenty-six adolescents with early onset psychosis received either 30 hours of individual cognitive remediation training (CRT), consisting of multiple individualized teaching components targeting cognitive differentiation, attention, memory and social perception with the goal of improved general problem-solving, or a standard treatment of schizophrenia psycho-education (Ueland & Rund, 2005). Participants’ functioning was measured at baseline, post-treatment, and at 1 year follow-up using specific measures of attention and memory. The Wisconsin Card Sorting Test (WCST) was used as a general ‘executive functioning’ measure with the variables of interest targeting inhibition, cognitive flexibility, and response monitoring. The CRT group displayed specific improvements in all measured ECFs, compared to the psycho-education group, even at 1 year follow-up. Participants
also showed improved levels of psychosocial functioning, as indicated by informant report.

Similarly, other individualized rehabilitation programs with modules targeting specific ECFs have illustrated positive results. The Neuropsychological Educational Approach to Rehabilitation (NEAR; Medalia, Revheim, & Casey, 2002), focused on improving basic cognitive elements such as categorization, working memory, and attention using computer software programs, with the goal of improving individuals’ ability to problem solve, think critically, and employ cognitive flexibility. More advanced training involved computer simulations of real life problems in an attempt to provide transfer of knowledge (e.g., simulation of managing a retail shop or planning a family vacation). Researchers have reported improved attention, working memory, and general problem-solving skills (Medalia, Revheim & Casey, 2001), which have persisted beyond the training sessions (Medalia et al., 2002). More recently, a randomized control trial was conducted including 40 individuals with schizophrenia (Hodge et al., 2010). Participants were assigned to either Waitlist control or NEAR group, and were tested at baseline, post-treatment, and at 4 month follow-up using various measures of ECFs including the Trail Making Test, an element of the Delis-Kaplan Executive Functioning System (D-KEFS), an attentional task, and a verbal learning task, as well as various measures of psychosocial functioning including clinician and self-reports. Following 15 weeks of treatment, NEAR participants showed improvements in memory, attention, and cognitive flexibility, persisting for up to 4 months, with low to moderate effect sizes. These improvements were accompanied by a significant improvement in psychosocial functioning.

Importantly, meta-analytic work also supports that rehabilitation targeting ECFs can
improve several ECFs, can generalize to improvements in other everyday functions, and can persist beyond the treatment period. McGurk, Twamley, Sitzer, McHugo, and Mueser (2007) conducted a meta-analysis of 26 random control trials of cognitive rehabilitation including 1,151 individuals with schizophrenia. Overall, rehabilitation was associated with significant improvements in attention (effect size $d = 0.41$), working memory ($d = 0.52$), and reasoning and problem-solving ($d = 0.47$), as well as improvement in psychosocial functioning ($d = 0.35$). For those studies that reported ECF data at follow-up, improvements were generally in the medium effect size range an average of 8 months post treatment.

This body of literature not only describes methods used to improve individual ECFs, such as cognitive flexibility, but also illustrates that targeting these functions can impact the broader category of psychosocial functioning, such as general life skills, employment, disability, as well as social behaviour and adjustment (McGurk et al., 2007). As an extension of this conclusion, one could propose that targeting the ECF deficits already established among many offenders could positively impact future undesirable or criminal behaviours (for a further discussion see Ross & Hoaken, 2010). A logical query following this argument is whether or not current correctional rehabilitation is indeed targeting, or measuring, ECF deficits of offenders. This question is paramount to the current dissertation and will be addressed throughout. However, before this question can be addressed, it is helpful to appreciate the current landscape of forensic rehabilitation.

To assist the reader less familiar with forensic rehabilitation, the philosophy of rehabilitation efforts is first reviewed, followed by a description of current rehabilitation programming, as well as a review of the effectiveness of forensic rehabilitation, and a discussion of the limitations of the forensic rehabilitation literature.
The Landscape of Current Correctional Rehabilitation Programming

Substantial resources have been employed in the creation, implementation, and study of correctional rehabilitation programming. To illustrate, Canada spent nearly 26 million dollars on correctional rehabilitation programming in 2007 (Nafekh, Allegri, Stys, & Jensen, 2009). Of the 2007 programming expenditures, nearly 47% of that cost was allocated to three rehabilitation program categories: 1) substance abuse programs, 2) violence prevention programs, and 3) family violence prevention programs. These program categories appear to represent the primary means of rehabilitation of male inmates in Canada. These programs are typically prescribed to incarcerated males, coming from both Aboriginal and non-Aboriginal origins. According to the National Correctional Programs Referral Guidelines (Correctional Service Canada, 2009), correctional program selection is determined by the level of actuarial risk of an inmate, as calculated by the use of various risk assessment measures that attach specific statistical weighting to different offender variables, in conjunction with offender’s needs.

The principal goal of each of the three forensic rehabilitation programs is reduction in re-offending. Examination of the curriculum suggests that reduction in recidivism is targeted by focusing on improvements in goal setting, planning, consequence awareness, communication, problem-solving, self-control, self-management, self-monitoring, assertiveness, understanding the link between thoughts and actions, and criminal attitudes, beliefs, and associations (Correctional Service Canada, 2004; Reintegration Programs Division, 2009). Of note, these are some of the skills that are often loosely referred to as ‘cognitive skills’ within the forensic literature; as well, many of these skills also appear to overlap with some of the functions defined as ECFs. Targeting this group of skills originated from the tenets laid out by the cognitive skills program (Ross &
Fabiano, 1985) coined more than 25 years ago. Arguably all North American and European correctional rehabilitation programs, ranging from anger management, substance abuse, to violence prevention programs, have been based on the original cognitive skills program. The overarching goal is to increase prosocial thoughts and actions in an effort to reduce recidivism. The implied assumption is that improving the aforementioned skills will ultimately decrease criminal behaviours. For example, impulsivity could be reduced by teaching consequential thinking, and rigid thinking could be minimized by teaching offenders creative thinking skills to provide them with prosocial alternatives in responding to interpersonal problems. Although the cognitive skills program (under the name Alternatives, Associates, and Attitudes in its current incarnation) is still an active rehabilitation program within the North American correctional system, there has been a move away from general skills programs, towards offence-focused rehabilitation programs which adhere more to the RNR principles.

The current dissertation focused on three offence-focused rehabilitation programs: substance abuse, violence prevention, and family violence prevention programs. For these three categories of correctional programming, there are multiple levels of program intensity available (i.e., high, moderate, and maintenance). As can be gleaned, high intensity refers to programs that are longer in duration and meet more frequently, compared to moderate intensity programs; whereas, maintenance programs act as one to four session boosters when high or moderate intensity programs have previously been completed. Assigned program intensity is primarily determined by the inmate’s level of risk, with increased risk resulting in assignment to a higher intensity of program. When an inmate displays multiple needs, high intensity programs take priority. An offender can be assigned to only one high intensity program, but where multiple needs are
identified, an inmate can be assigned additional programs, but at lower intensity. Thus, within correctional facilities, moderate intensity programs appear to be the most prevalent and frequent. To illustrate, over a 3 year period, 660 Canadian inmates were enrolled in a high intensity version of the above mentioned programs (targeting violence or substance abuse), while 2,187 inmates were enrolled in the moderate intensity versions of these same programs (Reintegration Programs Division, 2009). As a result, it seems reasonable that a substantially greater proportion of inmates are being subjected to moderate intensity programming of the violence, family violence, or substance abuse variety; consequently, the current dissertation focused specifically on the moderate intensity versions of these programs. Although the family violence prevention program represents one of the more commonly prescribed programs overall, fewer than 100 inmates enroll in this program per year across Canada (Reintegration Programs Division, 2009). Thus, the current dissertation focused primarily on the substance abuse and violence prevention programs within several analyses. A brief description of the moderate intensity programs will follow, as described in the Correctional Service Canada (CSC) documents by the Reintegration Programs Division (2009) and Nafekh, Allegri, Stys and Jensen (2009).

The national substance abuse program (NSAP) was developed for male offenders whose crimes were related to substance abuse and who demonstrated a need for substance abuse treatment. Because of the high prevalence of drug abuse among offenders (Mumola, 1999), substance abuse treatment programs may be the most common form of rehabilitation offered to this population. According to CSC, approximately 50% of their offenders exhibit a direct link between their substance use and their criminal behaviour (Reintegration Programs Division, 2009). The moderate intensity program was designed to cover four modules over 26 group sessions with one individual session, delivered 4 to 5
times per week, typically over a two month duration. The modules reflect the stages outlined by the Transtheoretical Model of Change (Prochaska, DiClemente, & Norcross, 1992). More specifically, the modules focus on contemplating the idea and actions of change, identifying risk, self-management, goal-setting, problem-solving, as well as understanding attitudes and beliefs related to substance use, and relapse prevention. The program has been offered in penitentiaries across Canada, as well as in other probation agencies (e.g., Norway, Switzerland). Between 2006 and 2008, 1,646 Canadian inmates were enrolled in the moderate intensity NSAP (hereafter referred to as ‘NSAP-M’) achieving an 84% completion rate (Reintegration Programs Division, 2009). In a large scale \(N = 48,601\) internal CSC program evaluation study including incarcerated males, predominately non-aboriginals, between the years 1997 and 2007, with a three year follow-up period \(n = 23,724\), NSAP-M resulted in reductions in re-offending in the range of 27% to 46% among non-Aboriginal participants (Nafekh et al., 2009).

The violence prevention program (VPP) was developed for male offenders who were considered at risk to commit future violent crimes and for those offenders who had already committed at least two violent offences, or had been convicted of murder. The moderate intensity program was designed to cover eight modules over 36 group sessions with three individual sessions, delivered 3 to 5 times per week, typically over a three month duration. The modules target skill improvement in self-management, emotion management, problem-solving, goal-setting, communication, as well as understanding attitudes and beliefs related to violence. The program has been offered in penitentiaries across Canada. Between 2007 and 2008, 243 Canadian inmates were enrolled in the moderate intensity VPP (‘VPP-M’) achieving a 95% completion rate (Reintegration Programs Division, 2009). Significant re-offending reductions in the range of 41% to
52% have been attributed to VPP-M in non-Aboriginal offenders (Nafekh et al., 2009).

The moderate intensity family violence prevention program (abbreviated in the forensic literature as ‘MIFVPP’) provides treatment for male inmates who displayed risk for violence in their intimate relationships. A male offender must have committed at least one confirmed act of family violence against an intimate partner or family member. Actuarial data (i.e., statistical information obtained through the collection of offender information regarding risk variables) is used to determine risk of future family violence. The program was designed to cover six modules over 29 group sessions with three individual sessions, delivered 2 to 4 times per week, typically over a two month duration. The modules attempt to increase inmates’ motivation and awareness of family violence, as well as improving skills in emotion management, communication, social skills, relapse prevention, and developing healthy relationships. The program has been offered in penitentiaries across Canada and the United Kingdom. Between 2006 and 2008, 298 Canadian inmates were enrolled in the MIFVPP achieving an 81% completion rate (Reintegration Programs Division, 2009). Significant re-offending reductions in the range of 36% to 57% for non-Aboriginal offenders have been attributed to MIFVPP (Nafekh et al., 2009).

Although each of the program categories target a different problem area or offence type (i.e., substance use, general violence, or family violence), examination of the program descriptions reveal considerable curriculum overlap, particularly in the areas of planning, goal setting, problem-solving, and self-monitoring skills. The foundational skills are taught within each program, with the additional offence-focused material incorporated. For example, within the VPP, participants are taught how to set appropriate prosocial daily goals and plan how to obtain those goals, with the added emphasis of
avoiding violence in achieving those goals. In addition to general cognitive skills, elements that are said to address offenders’ motivation to change also appear across program curricula (Reintegration Programs Division, 2009). Program delivery also shares many similarities. Programs are conducted using a structured group format with lecture type discussions to disseminate information, and employ cognitive-behavioural techniques to practice skills. In some ways these programs share similarities with the previously described rehabilitation programs for individuals with schizophrenia. For example, they operate using regular training sessions over several weeks, targeting similar outcomes of improving general problem-solving abilities using practice exercises. Where the programs seem to differ is in the way they target general problem-solving. It is in this domain that the schizophrenia rehabilitation literature describes measuring and targeting specific ECFs; within the forensic rehabilitation literature, conversely, the focus appears to be more on concepts such as antisocial thoughts, attitudes, and actions. This focus on antisocial behaviour may have come as a result of, or contributed to, the use of recidivism as the primary outcome measure among forensic rehabilitation programming. Thus, much of the literature describing the efficacy of correctional programming is instead limited to recidivism data, and thus often fails to include other potentially relevant variables, such as ECFs.

Program effectiveness has been almost exclusively determined by re-offending rates of program completers, and compares those rates to offenders not enrolled in programming, typically one to two years after program completion. Internal investigations, such as Nafekh et al. (2009) typically yield higher reductions; however, reported success of correctional programs, without distinction of type or focus, has been generally positive, if somewhat variable. Most of the evidence supporting the
relationship between forensic rehabilitation programs and the reduction of recidivism has come primarily from studies of incarcerated adult males out of Canada, the United States of America (USA), Britain, and New Zealand. Meta-analyses have been particularly important in depicting overall recidivism reductions. Dowden and Andrews’ (2000) meta-analytic work examined 35 studies of male offenders, consisting of 70% adults. They reported that correctional interventions (i.e., group or individual treatment) reduced recidivism by 12%. When 18 meta-analyses published between 1985 and 2000 were examined, McGuire (2002) found that correctional programming reduced recidivism by 10%, results mirrored by Lipsey and Cullen’s (2007) meta-analytic review. Whereas another meta-analytic work, predominantly examining program outcome studies originating from the USA, suggested reductions of 8% are more typical (Aos, Miller, & Drake, 2006). Primary outcome studies typically report similar reductions and effect sizes; however, some have been more variable, reporting effects on both the high and low extremes (e.g., Anstiss, 2003). Collectively the meta-analytic and primary outcome studies reveal that programs that more closely adhere to the RNR principles are associated with greater program effectiveness.

Programs that comply with the principles of effective correctional intervention, or RNR principles, achieve larger reductions in recidivism (Gendreau et al., 2006; also see Lipsey, 1992; Lowenkamp, Latessa, & Smith, 2006), and result in lower rates of institutional misconduct (French & Gendreau, 2006). Specifically, studies have reported average reductions in recidivism of 28% to 30% (Andrews et al., 1990; Dowden & Andrews, 1999). Furthermore, the applicability of these principles has been established across diverse offender populations, including male offenders (Andrews & Bonta, 2006; Gibbons, 1999), female offenders (Dowden & Andrews, 1999), violent offenders
PRISON REHABILITATION AND EXECUTIVE COGNITIVE FUNCTIONS


Although statistically rigorous meta-analyses assist with interpretation of the vast number of primary outcome studies, some methodological issues common within the large body of literature examining forensic rehabilitation programs result in problems that even the most sophisticated meta-analysis is unable to overcome. Some of these issues are outlined below, with the goal of providing guidance for future forensic programming research. These issues were strongly considered throughout the planning of the current dissertation and will also be discussed in later sections.

Limitations of Forensic Rehabilitation Research

Methodological problems within the body of forensic rehabilitation literature may unwittingly bias outcomes. Methodological problems have been identified among both positive and negative outcome evaluations of forensic rehabilitation programs (Ogilvie, Stewart, Chan, & Shum, 2011), particularly problems relating to comparison groups. The importance of including an appropriate control group within a treatment study may seem obvious to the reader (e.g., for the purpose of eliminating conjecture that change occurred due to unmeasured variables); however, many forensic researchers fail to include a comparison group, while others include poorly matched comparison groups (e.g., Hodel & West, 2003; Hollin et al., 2008; Mullin & Simpson, 2007). As illustrated by Wilkinson (2005), a comparison control group should be as similar to the experimental group as possible in order for a true evaluation of a treatment to occur. Unfortunately, among forensic studies, outcomes from treated individuals are often compared to outcomes from individuals who did not qualify for treatment, such as those deemed less at risk of re-offending. As a result, researchers may run the risk of underestimating the effectiveness
of treatments, by comparing higher risk treated individuals to lower risk untreated individuals.

Related to the problem of selecting the most appropriate participants to include in a treatment study, is the issue of program dropout rates. Dropout rates of up to 50% are not unheard of within correctional interventions (Losel, 2001). Not surprisingly, program dropout rates have been associated with lower offender motivation for change, a complex offender characteristic too often disregarded (Wilkinson, 2005). Dropouts result in distinct consequences for researchers and practitioners. For researchers, decisions must be made regarding whether or how to include dropouts within a treatment evaluation; decisions that vastly affect outcomes. For example, including only program completers may over represent positive, effective program aspects, whereas, including both completers and non-completers may skew results in the opposite direction (McMurran & McCulloch, 2007). For practitioners, program dropouts mean that individuals deemed in need of rehabilitation are not being exposed to the appropriate information. Increased program dropouts may represent an issue of low motivation, or readiness, for change. Offenders’ level of readiness for change has been linked to both treatment completion, as well as treatment outcome (Baxter, Marion, & Gouguen, 1995). Addressing the issue of dropout requires evaluation and understanding of offenders’ motivation. Together, these observations draw attention to the need for appropriate comparison groups, as well as the consideration of offenders’ level of motivation.

The body of literature examining correctional programming also brings to the forefront the question of whether using recidivism rates as the sole outcome measure is most appropriate. Traditionally, forensic research examining the effects of rehabilitation published outside of correctional institutions often restricts outcome data to recidivism
rates. This may be in part because reducing re-offending is the ultimate goal of these programs, and appears most relevant to public safety. However, implied in this type of evaluation is that programming has a direct impact on recidivism. Important to our conceptualization of program effectiveness, Friendship, Falshaw, and Beech (2003) discussed recidivism as a long-term goal, with several short-term, and more immediate “intermediary treatment targets” being affected by programming (p. 119). As follows, it may in fact be these intermediary treatment targets that are impacting recidivism, or in the least, contributing to the impact. Measuring how rehabilitation programs influence recidivism is important, but does not help researchers understand what mechanisms of change are involved. In addition to a host of other variables, ECFs could represent important intermediary treatment targets; however, consistent and appropriate measurement of these constructs within correctional facilities is required to address this question.

Discussions of intermediary treatment targets, as opposed to recidivism rates, within forensic rehabilitation research has included measurements of relevant constructs such as treatment motivation, social information processing, criminal attitudes, social skills, work behaviour, and self-control in risky situations (e.g., Andrews & Bonta, 1998; Bettman, 2000; Hanson & Harris, 2000). Measurement and reporting of changes of such targets can be seen in some more detailed internal correctional publications (e.g., Nafekh et al., 2009). Special attention has been paid to the evaluation of program specific attitudes, criminal attitudes, and beliefs of offenders before and after program completion. Many of these measurements rely solely on offender self-report, or generic informant reports which measure program performance, treatment gain, and participant progress. Constructs such as attitude change and participant progress may be important
intermediary treatment targets; however, the current author proposes that the measurement of ECF deficits should also be added to the list of relevant and beneficial pre/post program measurements. However, deciding how best to assess ECF deficits in an accurate and valid manner is a complex endeavor that the literature has yet to indicate a gold standard (Chan, Shum, Touloupoulou, & Chen, 2008). Before the assessment of ECFs can be adequately incorporated into forensic research, understanding how best to measure these functions must be appreciated.

**Measuring ECFs**

Researchers have examined the question of how to most appropriately measure ECFs from a variety of perspectives. Some have taken perhaps an overly simplistic approach and discussed measuring ECFs based purely on subjective versus objective characteristics (i.e., perception of deficits versus actual performance). Others, more recently, have discussed utilizing a model that incorporates both performance and function of the individual, and attempts to capture relevant aspects of other theoretical models and various modes of measuring ECFs. Together these models provide a framework for selecting the most appropriate methods of measuring ECFs. What follows is a brief description of these models and how they were merged together to inform measurement selection for the current dissertation.

As previously discussed, the concept of ‘problem solving’ has been used synonymously with the concept of ECFs within some of the forensic literature. In one model, assessment of this concept has been divided into two sub-groups respectively associated with ‘process’ and ‘outcome’ (McGuire, 2001). Process measures are designed to access general cognitive and behavioural activities that facilitate solving problems. They rely primarily on self-report, and provide information regarding
individuals’ perceptions of how they approach various challenges. By contrast, outcome measures, under this definition, entail the assessment of performance or problem-solving competence, as judged by the products of the activity. One can make a similar distinction when thinking about evaluating ECFs by categorizing measures as subjective (i.e., process measures) or objective and based on performance (i.e., outcome measures). This model implies that there is relevant, yet distinct, data to be obtained by both subjective self-report process measures and objective performance outcome measures. This distinction begins to sort out the various categories of measures a researcher should include; however, within each of these categories there exists a multitude of measures claiming to assess ECFs.

The existence of multiple ECF measures does not necessarily represent disagreement among ECF researchers; instead this again highlights that there are several distinct ECFs, and that ECFs are not represented by a single construct. Factor analytic studies confirm this later problem, indicating that measures of ECFs do not load onto a single, overarching executive construct (Royall et al., 2002). Because ECFs include a collection of abilities previously defined as cognitive flexibility, strategy formation, attention, working memory, response monitoring, and inhibition, it follows that the measures should adequately evaluate these functions. Unfortunately, most existing measures do not have the specificity to isolate a single ECF (Ogilvie et al., 2011). In most cases, a task that has been deemed an appropriate measure of executive functions is said to assess several ECFs, with certain variables more likely to tap into a specific ECF. To illustrate, the D-KEFS is a neuropsychological measure used to measure ECFs. One of the nine subtests is the Verbal Fluency Test (described in the Methods section) which overall is said to assess working memory, response monitoring, inhibition, and cognitive flexibility.
However, specific variables within this subtest are said to tap more specifically isolated ECFs; for example, Category Switching Total Correct Responses scaled score is thought to measure cognitive flexibility. Yet, there is not always a one to one relationship; instead, some variables represent a smaller combination of ECFs. Thus, measurement selection often involves choosing tasks that collectively, or in combination, measure the ECFs of interest.

Thus far the inclusion of subjective self-report, as well as objective performance measures has been endorsed. However, simply including both a process and outcome measure may not provide an adequate account of how the individual with ECF deficits is affected, nor how they may be rehabilitated (Chan et al., 2008; Lewis, Baggage, & Leatham, 2011). Another model which is not specific to ECFs but has been emphasized in the general cognitive rehabilitation literature (Chan et al., 2008; Lewis et al., 2011) takes in to consideration the various ways cognitive deficits can be manifested. The International Classification of Functioning, Disability and Health (WHO, 2000) provides a framework to describe an individual’s functioning, or level of functionality, from various perspectives. The model states that a disease affects an individual on four main levels. The first two levels describe pathology and resulting impairment of the disease within the individual. These levels define the diagnosis and indicate what is impaired (e.g., specified area of the brain). The remaining two levels describe the disability, as well as the handicap, the latter sometimes referred to as participation of the individual, which results from the disease or the external consequences of the disease. These levels speak to how the disease affects the functionality of the individual (e.g., effects on everyday social, vocational, or educational components).
Chan et al. (2008) provides a persuasive argument that ECF deficits should be conceptualized using the levels of functionality. The authors criticize traditional research examining ECFs of individuals, arguing that most extant research use measures that focus exclusively on performance at the pathological or impairment level (also see Whyte et al, 1996), and ignore the functional status of the individual. In support of this argument, individuals with lesions in their prefrontal cortex have been found to perform similarly to controls on many traditional neuropsychological measures (referred to in this model as measures of impairment), but exhibit difficulties within everyday activities (representing disability or participation problems in this model; Shallice & Burgess, 1991). As such, the authors argue that assessment of ECF deficits should span across the continuum of impact (i.e., including measures capturing impairment, disability, and participation deficits).

Several existing measures of ECFs, which span the various categories of measures, have been classified under these various levels of functionality. Under these definitions, measures that focus on specific skills with restrictions applied to the environment, sequencing, and instructions appear to tap the *impairment* level. Behavioural performance measures such as variations of the Stroop task, verbal fluency type tasks, tower tasks, and WCST, fall under this category. Whereas, behavioural performance measures that provide less structure and attempt to represent more cognitive processes engaged during everyday activities, such as the Six Elements Test (SET), appear to tap the *disability* experienced by the individuals. Finally, measures that utilize a naturalistic environment or ask about functioning in the naturalistic environment are thought to tap the level of an individuals’ *participation* that is affected. Behavioural performance measures, such as the Iowa Gambling Task (IGT), and self-report measures, such as the Dysexecutive
Questionnaire or Behavior Rating Inventory of Executive Function (BRIEF), fall under this category.

Taken together these models emphasize that no single measure comprehensively assesses all ECF domains or levels of impact, yet there is promise in capturing the various ECFs from both a performance and functional perspective with the use of a battery of assessment tools including process and outcome measures. Thus, for the present dissertation the ECFs of interest were captured using both self-report (i.e., BRIEF-Adult) and performance measures, which were also selected to represent the functioning levels of impairment, disability, and participation. More traditional neuropsychological performance measures were selected to represent impairment in functioning (i.e., subtests of the D-KEFS), a less structured performance measure was selected to represent the construct of disability (i.e., Modified SET), and a naturalistic performance measure was selected to represent the construct of participation (i.e., IGT).

**Current efforts to measure ECFs of offenders.** Research examining forensic rehabilitation programming aimed at improving aspects of ECFs rarely employs measures from more than one of the assessment categories named above (i.e., neuropsychological, self-report, and naturalistic measures). Forensic practitioners and researchers have commonly used self-report or informant report ECF measures (e.g., Pugh, 1993; Wilkinson, 2005; Young, Chick, & Gudjonsson, 2010). Reliance on a single category of measure may be misleading. Particularly problematic, self-reports rely on a subjective appraisal of performance. Problems with various ECFs can result in impaired self-awareness of deficits, whereby the individual is not able to realistically appraise his or her cognitive state (Wedding & Reeves, 1997). However, other categories of measures are not immune to problems. Reliance solely on traditional neuropsychological measures of
ECFs, which have the advantage of documented psychometric properties and normative data, may not always inform professionals how the individual would function in the real world.

Some researchers have utilized two of the categories of measures in an attempt to more accurately capture ECF deficits in different functional areas. Hodel and West (2003) used a combination of self-report and performance measure of ECFs to conclude correctional program effectiveness. Over the 5 week program period, significant improvements in attention and memory were observed and reported by participants. Mullin and Simpson (2007) used an extensive neuropsychological measure of ECFs, as well as an informant report to predict correctional program outcome. Measures of ECFs predicted a large proportion of the variance for the outcome variables. In a well-designed study by Serin, Gobeil, and Preston (2010), ECFs of violent offenders participating in correctional rehabilitation programming were assessed using self-report and vignettes. Vignettes added the element of assessment of behaviour in a naturalistic environment; however, they were still susceptible to the biases of self-report measures. Measuring ECFs provided a richer base of information within these later studies, which could be used to improve programming content, approach, and effectiveness, and ultimately improve offenders’ outcomes. These improvements could not be made from recidivism data alone. These studies are featured to highlight the value, yet infrequency, of exploring offenders’ ECFs in relation to correctional programming. They also call attention to the importance of accurately and appropriately measuring ECFs among offenders, an issue addressed in the conceptualization of the present dissertation, detailed below.

**The Current Study**
As reviewed, there is substantial evidence depicting deficits in ECFs among populations of offenders, as well as illustrating the link between ECF deficits and criminal behaviour. Among individuals with schizophrenia, who share similar deficits, evidence of rehabilitation efforts have surfaced that demonstrate improvement in specific targeted ECFs, as well as psychosocial functioning. The argument was made that amelioration of offenders’ ECFs may be important to decreasing undesirable behaviours such as re-offending. Yet, there is a dearth of empirical research to address whether ECFs are being affected by correctional programming. Barriers toward this goal were discussed, such as difficulties in accurately and appropriately measuring ECFs. In addition, the overreliance of recidivism rates as the primary outcome measure poses further barriers to assessing a program’s ability to impact offender ECF deficits. Additional methodological problems within the forensic literature, such as inadequate or non-existent control groups and high dropout rates, also muddy researchers’ ability to make conclusions regarding functioning of offenders. The current study sought to examine offenders’ ECFs, specifically cognitive flexibility, strategy formation, response monitoring, and inhibition, paying close attention to the issues outlined above.

There is little existing literature to support, or deny, the notion that ECFs are amenable to change among offenders, particularly as a result of correctional programming. Thus, the primary goal of the present dissertation was to conduct a methodologically sound study to investigate whether ECFs were impacted by correctional programming among populations of offenders. This was accomplished by the use of several experimental groups (i.e., those who completed various correctional programs) and a waitlist control group (i.e., those who met criteria but had not yet begun programming), and non-waitlist control group. Special attention was paid to ensuring that
the control group was an appropriate comparison. Because the use of convenient, and arguably dissimilar, control groups is common within this literature, the current dissertation also collected data from a non-waitlist control group so that comparisons could be made between the waitlist and non-waitlist control participants. In addition, particular focus was devoted to the operationalization of ECFs, discussed in detail in the methods section.

The specific scientific questions and associated hypotheses regarding ECFs were as follows:

1a) Does an inmate’s performance on measures of ECFs change over the length of a correctional program (i.e., two to seven months), beyond what is likely as a result of learning effects (i.e., compared to the changes seen in the control group)? Literature highlighting the efficacy of programs targeting ECF impairments in schizophrenia populations suggested that ECFs can be improved; thus, subtle, yet significant improvements were hypothesized. Because little to no previous research has examined which ECFs are targeted by correctional rehabilitation programming or whether changes in ECFs after correctional rehabilitation occur, only the general hypothesis of improvement across all ECFs measured was made.

1b) If changes in ECFs are apparent after programming, do correctional programs differentially impact ECFs (e.g., NSAP-M versus VPP-M)? Preliminary evidence suggests that these program types differentially impact recidivism, thus the possibility exists that changes in ECFs may also be differentially impacted. Following their respective reductions in recidivism, it is hypothesized that VPP-M will have the greatest positive impact on ECF deficits, followed by NSAP-M.
2) As a secondary exploration, the current dissertation sought to investigate other factors that more typically represented variables explicitly targeted and measured pre and post correctional programming; specifically, the factors of motivation for change and general criminal attitudinal change, as discussed by Losel (2001) and above. Thus, special attention was paid to whether motivation for change, or criminal attitudes, was impacted over the course of correctional programming, compared with changes seen among the control group. Support for special examination of these factors came from the following. The extant literature suggests that greater program benefit is derived when motivation for change is present; however, little is known about whether motivation levels relate to ECFs, particularly related to changes in ECFs as a result of intervention. There is some evidence to suggest that certain ECF deficits are associated with increased correctional treatment dropout (Fishbein et al., 2009) suggesting a possible link between ECFs and motivation. Similarly, attitude change is cited as a primary focus of many of the programs; however, little is known about whether criminal attitudes relate to ECFs, particularly related to changes in ECFs as a result of intervention. In at least one study, the construct of self-control, an attitudinal construct similar to impulsivity, has been associated with criminal attitudes (Wormith, 1984). As with motivation for change, this investigation is also exploratory in nature.

Each of the above research questions has important implications for offender rehabilitation, which not only affects the individual offender, but also those who are charged with the goal of rehabilitation, as well as the community that must reintegrate the offender. Uncovering the ECFs that are not currently being enhanced by correctional programming will assist both practitioners and researchers in their attempts to advance programming. In addition, identifying specific ECFs that are being positively impacted
by programming will give further evidence for the continuation of specific efficacious programs and/or program elements, and may assist researchers to understand the change mechanisms that relate to recidivism reduction.

**Method**

**Participants**

After receiving ethical, scientific, and administrative approval from the internal review board of The University of Western Ontario (see Appendix A) and CSC, 208 incarcerated males from two federal penitentiaries were invited to participate in the current study. Of those invited, 58 did not begin the study due to disinterest ($n = 53$), inadequate English proficiency ($n = 3$; all were of Inuit descent), or pending appeal for their offence ($n = 2$). Thus, 150 inmates completed the pre assessment of the study. Sixteen resided at Bath Institute, while the remainder resided at Fenbrook Institution ($n = 134$). Of those who completed the pre assessment, 132 were recruited from rehabilitation program waitlists. Twenty-seven participants who began the study failed to complete the post assessment measures due to transfer to another institution ($n = 13$), release to the community ($n = 3$), or refusal ($n = 11$). Demographic data of those who did not complete the study (referred to as ‘Dropouts’) is presented below, followed by demographic data of participants.

Dropouts ranged in age from 21 to 69 years ($M = 36.44$, $SD = 12.64$), with the majority being single (56%), and identified as Caucasian (70%). The average years of education among Dropouts was 11.46 ($SD = 2.00$) ranging between 9 and 17.5 years. An estimate of intelligence revealed a mean IQ of 86.41 ($SD = 11.72$), ranging from 57 to 101, mirroring mean IQ of other incarcerated samples (e.g., Fishbein et al., 2009).

---

1 A third penitentiary (Beaver Creek Institution) was also utilized, but only to complete post assessment measures in the rare cases that the participant had been transferred to this proximate facility.
Dropouts were currently serving a mean sentence length of 2929.07 days ($SD = 3342.18$), or approximately eight years (for those who had received a ‘life’ or indeterminate length sentence, 25 years was used, since it represented the most common length before parole was considered). Dropouts had served an average of 2436.44 days ($SD = 2880.89$) of their current sentence, with a range of 182 to 9813 days. In addition, eight (29%) Dropouts had previously been federally incarcerated. The majority of Dropouts had a calculated risk score of 4 (52%), representing high static and high dynamic risks for re-offending. The concepts of this risk assessment are briefly described in a later section.

Of the 150 initial participants, 123 inmates completed both pre and post study measures. Participants ranged in age from 19 to 58 years ($M = 35.63$, $SD = 9.78$), with the majority being single (58%), and identified as Caucasian (65%; see Table 1 and Table 2 for detailed group demographics). The average years of education among the sample was 11.47 ($SD = 2.53$) ranging between 6 and 20 years. An estimate of intelligence revealed a mean IQ of 90.59 ($SD = 12.19$), ranging from 54 to 126. Participants were currently serving a mean sentence length of 3347.07 days ($SD = 3327.70$), or approximately nine years. Participating inmates had served an average of 1640.28 days ($SD = 1965.71$) of their current sentence, with a range of 74 to 9926 days. In addition, 35 (28%) participants had previously been federally incarcerated. Forty-six percent of participants had a calculated risk score of 4 (representing high static and dynamic risk). During the 12 months of participant recruitment, 17 targeted rehabilitation groups were run across the two prisons, eight NSAP-M (five at Fenbrook; three at Bath), seven
Table 1

Mean (and Standard Error) of Demographic Information

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>NSAP-M n = 37</th>
<th>VPP-M n = 23</th>
<th>MIFVPP n = 9</th>
<th>Controls n = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>34.03 (1.55)</td>
<td>31.44 (1.96)</td>
<td>42.56 (3.14)</td>
<td>37.56 (1.33)</td>
</tr>
<tr>
<td>Education years</td>
<td>10.93 (.42)</td>
<td>11.52 (.53)</td>
<td>12.00 (.85)</td>
<td>11.68 (.36)</td>
</tr>
<tr>
<td>IQ</td>
<td>88.03 (1.92)</td>
<td>92.13 (2.44)</td>
<td>89.22 (4.11)</td>
<td>91.50 (1.65)</td>
</tr>
<tr>
<td>Risk level score</td>
<td>3.14 (.15)</td>
<td>3.30 (.18)</td>
<td>3.56 (.29)</td>
<td>3.19 (.13)</td>
</tr>
<tr>
<td>Index sentence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Days served to date</td>
<td>484.51 (267.77)</td>
<td>781.39 (339.63)</td>
<td>606.00 (542.96)</td>
<td>1842.64 (230.35)</td>
</tr>
<tr>
<td>Total days served</td>
<td>856.65 (309.88)</td>
<td>1333.17 (393.03)</td>
<td>1345.11 (628.30)</td>
<td>2407.38 (266.57)</td>
</tr>
<tr>
<td>Number of programs Completed</td>
<td>.60 (.22)</td>
<td>.57 (.28)</td>
<td>.78 (.45)</td>
<td>1.24 (.19)</td>
</tr>
<tr>
<td>Days between Pre and Post</td>
<td>93.11 (7.89)</td>
<td>115.35 (10.00)</td>
<td>162.75 (16.96)</td>
<td>114.28 (6.78)</td>
</tr>
</tbody>
</table>

Note: NSAP-M = national substance abuse program - moderate intensity; VPP-M = violence prevention program – moderate intensity; MIFVPP = moderate intensity family violence prevention program.
Table 2

*Number of Participants of each Ethnicity and Marital Status*

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>NSAP-M n = 37</th>
<th>VPP-M n = 23</th>
<th>MI FVPP n = 9</th>
<th>Controls n = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>24</td>
<td>17</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>African Canadian</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Inuit</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>22</td>
<td>13</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>Married/</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Cohabitating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated/Divorced</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Widowed</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
VPP-M (five at Fenbrook; two at Bath), and only four MIFVPP (three at Fenbrook; one at Bath). NSAP-M groups typically consisted of 10 to 15 participants, while VPP-M and MIFVPP consisted of 5 to 10 participants. Pre assessment data was obtained from 30 inmates who began the violence prevention program (VPP-M group); however three inmates began the program but did not complete and four were unavailable for follow-up (i.e., due to transfer, release, or refusal). Pre assessment data was obtained from 10 inmates who began the family violence program (MIFVPP group), however one inmate began the program but did not complete. Pre assessment data was obtained from 40 inmates who began the substance abuse program (NSAP-M group); however, three inmates were unavailable for follow up. As alluded to, there were four inmates who did not complete their assigned correctional program. These participants were excluded from analysis. In cases of larger samples, comparison of completers and non-completers have been recommended (e.g., McLean & Grace, 1998); however the small sample size of non-completers did not support this.

Pre and post assessment measures were collected from 32 participants who met the admission criteria for an above mentioned program but were not assigned to begin the program during the study and thus were assigned to the Waitlist Control group. Eighteen participants represented a convenient sample of inmates who were not currently enrolled or awaiting rehabilitation programming to be used as a Non-Waitlist Control comparison. Therefore, the final groups considered in the various analyses consisted of NSAP-M (n = 37), VPP-M (n = 23), MIFVPP (n = 9), Waitlist Control (n = 32), and Non-Waitlist Control (n = 18). All data was obtained during a 15-month period beginning in March 2010.
Prior to data collection several sources were consulted to determine the sample size needed to conduct repeated measure analyses with appropriate effect size and power. When recidivism rates have been used as the primary outcome in forensic rehabilitation research, effect sizes have typically been small. How these effects sizes translate to the virtually non-existent repeated measure designs of ECFs and rehabilitation programming studies was difficult to predict. In non-forensic populations, such as Schizophrenia, rehabilitation programs report medium effect sizes (e.g., McGurk et al., 2009; Penades et al., 2006). Thus, the current author felt that it was reasonable to expect between a small to moderate effect size (e.g., $0.3 < \text{d} < 0.7$). Therefore, during the planning stages, an effect size of 0.5 was used. Germane to the present dissertation, Tran (1997) discussed two types of repeated measures hypotheses (e.g., treatment main effects, and consistency of treatment differences) and the relevant sample sizes required. Because a variety of analyses were proposed, the analysis demanding the greatest number of participants per cell was reported here. Using the anticipated effect size of 0.5 and desired power of 0.8, Tran suggested that each treatment group would require between 31 to 54 participants. Consultation with a statistician (Bradley Corbett, Ph.D., at The University of Western Ontario) suggested that previous research supported the lower end of that suggestion (e.g., 31 participants per cell; Guilford & Frunchter, 1978). A priori examination using the program G*Power (3.1.2) confirmed that these parameters would achieve the appropriate power, and in fact, suggested lower samples sizes in some cases (e.g., 9, 24, and 59) depending on whether between, within, or interactions effects were being examined. Every effort was made to obtain at least the minimum of 31 participants per group; however, as discussed above, this was not always possible due to the infrequency of some rehabilitation programs, attrition of participants, length of time required to
participate in the study (e.g., three plus hours), and restrictions placed on inmate movement.

**Measures**

**Measures of ECFs.** In line with the current conceptualization of ECFs, several measures were selected to capture various functions. Although most individual measures cannot claim to isolate a single ECF, due to task completion typically requiring additional cognitive processes (Ogilvie et al., 2011), a problem discussed in more detail earlier, variables within measures have been empirically supported to measure specific ECFs, or combinations of ECFs. The measures and subsequent variables are described below.

ECFs were assessed using traditional neuropsychological measures, a self-report measure, and a naturalistic measure, administered in random order. Following the criteria outlined by the WHO (2000) and discussed by Chan et al. (2008), the measures represented the functional levels of impairment, disability, and participation, and are presented under those categories.

**Impairment measures.** The Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan & Kramer, 2001a) is a neuropsychological measure used to identify specific ECF deficits. The measure consists of nine stand-alone tests, many of which are improvements on preexisting tests of ECFs. Based on literature citing use of select D-KEFS subtests, or measures similar in nature, three subtests were chosen for the current study: Verbal Fluency Test (VFT), Color-Word Interference Test (CWIT), and Tower Test (TT). Normative data was compiled using over 1,000 individuals ranging from 16 to 89 years of age, representing a wide range of racial, ethnic and educational backgrounds, matched to the 2000 United States Census (Delis, Kaplan, & Kramer, 2001b). Internal consistency reliabilities for adults ranging between 20 to 79 years of age were generally
high for the VFT (from 0.43 to 0.90 across various ages of adults), the CWIT (0.72 to 0.86), and the TT (0.56 to 0.78; Delis et al., 2001b). The test-retest reliability estimates, obtained based on an average administration interval of 25 days, for adults between 20 and 89 years of age, across the various conditions of each task, ranged between 0.24 to 0.88 for the VFT, 0.49 to 0.86 for the CWIT, and 0.38 to 0.41 for the TT. The D-KEFS has been used with offender populations as a predictor of correctional program outcome (Mullin & Simpson, 2007), and to identify neuropsychological deficits among repeat driving while impaired offenders (Ouimet et al., 2007).

a) VFT: Based on the Controlled Oral Word Association Test (Benton & Hamsher, 1976; Spreen & Benton, 1969), the VFT involved three conditions: Letter Fluency, Category Fluency, and Category Switching. Participants were asked to generate words beginning with specific letters, words belonging to specific semantic categories, and words that switched between two semantic categories, respectively. Each trial was allotted 60 seconds. The VFT measured the participants’ ability to generate words fluently in an effortful, phonemic format (Letter Fluency), from over-learned concepts (Category Fluency), while simultaneously shifting between over-learned concepts (Category Switching; Delis et al., 2001a). Overall performance on this task assesses working memory, and concept generation (Royall et al., 2002) which included the ability to maintain a cognitive set and response monitoring, inhibit irrelevant associations, and uphold cognitive flexibility (Salthouse, 2005). Variables of interest included the Category Switching Total Correct Responses scaled score (henceforth referred to as ‘VF Switching’), which measures cognitive flexibility or set shifting, and the Category Switching Versus Category Fluency Contrast scaled score (‘VF Switching Contrast’) which further isolates cognitive flexibility deficits by comparing a higher order task with
a more basic task. Post treatment procedures utilized an alternative form of the VFT as suggested by Delis et al. (2001a).

b) CWIT: This test represented a restructured version of the Stroop (Golden, 1978; Trencerry, Crosson, DeBoe & Leber, 1989) in which participants were shown a list of different colours that were typed in incongruent ink colours. Participants were required to say the colour of the ink the word was printed in, opposed to reading the printed word. A colour naming and colour reading baseline measure was also included, as well as a switching condition where the participant switched between saying the word and saying the ink colour. Overall performance on this task assessed working memory, inhibition and cognitive flexibility (Delis et al., 2001a; Royall et al., 2002; Salthouse, 2005).

Variables of interest included the Inhibition scaled score (‘CW Inhibition’) and the Inhibition/Switching scaled score (‘CW Inhibition/Switching’), which are both considered the higher order cognitive tasks measuring inhibition and cognitive flexibility; as well as, the Inhibition/Switching Versus Inhibition Contrast scaled score (‘CW Inhibition/Switching Contrast’), which allows for further identification of the predominant impairment.

c) TT: Participants constructed a tower of a particular design by moving disks of different sizes across three vertical pegs in the fewest number of moves. Similar to the Tower of Hanoi and Tower of London tasks (Humes, Welsh, Retzlaff & Cookson, 1997), performance on the TT indicated participants’ working memory, planning, inhibition, and goal management skills which included establishing and maintaining a cognitive set and rule learning abilities (Royall et al., 2002; Salthouse, 2005). Total Achievement scaled score (‘TT Achievement’) was the variable of interest which measures various aspects of
ECFs including, spatial planning, rule learning, inhibition, as well as establishing and maintaining a cognitive set.

**Participation measures.** The Behavior Rating Inventory of Executive Function – Adult Version (BRIEF-A; Roth, Isquith, & Gioia, 2005) is a self-report measure containing 75 items within nine independent theoretically and empirically derived scales that measure various ECFs, including inhibition, shifting, emotional control, self-monitoring, initiation, working memory, planning/organizing, task monitoring, and organization. For the current study, the Global Executive Composite $t$ score (‘BRIEF Composite’), which is a composite of all nice subscales, represents the variable of interest. Commonly, BRIEF-A $t$ scores at or above 65 are considered an indication of substantial ECF deficits (Roth et al., 2005). The BRIEF-A has a normative sample of over 1,000 American adults from a wide range of racial, ethnic and educational backgrounds, matched to the 2002 United States Census (Roth et al., 2005). Although males and females, ranging in age between 18 and 90 years of age, were included in the sample, gender accounted for less than 2% of the variance in any scale. Separate norms based on age are provided. The BRIEF-A has been shown to be sensitive to subtle executive changes (Rabin et al., 2006). In a mixed clinical sample, the alpha coefficients for internal consistency ranged from 0.80 to 0.94 for the various subscales, while test-retest reliability after a mean of four weeks ranged from 0.82 to 0.93 (Roth et al., 2005). The BRIEF-A showed convergence (i.e., correlations ranging from 0.38 to 0.80) with similar measures, such as the Dysexecutive Questionnaire (Wilson, Alderman, Burgess, Emslie, & Evans, 1996).

The Iowa Gambling Task (IGT; Bechara, 2007), the second measure in the participation category, is a computerized behavioural task aimed at measuring risky
decision making. Participants are instructed to select cards from the four ‘decks’, switching as often as desired, in order to maximize total winnings, expressed in dollars at the top of the screen. Each card selected offered the participant a reward or combination of reward and punishment (e.g., “Win 100”, or “Win 100, Lose 200”). Although money could be lost from selections from each deck, two of the decks were associated with overall financial gain while the two remaining decks were associated with overall financial loss. The pattern of financial contingencies was complex and participants were not typically explicitly aware of the pattern. Nonetheless, non-impaired participants tend to show a progressive increase in selections made from the advantageous decks. However, individuals with frontal lobe impairments (e.g., medial orbitofrontal or ventromedial prefrontal cortex lesions) did not demonstrate this improvement, and instead consistently lost money (Bechara, Damasio & Damasio, 2000), as well as showed deficits in forming strategies that promoted delayed gratification, such as drawing cards from low risk decks (Bechara, Damasio, Damasio, & Anderson, 1994; Bechara et al., 1998). To date, no known studies have directly examined the reliability of the IGT; however, evidence suggests that the IGT simulates real-life reward-dependent decision making, and represents what some authors refer to as the ‘hot’ components of ECFs (Buelow & Suhr, 2009; Seguin & Zelazo, 2005), as emotional processing is associated with performance on the task. Whereas the ‘cold’ components of ECFs are more likely to be elicited by abstract or decontextualized tasks, such as those in the D-KEFS, the ‘hot’ components are elicited by tasks that involve the regulation of affect and motivation, such as gambling tasks involving losses and rewards. The normative standards used represent USA Census-matched scores which are recommended when “inferences regarding the adequacy of an examinee’s capacity for everyday functioning” is desired (Bechara, 2007,
This normative sample consisted of 264 non-impaired adults matched to the 2002 United States Census on age and education. Although males and females, ranging in age between 18 and 90 years of age, were included in the sample, gender accounted for less than 1% of the variance. Completion of the IGT consists of selecting 100 cards, with each selection of 20 cards referred to as a ‘Block’ (e.g., Block 1 represents cards 1 to 20 selected, whereas Block 4 represents cards 61 to 80). Recent research suggests that the first 40 cards selected (or Blocks 1 and 2) represent decisions made under ambiguity because the outcome is unclear; whereas later selections represent decisions made under risk because outcomes depend on known probabilities (Brand, Recknor, Grabenhorst, & Bechara, 2007). Decisions under risk, and not decisions under ambiguity, are said to be influenced by ECFs, particularly cognitive flexibility and response monitoring (Brand, Labudda, & Markowitsch, 2006). Therefore, the \( t \) scores derived from Block 4 or the 61\textsuperscript{st} to 80\textsuperscript{th} card selection (‘IGT Net 4’), and Block 5 or the 81\textsuperscript{st} to 100\textsuperscript{th} card selection (‘IGT Net 5’) represented the variables of interest. A \( t \) score less than or equal to 39 represented impaired functioning, while a \( t \) score ranging from 40 to 44 represented below average functioning, and a \( t \) score greater than or equal to 45 represented non-impaired functioning. Relevant to the present study, poor IGT performance has been demonstrated in populations of aggressive and impulsive individuals (Best, Williams, & Cocaro, 2002).

**Disability measure.** The Modified Six Elements Test (MSET; Wilson et al., 1996) is a simplified version of the original SET (Shallice & Burgess, 1991). It included a total of six tasks, comprising of three activity types: dictation, picture naming, and arithmetic. Unlike some traditional neuropsychological measures, few restrictions are applied to the completion of the test, with the exception of time (e.g., 10 minute limit). For optimal achievement, the activities must be started in a specified time period, and no two tasks of
the same type may be done in sequence. Successful completion of this task required the ability to think ahead, monitor previous and future responses, adhere to task rules, as well as monitor the passage of time, which measured the ECF abilities of planning, strategy formation, and self-monitoring (Burgess, 1997). Overall profile scaled score (‘MSET Profile’) represented the variable of interest. The MSET is a subtest of the Behavioural Assessment of the Dysexecutive Syndrome (BADS) which has normative data; however, since only one subtest was used in the current study comparison to the normative or patient sample was not possible. Good psychometric properties have been reported (Wilson et al., 1996). For instance, inter-rater reliability ranged from 0.88 to 1.00. In addition, the MSET successfully differentiated between the performance of individuals with a brain injury from non-injured controls ($t = 10.6, p < .0001$). Test-retest reliability after 6 to 12 months among normal controls was 0.33. The SET has been used to detect impaired cognitive functioning in frontal lobe lesion patients (Gouveia, Brucki, Malheiros & Bueno, 2007) and individuals with schizophrenia, (van Beilen, Withaar, van Zomeren, van den Bosch, & Bouma, 2006). Deficits in these populations often represent greater likelihood to break the rules of the task, that is, in the plan-following processes, rather than in planning the strategic approach to solve it. The MSET has been used with incarcerated adult males, who performed significantly worse than non-incarcerated controls (Barbosa & Monteiro, 2008).

**Additional Measures.** The Kaufman Brief Intelligence Test, Second Edition (KBIT-2; Kaufman & Kaufman, 1997) is a measure of intelligence that provides an overall estimate of IQ, as well as specific verbal and nonverbal intelligence scores. Unlike other measures of intelligence, the KBIT-2 can be administered in approximately 20 minutes. Verbal intelligence is assessed using two subtests (Verbal Knowledge and Riddles) which
measures verbal, school-related skills by assessing word knowledge, range of general information, verbal concept formation, and reasoning ability. Nonverbal intelligence is assessed using one subtest (Matrices) which measures the ability to solve new problems by assessing the ability to perceive relationships and complete visual analogies. Test items are free of cultural and gender bias. The KBIT-2 has a normative sample of over 600 American adults from a wide range of racial, ethnic and educational backgrounds, matched to the 2001 population demographics of the USA (Kaufman & Kaufman, 1997). Separate norms based on age are provided. Psychometric properties, as reported by Kaufman and Kaufman (1997), were satisfactory. The KBIT-2 showed generally high internal consistency for adults 19 to 90 years of age on the IQ composite score (0.95), as well as on the Verbal score (from 0.88 to 0.96 across different ages of adults) and on the Nonverbal Score (from 0.87 to 0.93). In addition, the KBIT-2 correlated in the moderate to high range with other well-established tests of cognitive ability and academic achievement such as the Wechsler Abbreviated Scale of Intelligence, Wechsler Adult Intelligence Scale: Third Edition, and Wide Range Achievement Test: Third Edition. Test-retest reliability estimates of verbal, nonverbal, and IQ composite scores, obtained based on an average administration interval of four weeks in adult groups ranging between 22 and 89 years of age, were high (i.e., ranging from 0.85 to 0.92). IQ was not hypothesized to change over the two to seven month pre/post study period; thus, the KBIT-2 was administered only during the pre assessment procedure, as a method of ensuring that performance on ECF measures was not due to differences in intellectual ability. The overall estimate of IQ (‘IQ’) was used in the current study.

The University of Rhode Island Change Assessment Scale (URICA; McConnaughy, Prochaska, & Velicer, 1983) is theoretically based on the Transtheoretical Therapy Model
(Prochaska & DiClemente, 1984), and was designed to assess participants’ readiness and motivation to change an undesired behaviour. The URICA consists of 32 self-report items with multiple items corresponding to each stage or subscale: Precontemplation (PC), Contemplation (C), Action (AC), and Maintenance (M). All items were measured using a 5-point likert format (1 = strongly disagree and 5 = strongly agree). In the present dissertation, the last two sentences of the instructions were modified to read “For all the statements that refer to your "problem", answer in terms of the issues that led to your involvement with the criminal justice system. In these questions, the word "here" refers to this institution and the programs offered here.” Subscale scores were calculated by summing relevant question scores, resulting in an assessment of the level of engagement of each stage. For the present study, a composite score (‘URICA Composite’) was of most interest ([C + AC + M] – PC) as supported by previous research as a practical application of the URICA (Amodei & Lamb, 2004; Blanchard, Morgenstern, Morgan, Labouvie & Bux, 2003; Field, Adinoff, Harris, Ball & Carroll, 2009; Pantalon, Nich, Frankforter, & Carroll, 2002). Reliability has been assessed in a two part study by Dozois, Westra, Collins, Fung, and Garry (2004) resulting in the following Cronbach’s alphas for each subscale: PC 0.77 and 0.73, C 0.80 and 0.79, AC 0.84 and 0.90, as well as, M 0.82 and 0.81, which are relatively consistent with the original reports (McConnaughey et al., 1983; McConnaughey, DiClemente, Prochaska, & Velicer, 1989). Moderate support has been found for the URICA’s convergent and discriminant validity (Amodei & Lamb, 2004; Dozois et al., 2004). The URICA has been used with groups of young offenders (Hemphill & Howell, 2000), incarcerated males (Serin & Kennedy, 1997), and a modified version was used with treatment court-ordered male batterers (Levesque, Gelles, & Velicer, 2000), to assess their readiness for change.
Motivation for change has been linked to treatment outcome (Baxter et al., 1995; Clarke, Simmonds, & Wydall, 2004); thus, the current author sought to explore its effects on changes in ECFs.

The Criminal Sentiments Scale-Modified (CSS-M; Shields & Simourd, 1991) is a 41 item, 3-point likert scale self-report measure designed to assess antisocial attitudes, values, and beliefs concerning criminality. Answers are coded so that higher scores reflect greater criminal attitudes, or less prosocial attitudes. The CSS-M consists of modified scales originally founded within the Connecticut Correctional System (Gendreau, Grant, Leipeiger, & Collins, 1979; Wormith & Andrews, 1984). The CSS-M consists of a total score and five subscales; however, frequently the first three subscales (Attitudes toward the Law, Attitudes toward the Court, Attitudes toward the Police) are combined to form the Law-Court-Police (LCP) subscale. The remaining two subscales consisted of questions related to Tolerance for Law Violations (TLV) and Identification with Criminal Others (ICO). The LCP subscale was thought to assess more general attitudes for law and the criminal justice system as a whole; while the TLV assessed specific justifications for illegal behaviour, and the ICO assessed attitudes about others who engage in illegal conduct. The total score (‘CSS-M Total’), which is the variable of interest for the present study, is a combination of the aforementioned scales. Normative scores for offenders and non-offenders are available for the original CSS (Andrews & Wormith, 1984); however, the CSS employed a 5-point likert scale and thus cannot be adequately compared. However, a clinical cut-off score of 19 has been established elsewhere (Morgan, Fisher, Duan, Mandracchia & Murray, 2010). The CSS reliability has been established with samples of adult offenders (Roy & Wormith, 1985; Witte, DiPlacido, Gu, & Wong, 2006; Wormith & Andrews, 1984). Internal consistency was
reported to range between 0.70 and 0.76 for the subscales and CSS-M total (Simourd, 1997). Change in criminal attitudes is a fundamental tenet of correctional programming, and has been linked to antisocial behaviour (Simourd & van de Ven, 1999) and self-control (Wormith, 1984). As there is an established link between ECFs, which includes constructs similar to self-control, and antisocial behaviour, the current author also sought to examine whether a connection between criminal attitudes and ECFs existed.

Participants were also asked to indicate their date of birth, marital status, ethnicity, and years of education using a demographic questionnaire. Information regarding current and past program enrolment and completion, as well as current and past sentence length was obtained through a review of the participant’s correctional file. In addition, information was obtained regarding each inmate’s risk level. Upon entry to a federal institution, each inmate is given a rating of risk (i.e., low, medium, or high) based on static (e.g., criminal history, recidivism prediction) and dynamic (e.g., employment, associates, substance use history) factors. For the current dissertation, static and dynamic risk levels were coded and combined to obtain one risk level score, ranging from 0 (designating a low/low static/dynamic level) to 4 (designating a high/high static/dynamic level). All file reviews were performed subsequent to testing to avoid bias.

Procedure

Participants were recruited from two medium security federal penitentiaries within Canada. Fenbrook Institution is located in Gravenhurst, Ontario and housed approximately 450 male inmates at the time of the study. Bath Institution is located in Bath, Ontario and housed approximately 350 male inmates. Participants were also accessible if they resided at Beaver Creek Institution, a minimum security federal penitentiary; however, only post data was obtained at Beaver Creek in the event that an
inmate was relocated there upon rehabilitation program completion. Incarceration in any Canadian federal facility required a sentence of over two years, and was not dependent on the type of crime committed.

Inmates whose names appeared on the waitlist for upcoming moderate intensity correctional rehabilitation programs, specifically VPP-M, MIFVPP, and NSAP-M, were sent letters of invitation through the internal mail system or were approached in person. After a brief in-person discussion regarding testing procedures and scheduling, interested participants were escorted by the researcher to a quiet testing room at the agreed upon testing session. Study inclusion criteria included ability to speak and understand English, and normal or corrected to normal vision. Before the participant consented, the voluntary nature of the study was made explicit, as was the fact that participation had no bearing on any subsequent correctional decisions. Previous computer experience was assessed. When required, the participant was given a short tutorial and several practice sessions on the basics of using a computer mouse. A voice automated computer version of the three self-report questionnaires (BRIEF-A, CSS-M, and URICA) was used to assist with literacy and attentional problems sometimes found in this population. All other measures were completed with the experimenter. Measures were randomly ordered. No monetary compensation was provided, as prescribed by CSC.

Initial testing sessions were approximately 2 hours and 15 minutes in length. The majority of participants completed the testing in one session; however, 17 participants required testing to be broken up into two sessions (i.e., one in the morning and one in the afternoon). On seven occasions testing sessions were broken up by more than 24 hours. Following the completion of the testing session, the offender’s institutional file was reviewed to ascertain number of previous federal incarcerations, sentence length, verify
institutional program enrolment information, and demographic material not obtained during the testing session.

Whenever possible, all inmates who appeared on the waitlist for a targeted program were invited to participate. Some inmates were unavailable during the recruitment periods due to onsite work detail, illness, or segregation status. Programs typically enrolled 5 to 15 participants; however, the waitlists contained two to four times this number of inmates. Thus, in most cases participants completed pre testing measures before certainty of program enrollment. In general, recruitment was started approximately one month before the scheduled start of a program; however, on most occasions program start dates were delayed. Because programs began somewhat unpredictably, the researchers waited up to five months to ascertain whether the participant had begun the program. If a participant began a program within five months of his initial assessment, he was assigned to the appropriate group for the study: VPP-M, MIFVPP, or NSAP-M. If after five months the participant had not begun the program which he was on the waitlist for, the participant was assigned to the Waitlist Control group.

For those who began programming, all measures, with the exception of the demographics and KBIT-2, were re-administered to participants following the completion of the correctional program. Waitlist and Non-Waitlist Controls also completed post measures. If the researchers were notified that a Control participant was receiving early release or being transferred to another institution, the post measures were completed at that time; otherwise, post measures were completed on a similar schedule to those who had completed programming. On three occasions the post measures were completed at a
nearby minimum security institution (Beaver Creek Institution). Post testing sessions were approximately 1 hour and 15 minutes in length.

The present dissertation employed a quasi-experimental design. Although randomized designs are held to be the gold standard, there is increasing recognition that randomization is not the only method of obtaining meaningful data (Munro, 2005; Slade & Priebe, 2001). Ethical concerns related to random assignment, particularly when the intended intervention is behaviour change related to public safety, preclude its use with certain populations or interventions. Quasi-experimental design has been shown to be an appropriate method for forensic interventions. When compared to treatment outcome studies with offenders using randomization, studies using non-randomized procedures showed no statistical difference in treatment outcome (Lipsey, Chapman & Landenber, 2001).

Results

The following analyses were conducted to determine whether ECFs improved after completion of rehabilitation programs, and whether changes were above and beyond what would be expected due to learning effects. More specific analyses were also used to test whether differences in ECFs after rehabilitation existed between inmates who completed specific rehabilitation programs.

Prior to analyses, the data was screened for univariate outliers by comparing the variable means to the 5% trimmed means (i.e., the means with the top and bottom 5% of cases removed; Field, 2009a). Trimmed means were similar to untrimmed means; thus, univariate outliers were not present. Multivariate outliers were examined using Mahalanobis distance, which is a chi-square statistic. None of the differences were significant (i.e., no p-values were < .001) indicating that there were no multivariate
outliers. The assumptions for conducting repeated-measures analyses of variance (ANOVAs) and multivariate analyses of variance (MANOVAs) were also examined for any violations (i.e., normality, homogeneity of variance, homogeneity of covariance matrices, multicollinearity; Field, 2009a). The validity of conclusions from MANOVA are based on the assumption of the absence of multicollinearity, and this assumption was not violated as indicated by the observation that correlations among ECF scores did not exceed 0.64 (see Table 3). Homogeneity of variance or of covariance matrices was also not violated, as indicated by nonsignificant Mauchly’s Test, Levene’s Test, or Box’s M, when appropriate. Normality is discussed in the following section. Self-report measures were also examined for measurement reliability using Cronbach’s alpha which was found to be acceptable for all three self-report measures (i.e., 0.97 for the BRIEF-A, 0.91 for the CSS-M, and 0.85 for the URICA). The assumption of normality, as well as the composition of the control group, is discussed below.

**Assumption of Normality Violation for ANOVA and MANOVA**

Skewness and kurtosis $z$ scores were calculated to check for normality. Violations of the assumption of normality were deemed to be present if the absolute $z$ score value was $> 3.2$ (Ainsworth, 2005). Pre and post assessment CW Inhibition were positively skewed. Pre and post assessment CW Inhibition/Switching were negatively skewed with the post score also being platykurtic. Pre and post assessment TT Achievement were both positively skewed and platykurtic. Pre and post assessment MSET Profile were negatively skewed. In addition, post assessment CW Inhibition/Switching Contrast was platykurtic. Appropriate transformations were attempted (e.g., finding the square root for negatively skewed variables, multiplying by exponents for positively skewed variables,
### Table 3

**Correlations between Pre and Post ECF variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>VF S</th>
<th>VF SFC</th>
<th>CW I</th>
<th>CW I/S</th>
<th>CW I/SC</th>
<th>TT A</th>
<th>MP</th>
<th>IN4</th>
<th>IN5</th>
<th>BC</th>
</tr>
</thead>
<tbody>
<tr>
<td>VF S</td>
<td>(.28*/.35*)</td>
<td>.49**/.58**</td>
<td>.30*/.36**</td>
<td>.21/.52**</td>
<td>-.07/.18</td>
<td>.04/- .01</td>
<td>.09/.22</td>
<td>.13/.12</td>
<td>-.03/-.15</td>
<td>-.31**/- .04</td>
</tr>
<tr>
<td>VF SFC</td>
<td>.42**/.59**</td>
<td>(.14/.10)</td>
<td>-.05/.11</td>
<td>-.01/.23</td>
<td>.05/.14</td>
<td>-.03/- .07</td>
<td>.10/.05</td>
<td>-.07/.16</td>
<td>-.25**/-.07</td>
<td>-.06/-.02</td>
</tr>
<tr>
<td>CW I</td>
<td>.30**/.38**</td>
<td>-.09/.20</td>
<td>(.80**/.83**)</td>
<td>.72**/.51**</td>
<td>-.23/- .50**</td>
<td>.19/.09</td>
<td>.18/.03</td>
<td>.20/.16</td>
<td>.04/22</td>
<td>-.19/.01</td>
</tr>
<tr>
<td>CW I/S</td>
<td>.33**/.23</td>
<td>.06/.10</td>
<td>.64**/.48**</td>
<td>(.88**/.81**)</td>
<td>.50**/.49**</td>
<td>.17/.15</td>
<td>.20/.25</td>
<td>-.02/.22</td>
<td>.13/- .06</td>
<td>-.07/14</td>
</tr>
<tr>
<td>CW I/SC</td>
<td>-.09/- .22</td>
<td>.13/-.18</td>
<td>-.28/- .58**</td>
<td>(.63**/.62**)</td>
<td>.00/.07</td>
<td>.07/23</td>
<td>-.22/07</td>
<td>.13/- .29*</td>
<td>.15/.13</td>
<td></td>
</tr>
<tr>
<td>TT A</td>
<td>.24*/-.15</td>
<td>.01/- .05</td>
<td>.11/-.15</td>
<td>.16/- .01</td>
<td>.08/12</td>
<td>(.25*/.49**)</td>
<td>.30*/12</td>
<td>-.08/11</td>
<td>.15/26</td>
<td>-.24*/-.06</td>
</tr>
<tr>
<td>MP</td>
<td>-.10/.26</td>
<td>-.19/11</td>
<td>-.03/06</td>
<td>.01/01</td>
<td>.03/-.07</td>
<td>.04/14</td>
<td>(.41**/.44**)</td>
<td>-.05/-.31*</td>
<td>.25*/-.11</td>
<td>-.27*/-.01</td>
</tr>
<tr>
<td>IN4</td>
<td>-.01/-.31</td>
<td>.01/42**</td>
<td>.12/20</td>
<td>.11/29*</td>
<td>-.02/02</td>
<td>.07/.06</td>
<td>.07/.33*</td>
<td>(.30*/.49**)</td>
<td>.39**/.43**</td>
<td>-.13/03</td>
</tr>
<tr>
<td>IN5</td>
<td>.07/11</td>
<td>.06/18</td>
<td>.32**/29**</td>
<td>.29**/32*</td>
<td>-.01/00</td>
<td>.10/.10</td>
<td>.01/13</td>
<td>.64**/.71**</td>
<td>(.33*/.10)</td>
<td>-.24*/-.12</td>
</tr>
<tr>
<td>BC</td>
<td>-.05/04</td>
<td>.16/- .08</td>
<td>-.20/.13</td>
<td>-.06/28*</td>
<td>.16/12</td>
<td>-.03/13</td>
<td>.10/08</td>
<td>-.01/04</td>
<td>-.22/13</td>
<td>(.70**/77**)</td>
</tr>
</tbody>
</table>

*Note.* VF S = VF Switching; VF SFC = VF Switching Contrast; CW I = CW Inhibition; CW I/S = CW Inhibition/Switching; CW I/SC = CW Inhibition/Switching Contrast; TT A = TT Achievement; MP = MSET Profile; IN4 = IGT Net 4; IN5 = IGT Net 5; BC = BRIEF Composite. Numbers appearing on the left of the slash are correlations among Programmers, while numbers on the right of the slash are correlations among Controls. Data below the diagonal of the table represent the correlations among Pre scores, and data above the diagonal of the table represent the correlations among Post scores. The data in the diagonal are the correlation between the pre and post scores of the same variable (e.g., in the top left cell, .28 is the correlation between VF Switching Pre and VF Switching Post among Programmers, and .35 is the correlation among Controls). **p<.01 *p<.05
and finding the reciprocal for platykurtic variables), but these transformations were unable to normalize most variables. Although normally distributed data are preferred, MANOVA is robust in the face of most violations of the assumption of normality, as long as sample sizes are not overly small (i.e., consisting of fewer than 20 participants; Lindman, 1974).

**Group Membership: Waitlist Controls and Non-Waitlist Controls**

Waitlist Controls and Non-Waitlist Controls were combined to form one control group. The non-waitlist control group consisted of 18 inmates who were not on a program waitlist but who instead volunteered to participate in the study; therefore this subsample constituted a convenience sample. Although the non-waitlist control group was originally intended to be kept analyses separate from the waitlist control group, revealed no significant group differences between Waitlist Controls and Non-Waitlist Controls in marital status or ethnicity. As indicated by a series of *t*-tests, there were no significant group differences found on age, years of education, IQ, days between pre and post assessment, risk level score, number of days served of current sentence, total number of days incarcerated including prior sentences, URICA Composite pre score, or CSS-M Total pre score. There was a significant difference between the number of past correctional programs completed prior to the study, *t*(51) = -2.72, *p* = .01, with Waitlist Controls having completed a mean of .89 programs (*SD* = 1.34) and the Non-Waitlist Controls having completed 2.17 programs (*SD* = 2.01). Although this difference could conceivably affect initial assessment of ECFs (i.e., additional programming could result in better ECF performance due to prior learning or amelioration of deficits), a strength of the pre/post design is its ability to control for participants’ initial performance. In an effort to maintain the largest appropriate sample, avoid violations of assumptions, and
maintain power, Waitlist Controls and Non-Waitlist Controls were combined to make up the control group (henceforth referred to as ‘Controls’).

**Demographic Information**

Participants who completed the study \( (n = 123) \) and those who did not complete the post measures (‘Dropouts’; \( n = 27 \)) were compared to determine whether group differences existed. Chi square analyses revealed no significant group differences in marital status or ethnicity. As indicated by a series of \( t \)-tests, there were no significant group differences in age, years of education, IQ, risk level score, URICA Composite pre score, or CSS-M Total pre score, number of past correctional programs completed, number of days served of current sentence, or total number of days incarcerated. Thus, no differences were found between study participants and Dropouts.

Programmers (i.e., participants who engaged in any relevant rehabilitation program during the current study; \( n = 69 \)) and Controls (\( n = 50 \)) were compared to determine whether there were group demographic differences. Chi square analyses revealed no significant group differences in marital status or ethnicity. As indicated by a series of \( t \)-tests, there were no significant group differences in age, years of education, IQ, days between pre and post assessment, risk level score, URICA Composite pre score, or CSS-M Total pre score. However, significant differences were found on the total number of rehabilitation programs previously completed, \( t(117) = 2.55, p = .01 \). Programmers had completed a mean of .61 \( (SD = 1.06) \) prior programs; whereas, Controls had completed a mean of 1.24 \( (SD = 1.64) \) programs. This difference was likely a result of the fact that the control group included Non-Waitlist Controls who, as determined above, participated in significantly more prior programming than Waitlist Controls. In addition, relative to Programmers, Controls had served significantly more days of their current sentence,
Thus, Programmers and Controls appeared to represent a relatively homogeneous group with the exception that Controls had, prior to study inclusion, completed twice as many rehabilitation programs, which is a logical consequence of their more than double total incarceration time.

**Hypothesis 1A: ECF Performance of Programmers and Controls**

The principal goal of the current dissertation was to longitudinally examine performance of ECFs of different groups of inmates, specifically those who participated in correctional rehabilitation programs compared to those who had not. The a priori plan of analysis was to use a single repeated-measures MANOVA with all 10 variables relating to ECFs acting as the dependent variables. However, examination of the correlations among those variables did not support this procedure (i.e., several variables were not adequately correlated with each other to be included in a single MANOVA using a customary correlation criteria of .4 or greater; see Table 3). Small correlations between naturalistic measures of ECFs and more traditional neuropsychological measures of ECFs, as well as among measures within each of these categories, have been previously established (e.g., Miyake et al., 2000; Wood & Liossi, 2007). The observed pattern of correlations supported including VF Switching, VF Switching Contrast, CW Inhibition, CW Inhibition/switching, and CW Inhibition/switching Contrast in a single MANOVA, IGT Net 4 and IGT Net 5 in another MANOVA, and TT Achievement, MSET Profile, and BRIEF Composite in separate ANOVAs. Thus, in order to examine whether inmates’ ECFs changed over the length of a correctional rehabilitation program,

---

2 Due to the large number of dependent variables being examined, data reduction procedures were also explored and are discussed in detail in the Endnote section following the References.
two repeated-measures MANOVAs and three repeated-measures ANOVAs were conducted with Program Status (i.e., Programmers vs. Controls) as the between-subjects variable, Time (pre vs. post) as the within-subjects variable, and ECF performance as the dependent variable(s). The interaction effect between Program Status and Time, as well as the main effect of Time, were relevant to the empirical questions at hand; thus, these were reported for all subsequent analyses. The main effect of Program Status is not reported here, or in subsequent analyses, because the interpretation does not help to answer the current empirical questions. Pillai’s Trace was reported for all analyses, as it is commonly used when the assumption of homogeneity of covariance matrices is not violated, as indicated by a nonsignificant Levene’s test and Box’s M.

Across the five repeated-measures analyses, two interaction effects were found (see Table 4). The multivariate interaction between Program Status and Time was significant for the IGT dependent variables (including the variables IGT Net 4 and IGT Net 5), indicating that completion of a rehabilitation program affected inmates’ IGT scores across time. Univariate analyses revealed that IGT Net 5 pre and post scores did not differ by group, $F(1,117) = .26, p = .61$, partial $\eta^2 = .002$. However, IGT Net 4 pre and post scores were found to differ by group, $F(1,117) = 5.41, p = .02$, partial $\eta^2 = .04$, with Controls showing an unexpected improvement between pre assessment ($M = 48.08, SE = 1.51$) and post assessment ($M = 52.58, SE = 1.55$), representing a significant change between pre and post, $t(49) = -3.40, p = .001$. Programmers did not differ between pre assessment of IGT Net 4 ($M = 49.17, SE = 1.29$) and post assessment ($M = 48.54, SE = 1.32$; $t(68) = .40, p = .69$; see Figure 1).

The multivariate interaction between Program Status and Time was also significant for the BRIEF Composite, indicating that the completion of a rehabilitation program
Table 4

*Results of the Five Repeated-Measures (RM) Analyses Comparing Pre and Post Scores by Group (Programmers vs. Controls)*

<table>
<thead>
<tr>
<th>Type of RM Analysis (&amp; Dependent Variables Included)</th>
<th>Interaction</th>
<th>Main (Pre/Post)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANOVA (VF Switching, VF Switching Contrast, CW Inhibition, CW Inhibition/Switching, CW Inhibition/Switching Contrast)</td>
<td>$F(5, 113) = .70, \ p = .63, \ \eta^2_p = .03, \ f = .18, \ P = .48$</td>
<td>$F(5, 113) = 19.82, \ p = .00, \ \eta^2_p = .47, \ f = .94, \ P = .99$</td>
</tr>
<tr>
<td>MANOVA (IGT Net 4, IGT Net 5)</td>
<td>$F(2, 116) = 4.00, \ p = .02, \ \eta^2_p = .06, \ f = .26, \ P = .81$</td>
<td>$F(2, 116) = 1.69, \ p = .19, \ \eta^2_p = .03, \ f = .17, \ P = .77$</td>
</tr>
<tr>
<td>ANOVA (TT Achievement)</td>
<td>$F(1, 117) = .47, \ p = .50, \ \eta^2_p = .004, \ f = .06, \ P = .20$</td>
<td>$F(1, 117) = 16.28, \ p = .00, \ \eta^2_p = .12, \ f = .37, \ P = .99$</td>
</tr>
<tr>
<td>ANOVA (MSET Profile)</td>
<td>$F(1, 117) = .21, \ p = .65, \ \eta^2_p = .002, \ f = .04, \ P = .12$</td>
<td>$F(1, 117) = 8.21, \ p = .01, \ \eta^2_p = .07, \ f = .27, \ P = .99$</td>
</tr>
<tr>
<td>ANOVA (BRIEF Composite)</td>
<td>$F(1, 117) = 6.05, \ p = .02, \ \eta^2_p = .05, \ f = .23, \ P = .98$</td>
<td>$F(1, 117) = 11.35, \ p = .001, \ \eta^2_p = .09, \ f = .31, \ P = .99$</td>
</tr>
</tbody>
</table>

*Note: \( \eta^2_p \) denotes partial eta squared, \( f \) denotes effect size, and \( P \) denotes power.*
Figure 1. IGT Net 4 t scores for Programmers and Controls pre and post rehabilitation (or equivalent time for Controls).
affected how inmates self-reported ECF impairments across time, $F(1, 117) = 6.05, p = .02$, partial $\eta^2 = .05$. More specifically, Controls were relatively consistent in their ratings at the pre assessment ($M = 51.60, SE = 1.37$) and post assessment ($M = 51.00, SE = 1.37$), representing a nonsignificant change, $t(49) = .68, p = .50$; whereas, Programmers reported a reduction of ECF impairments between the pre assessment ($M = 55.23, SE = 1.17$) and the post assessment ($M = 51.39, SE = 1.17$; see Figure 2), representing a significant change, $t(68) = 4.18, p = .000$. Because the BRIEF Composite is produced by combining all nine variables measured by the BRIEF-A, the two subsets of scores that make up the BRIEF Composite were also examined. The Behavioral Regulation Index (BRI) included measures of inhibition, shifting, emotional control, and self-monitoring. Among Programmers, there was a significant change between pre and post assessment on the BRI, $t(68) = 2.62, p = .01$. The Metacognition Index (MI) included measures of initiating, working memory, planning and organizing, as well as organization of materials. Among Programmers, there was also a significant change between pre and post assessment on the MI, $t(68) = 4.19, p = .000$.

For the remaining three repeated-measures analyses, no interactions were found; however, the multivariate main effect of Time was significant. Univariate analyses revealed significant differences between the pre and post scores of VF Switching Contrast, $F(1, 117) = 12.53, p = .001$, partial $\eta^2 = .10$, CW Inhibition, $F(1, 117) = 8.59, p = .004$, partial $\eta^2 = .07$, CW Inhibition/Switching, $F(1, 117) = 38.99, p = .000$, partial $\eta^2 = .25$, and CW Inhibition/Switching Contrast, $F(1, 117) = 6.33, p = .01$, partial $\eta^2 = .05$. In all cases the difference represented improvements across time. Comparisons of TT Achievement also showed an improvement across time ($p = .000$), as did MSET Profile scores ($p = .005$). However, in the absence of interaction effects, these significant main
Figure 2. BRIEF Composite t scores for Programmers and Controls pre and post rehabilitation (or equivalent time for Controls).
effects of Time stemming from the behavioural measurement of ECFs, were likely due to learning effects from repeated exposure to the testing materials, as opposed to improvement due to treatment effects.

**Hypothesis 1B: ECF Performance of NSAP-M, VPP-M, and Controls**

The decision to combine participants who had completed different treatment programs into a single group may have masked program-specific treatment effects. Therefore, further analyses were conducted to examine ECF performance across time between specific rehabilitation programs. There were a sufficient number of inmates who had completed the NSAP-M \((n = 37)\) and the VPP-M \((n = 23)\) rehabilitation programs to create more specific treatment groups for subsequent analyses. Similar to those analyses discussed above, five separate repeated-measures analyses were conducted with Program Status (i.e., NSAP-M vs. VPP-M vs. Controls) as the between-subjects variable, Time (pre vs. post) as the within-subjects variable, and ECF performance as the dependent variable(s).

Across the five repeated-measures analyses, no significant interaction effects were found (see Table 5). However, two interaction effects approached significance \((p < .06)\) and these trends are discussed in turn. First, the multivariate interaction between Program Status and Time approached significance for the repeated-measures MANOVA including IGT scores as the dependent variables. Thus, changes in IGT scores across time differed

---

3 In both MANOVAs the assumption of homogeneity of covariance matrices was violated, as indicated by a significant Box’s test, a test particularly sensitive to violations of normality. In cases with unequal sample sizes and where this assumption is violated, there are few tenable solutions (Field, 2009a). Hotelling’s Trace is robust when this assumption is violated and samples sizes are equal (Hakstian, Roed, & Lind, 1979). Pillai’s Trace is robust when this assumption is not violated and there are unequal sample sizes. Although no particular MANOVA test statistic is indicated when this assumption is violated and there are unequal sample sizes (Field, 2009a). In the case of both MANOVAs, Pillai’s Trace and Hotelling’s Trace generated identical conclusions. Therefore, Pillai’s Trace is reported to be consistent with the other analyses.
Table 5  

*Results of the Five Repeated-Measures (RM) Analyses Comparing Pre and Post Scores by Group (NSAP-M vs. VPP-M vs. Controls)*

<table>
<thead>
<tr>
<th>Type of RM Analysis (&amp; Dependent Variables Included)</th>
<th>Interaction</th>
<th>Main (Pre/Post)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MANOVA (VF Switching, VF Switching Contrast, CW Inhibition, CW Inhibition/switching, CW Inhibition/switching Contrast)</td>
<td>$F(10, 208) = 0.85, p = .58, \eta^2_p = .04, f = .20, P = .78$</td>
<td>$F(5, 103) = 16.20, p = .00, \eta^2_p = .44, f = .89, P = .99$</td>
</tr>
<tr>
<td>MANOVA (IGT Net 4, IGT Net 5)</td>
<td>$F(4, 214) = 2.37, p = .053, \eta^2_p = .04, f = .21, P = .81$</td>
<td>$F(2, 106) = 0.22, p = .80, \eta^2_p = .004, f = .06, P = .15$</td>
</tr>
<tr>
<td>ANOVA (TT Achievement)</td>
<td>$F(2, 107) = 2.39, p = .10, \eta^2_p = .04, f = .21, P = .88$</td>
<td>$F(1, 107) = 20.11, p = .000, \eta^2_p = .16, f = .43, P = .99$</td>
</tr>
<tr>
<td>ANOVA (MSET Profile)</td>
<td>$F(2, 107) = 1.91, p = .15, \eta^2_p = .03, f = .19, P = .76$</td>
<td>$F(1, 107) = 8.30, p = .01, \eta^2_p = .07, f = .28, P = .98$</td>
</tr>
<tr>
<td>ANOVA (BRIEF Composite)</td>
<td>$F(2, 107) = 2.93, p = .058, \eta^2_p = .05, f = .23, P = .94$</td>
<td>$F(1, 107) = 15.49, p = .000, \eta^2_p = .13, f = .38, P = .99$</td>
</tr>
</tbody>
</table>

*Note: $\eta^2_p$ denotes partial eta squared, $f$ denotes effect size, and $P$ denotes power.*
depending on whether an inmate had completed a specific rehabilitation program or not. Univariate analyses revealed that IGT Net 5 pre and post scores did not differ by group, $F(2,107) = .29, p = .75$, partial $\eta^2 = .005$. However, pre/post differences on IGT Net 4 approached significance, $F(2,107) = 2.96, p = .056$, partial $\eta^2 = .05$. Planned group comparisons, or contrasts, are not possible for repeated-measures designs with more than two groups within standard SPSS options; however, syntax can be written to allow comparisons (Field, 2009b). Three comparisons were performed in this manner. First, changes on IGT Net 4 from VPP-M pre and post scores were compared to changes on IGT Net 4 from NSAP-M pre and post scores, and were found not to differ significantly, $t = -.31, p = .75$. Second, changes on IGT Net 4 from Controls pre and post scores were compared to changes on IGT Net 4 from VPP-M pre and post scores, and were also found not to differ significantly, $t = -1.63, p = .11$. Lastly, changes on IGT Net 4 from Controls pre and post scores were compared to changes on IGT Net 4 from NSAP-M pre and post scores, and were found to differ significantly, $t = -2.28, p = .02$. Thus, Controls exhibited significantly more improvement across time on the IGT Net 4, compared to NSAP-M.

Previously, it was reported that Controls IGT Net 4 scores improved significantly across time (pre assessment: $M = 48.08$, $SE = 1.51$; post assessment: $M = 52.58$, $SE = 1.55$). On the other hand, although IGT Net 4 scores for NSAP-M appeared to decline (pre assessment: $M = 48.81$, $SE = 1.80$; post assessment: $M = 47.27$, $SE = 1.83$; see Figure 3), this decline was not significant, $t(36) = .73, p = .47$.

The multivariate interaction between Program Status and Time also approached significance for the repeated-measures ANOVA including the BRIEF Composite score as the dependent variable. Thus, across time, BRIEF Composite scores differentiated whether an inmate had completed a specific rehabilitation program or not. Three
Figure 3. IGT Net 4 t scores for NSAP-M, VPP-M, and Controls pre and post rehabilitation (or equivalent time for Controls).
comparisons were also performed using the SPSS syntax discussed above. First, VPP-M BRIEF Composite pre and post scores were compared to NSAP-M BRIEF Composite pre and post scores, and were found to not differ significantly, $t = 0.25$, $p = .81$. Second, Controls BRIEF Composite pre and post scores were compared to VPP-M BRIEF Composite pre and post scores, and were found to differ significantly, $t = -1.99$, $p = .05$ (see Figure 4). Lastly, Controls BRIEF Composite pre and post scores were compared to NSAP-M BRIEF Composite pre and post scores, and were also found to differ significantly, $t = -2.01$, $p = .05$. As previously reported, the decline seen in BRIEF Composite scores between pre and post assessment among Controls did not represent a significant change; however, further examination of the specific program groups revealed that significant decline between pre and post BRIEF Composite scores were present for VPP-M [$t(22) = 2.82$, $p = .01$], and NSAP-M [$t(36) = 2.70$, $p = .01$]. Looking at the subset of scores that compose the BRIEF Composite revealed that for VPP-M, significant differences were found between pre and post assessment on the BRI [$t(22) = 2.18$, $p = .04$], but not for the MI [$t(22) = 2.00$, $p = .06$], which only approached significance. Thus, VPP-M reported improvements in the areas of inhibition, shifting, emotional control, and self-monitoring. For NSAP-M the opposite pattern was evident, with significant differences found between pre and post assessment on the MI [$t(36) = 3.29$, $p = .002$], but not for the BRI [$t(36) = 1.17$, $p = .25$]. Thus, NSAP-M reported improvements in the areas of initiating, working memory, planning and organizing, as well as organization of materials.

For the remaining three repeated measures analyses, no interactions were found; however, the multivariate main effect of Time was significant for all three analyses. Not
Figure 4. BRIEF Composite $t$ scores for NSAP-M, VPP-M, and Controls pre and post rehabilitation (or equivalent time for Controls).
surprisingly, the univariate analyses revealed the identical trend as when Controls and Programmers were examined. More specifically, univariate analyses revealed significant differences between the pre and post scores of VF Switching Contrast, $F(1,107) = 13.11$, $p = .000$, partial $\eta^2 = .11$, CW Inhibition, $F(1,107) = 8.22$, $p = .005$, partial $\eta^2 = .07$, CW Inhibition/Switching, $F(1,107) = 36.02$, $p = .000$, partial $\eta^2 = .25$, and CW Inhibition/Switching Contrast, $F(1,107) = 5.24$, $p = .02$, partial $\eta^2 = .05$. In all cases the difference represented improvements across time. Comparisons of TT Achievement also showed an improvement across time ($p = .000$), as did MSET Profile scores ($p = .005$). Again, these improvements appear representative of learning effects.

In summary, across the two sets of analyses examining hypothesis 1a and 1b, only two ECF variables differed across the groups over time (i.e., only two interaction effects were present). The pattern of results indicates that Controls improved significantly on an element of the IGT, but reported no significant change in their ECF impairments over time. In comparison, those inmates who completed a rehabilitation group showed no significant change on an element of the IGT, but reported significantly fewer ECF impairments over time. Examination of specific rehabilitation programs helped to better interpret these patterns of results. Group differences seen across time on IGT Net 4 appeared to be driven by the difference between NSAP-M and Controls; however, as stated, only Controls showed significant changes (i.e., improvement) between pre and post assessment. Group differences seen across time on the BRIEF Composite score appeared to be driven by the difference between Controls and VPP-M, and Controls and NSAP-M; however, only VPP-M and NSAP-M showed significant changes (i.e., improvement) between pre and post assessment.

**Strength of Analyses and Effects**
Reliance purely on statistical significance is no longer acceptable within social science research. In order to make appropriate conclusions, one must understand the strength underlying their analyses; therefore, the power and effect sizes associated with the present analyses are also discussed. When F statistics are used to determine statistical significance, Cohen’s $f$ is the effect size index used. Cohen’s $f$ identifies the magnitude of the differences among the groups (Kotrlik, Williams, & Jabor, 2011). When using $f$ as an effect size, Cohen (1992) suggests that an $f$ of 0.10 represents a small effect, 0.25 represents a medium effect, and 0.4 or greater represents a large effect. Looking across the interaction effects reported in Table 4, one can see that the effects are generally small, indicating little practical difference between the groups among the variables of interest. However, a medium effect size was found for the interaction of Group by Time on IGT, which was also statistically significant. With regards to the analyses investigating learning effects (i.e., main effect of Time), effect sizes ranged from medium to large. When effect sizes are examined across the three groups (see Table 5), less effect size variability is seen among the interactions; although all effects are small indicating little practical differences between the groups among the variables of interest. With regards to the analyses investigating learning effects (i.e., main effect of Time), effect sizes ranged from large (for learning effects seen among D-KEFS variables), medium (for learning effects seen on the MSET Profile), and small (for learning effects seen on the IGT). In summary, generally only small effects were found for the interaction between Groups and Time, suggesting that prudence must be taken when making conclusions and stating implications regarding the magnitude of the effects. That being said, small effects are practically relevant when an intervention is effective under unfavorable conditions or in the case where a better alternative has yet to be developed (Prentice & Miller, 1992).
Power to detect an effect, if one in fact exists, is also important. Using G*Power (3.1.2), power is calculated separately for MANOVAs and ANOVAs, as well as for repeated-measures interaction effects, and within subject main effects. When two groups were compared, power ranged significantly. The interaction effects examining IGT performance and BRIEF-A self-report achieved adequate power (see Table 4). However, this was not the case for the interaction effects examining elements of the D-KEFS and MSET. Because two of the analyses were able to achieve sufficient power, the statistical significance criterion, nor sample size, appears to account for the low power seen in the other interaction effects; instead, small to negligible effect sizes are likely responsible (e.g., $\eta^2_p$ ranged from 0.03 to 0.002 in these cases). Within the two group analyses, the main effects obtained sufficient power. Among the analyses examining three groups (see Table 5), power was also sufficient to detect differences within both the interaction and main effects. Although authors often caution readers away from making strong conclusions when power is less than the standard cut-off of 0.8, this caveat may not need to be applied here due to a unique element of this dissertation. This research is unique in that it used essentially the same data for two sets of analyses, one in which the data was divided into two groups, and another where one of those groups was further divided to ultimately produce three groups of participants. In this case, when three groups were examined, power was sufficient to detect differences. The three group analyses allowed for larger effect sizes to be detected, indicating that combining the NSAP-M and VPP-M was likely diminishing the ability to detect differences. Therefore, inadequate power within the analyses examining two groups should not necessarily be looked upon as a limitation.

Clinical Meaningfulness of Inmates’ Performance
Although not formally discussed as a hypothesis, the clinical meaningfulness of ECF performance is briefly discussed to assist in the interpretation of the statistically significant and nonsignificant results reported. The issue of reporting clinical significance in addition to statistical significance is particularly germane in evaluations which look at changes after treatment (Bowersox, Saunders, & Wojcik, 2009). Results focused on between and within group statistical differences, as described above, help to detect significant changes in scores, but do little to inform the researcher or reader whether differences are meaningful. Determining whether participants’ scores moved from an impaired range to a non or less impaired range is an important element towards establishing the clinical significance of treatment. Fortunately, several of the measures used in the current dissertation apply standardized scores to assist in this endeavor. For each measure, mean pre and post scores of the relevant variables are discussed in terms of whether they represented impaired functioning; as well, the percentage of initially impaired scores is examined.

**D-KEFS.** Standardized scores for each subtest of the D-KEFS have a mean of 10 and standard deviation of 3 (Delis et al., 2001a). Examination of the current data (see Table 6) revealed that mean performance on the D-KEFS (i.e., VF Switching, VF Switching Contrast, CW Inhibition, CW Inhibition/Switching, CW Inhibition/Switching Contrast, TT Achievement) during the pre assessment fell within the normal range (i.e., $7 \leq M \leq 13$), although was predominately below the normative mean of 10. At the post assessment, inmates’ performance continued to remain in the normal range, with only half of the scores falling below the normative mean.

The expectation that mean scores falling in the normal range would show an overall improvement after rehabilitation is not reasonable, and may not be meaningful.
### Table 6

**Mean (and Standard Deviation) of ECF Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>NSAP-M (n = 37)</th>
<th>VPP-M (n = 23)</th>
<th>Controls (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CW Inhibition/Switching</td>
<td>Pre: 8.32(3.64)</td>
<td>Post: 9.16(3.50)</td>
<td>Pre: 8.39(3.65)</td>
</tr>
<tr>
<td>TT Achievement</td>
<td>Pre: 9.76(3.30)</td>
<td>Post: 10.59(2.66)</td>
<td>Pre: 9.13(2.65)</td>
</tr>
<tr>
<td>MSET Profile</td>
<td>Pre: 3.16(1.07)</td>
<td>Post: 3.19(0.91)</td>
<td>Pre: 2.83(1.27)</td>
</tr>
<tr>
<td>BRIEF Composite</td>
<td>Pre: 57.76(10.92)</td>
<td>Post: 53.97(8.90)</td>
<td>Pre: 52.61(8.38)</td>
</tr>
<tr>
<td>URICA Composite</td>
<td>Pre: 75.76(1605)</td>
<td>Post: 79.49(16.42)</td>
<td>Pre: 75.43(17.27)</td>
</tr>
</tbody>
</table>
Therefore, descriptive data is given regarding participants who initially displayed impaired scores (i.e., < 7). For VF Switching, 33.3% \((n = 23)\) of Programmers scored at least 1 SD below the mean, indicating impairment, prior to completing a rehabilitation program. At post assessment, 60.9% \((n = 14)\) of those initially impaired scored in the normal range. In contrast, 28.0% \((n = 14)\) of Controls displayed impairment during the pre assessment with 42.9% \((n = 6)\) scoring in the normal range at post assessment. For VF Switching Contrast, 33.3% \((n = 23)\) of Programmers scored in the impaired range at pre assessment, but 73.9% \((n = 17)\) scored in the normal range at post assessment; whereas, 32.0% \((n = 16)\) of Controls scored in the impaired range at pre assessment, with 81.3% \((n = 13)\) scoring in the normal range at post assessment. For CW Inhibition, 24.6% \((n = 17)\) of Programmers displayed initial impairment, with 58.8% \((n = 10)\) displaying normal performance at post assessment; whereas, 14.0% \((n = 7)\) of Controls displayed initial impairment, with 42.9% \((n = 3)\) displaying normal performance at post assessment. For CW Inhibition/Switching, 31.9% \((n = 22)\) of Programmers displayed initial impairment, with 27.3% \((n = 6)\) displaying normal performance at post assessment; whereas, 20.0% \((n = 10)\) of Controls displayed initial impairment, with 50.0% \((n = 5)\) displaying normal performance at post assessment. For CW Inhibition/Switching Contrast, 27.5% \((n = 19)\) of Programmers displayed initial impairment, with 42.1% \((n = 8)\) displaying normal performance at post assessment; whereas, 28.0% \((n = 14)\) of Control displayed initial impairment, with 57.1% \((n = 8)\) displaying normal performance at post assessment. Lastly, for TT Achievement, 21.7% \((n = 15)\) of Programmers displayed initial impairment, with 80.0% \((n = 12)\) displaying normal performance at post assessment; whereas, 12.0% \((n = 6)\) of Controls displayed initial impairment, with 100% displaying normal performance at post assessment. Therefore, although mean
performance did not show impairments, 22% to 33% of Programmers and 12% to 32% of Controls displayed initial impairments on ECFs as measured by the D-KEFS. At post assessment, 27% to 80% of Programmers who were initially impaired, and 43% to 100% of Controls initially impaired, improved their performance to fall within the normal range.

**IGT.** IGT scores are converted to age and education based $t$ scores. Bechara (2007) indicates that $t$ scores less than or equal to 39 represent ‘impaired’ functioning, scores ranging between 40 and 44 represent ‘below average’ functioning, while scores greater than or equal to 45 represent ‘non-impaired’, or normal, functioning. Examination of the current data revealed that mean IGT performance at the pre assessment was consistently in the normal range for all groups. This was also the case for the post assessment mean scores. Thus, even though statistical group differences were found on IGT Net 4, clinically all groups performed in the normal range. However, 20.3% ($n = 14$) of Programmers preformed in the impaired range on the IGT Net 4 during the pre assessment, with 64.3% showing marked improvement ($n = 3$ moving to ‘below average’ and $n = 6$ to ‘non-impaired’ range) at post assessment; whereas, 16.0% ($n = 8$) of Controls were initially impaired, with 75.0% showing improvement ($n = 2$ ‘below average’ and $n = 4$ ‘non-impaired’ range) at post assessment. For IGT Net 5, 18.8% ($n = 13$) of Programmers were initially impaired with 61.5% ($n = 8$) showing improvement ($n = 1$ ‘below average’ and $n = 7$ ‘non-impaired’ range); whereas, 18.0% ($n = 9$) of Controls were initially impaired with 100% showing improvement ($n = 1$ ‘below average’ and $n = 8$ ‘non-impaired’ range).

**MSET.** Establishing clinical significance for the MSET, which is a subtest of the BADS battery, proved more difficult. A limited range profile score is obtained for each subtest of the BADS. Profile scores are summed to produce a scaled score; however,
individual subtests cannot be readily converted to a scaled score. For the purpose of discussing clinical significance, previous research comparing normative samples and impaired samples provide some approximation. In a sample of 216 healthy adults ($M_{age} = 46.6$, $SD = 19.8$) the mean MSET Profile score obtained was $3.56$ ($SD = 0.78$), which was significantly higher than the impaired group, in this case comprising of brain injured individuals ($n = 92; M_{age} = 38.8$, $SD = 15.7$), who obtained a mean profile score of $1.99$ ($SD = 1.18$; Wilson et al., 1996). In another study, 68 normal adult controls obtained a mean profile score of $3.58$ ($SD = 0.56$) which was significantly higher than a brain injured group ($n = 30, M = 2.73, SD = 1.01$; Chan & Manly, 2002). Looking to a forensic sample, one study of 30 incarcerated non-violent adult ($M_{age} = 39.3$, $SD = 9.98$) male offenders obtained a mean profile score of $3.10$ ($SD = 0.92$; Barbosa & Monteiro, 2008). The non-incarcerated male adult ($M_{age} = 32.7$, $SD = 11.8$) comparison sample obtained a profile score similar to controls in other studies ($M = 3.83$, $SD = 0.46$). Taken together, healthy controls, who generally achieve scores greater than 3, perform better than inmates and those with brain injury. Whereas, inmates appear to perform better than brain injured individuals, who generally achieve scores less than 3. From the above research, one may extrapolate that scores less than 3 represent impairment. In the current sample, both VPP-M and Controls pre assessment MSET Profile scores were less than 3, but only slightly. In comparison, inmates’ post assessment MSET Profile scores were consistently above 3; thus, consistent with non-impaired controls. Using this definition, 26.1% ($n = 18$) of Programmers were initially impaired with 55.6% ($n = 10$) showing improvement (i.e., obtaining scores $> 3$) at post assessment; whereas, 32.0% ($n = 16$) of Controls were initially impaired with 56.6% ($n = 9$) showing improvement at post assessment.
**BRIEF-A.** On the BRIEF-A, standardized scores (t scores with $M = 50$, $SD = 10$) are used to establish a participant’s level of perceived impairment of ECFs. Traditionally, BRIEF-A t scores greater than or equal to 65 are considered clinically significant, with higher scores indicating concerns with ECFs (Roth et al., 2005). Examination of the current data revealed that self-reported ECF impairments during the pre assessment did not reach the threshold for concern (i.e., group means were not $\geq 65$). Similarly, during the post assessment group means continued to fall below this threshold. Thus, even though statistical group differences were found on the BRIEF Composite, clinically all groups performed in the normal range. Only 13.0% ($n = 9$) of Programmers self-reported a significant frequency of ECF problems to obtain a t score of $\geq 65$ at the initial assessment, with 44.4% ($n = 4$) obtaining marked improvement (i.e. t scores < 65); whereas, 10.0% ($n = 5$) of Controls were initially impaired, with 40.0% ($n = 2$) showing improvement.

In summary, the current sample of inmates appeared to be functioning in the normal range on many behavioural tasks measuring ECFs (i.e., D-KEFS, IGT), and self-report measure of ECFs (i.e., BRIEF-A) at both the initial and post assessment when overall means were examined. When impaired performance was isolated, 13.0% to 33.3% of Programmers showed initial impairment on specific variables. Similarly, 10.0% to 32.0% of Controls showed initial impairment. Among Programmers, 27.3% to 80% of those initially impaired, performed in the average range at post assessment; whereas, 40.0% to 100% of Controls performed in the average range at post assessment.

**Hypothesis 2: Motivation and Criminal Attitude Changes**

To uncover whether motivation for change, as assessed by the URICA, was impacted over the course of correctional programming, a repeated-measures ANOVA was
conducted with Program Status (i.e., Programmers vs. Controls) as the between-subjects variable, Time (pre vs. post) as the within-subjects variable, and URICA Composite as the dependent variable. A significant multivariate interaction effect between Program Status and Time was revealed, $F(1, 117) = 5.28, p = .02$, partial $\eta^2 = .04$. Therefore, self-reported motivation for change, across time, differed for Programmers and Controls. Recall that higher URICA scores represent increased motivation to change. URICA Composite scores decreased from pre assessment ($M = 75.22, SE = 17.54$) to post assessment ($M = 73.42, SE = 16.82$) for Controls, suggesting reductions in motivation for change; however, this reduction was not significant, $t(49) = 1.33, p = .19$. In comparison, URICA Composite scores increased from pre assessment ($M = 73.87, SE = 17.58$) to post assessment ($M = 77.58, SE = 15.08$) for Programmers, representing a significant improvement [$t(68) = -2.08, p = .04$], suggesting improved motivation for change after completing a rehabilitation program (see Figure 5). No significant multivariate main effect of Time was found.

As part of the exploratory nature of the investigation of motivation and criminal attitudes, the current researcher initially sought to investigate whether inmates’ self-reported motivation for change was related to their performance on ECFs. It was proposed that inmates’ pre assessment URICA score may be positively related to changes in their ECF scores, with the idea that perhaps an individuals’ initial level of motivation impacted their ability to improve their ECFs after completing rehabilitation. Table 7 shows the Pearson correlation coefficient between URICA Composite pre scores of Programmers and the change score from the 10 ECF scores (i.e., change score = time 2 ECF score – time 1 ECF score). No correlations were significant, suggesting that initial
Figure 5. URICA Composite scores for Programmers and Controls pre and post rehabilitation (or equivalent time for Controls).
Table 7

*Pearson Correlation Coefficients between ECF Change Scores and Pre URICA Composite Score*

<table>
<thead>
<tr>
<th>Variable</th>
<th>URICA Composite</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>VF Switching</td>
<td></td>
<td>.09</td>
<td>.31</td>
</tr>
<tr>
<td>VF Switching Contrast</td>
<td></td>
<td>-.02</td>
<td>.81</td>
</tr>
<tr>
<td>CW Inhibition</td>
<td></td>
<td>.04</td>
<td>.65</td>
</tr>
<tr>
<td>CW Inhibition/Switching</td>
<td></td>
<td>-.12</td>
<td>.21</td>
</tr>
<tr>
<td>CW Inhibition/Switching Contrast</td>
<td></td>
<td>-.11</td>
<td>.24</td>
</tr>
<tr>
<td>TT Achievement</td>
<td></td>
<td>.05</td>
<td>.61</td>
</tr>
<tr>
<td>MSET Profile</td>
<td></td>
<td>-.04</td>
<td>.69</td>
</tr>
<tr>
<td>IGT Net 4</td>
<td></td>
<td>.12</td>
<td>.18</td>
</tr>
<tr>
<td>IGT Net 5</td>
<td></td>
<td>.08</td>
<td>.38</td>
</tr>
<tr>
<td>BRIEF Composite</td>
<td></td>
<td>-.01</td>
<td>.91</td>
</tr>
</tbody>
</table>
levels of motivation for change were unrelated to changes in ECF performance following rehabilitation programming.

A similar procedure was conducted for criminal attitudes, as assessed by the CSS-M. To uncover whether criminal attitudes were impacted over the course of correctional programming, a repeated-measures ANOVA was conducted with Program Status (i.e., Programmers vs. Controls) as the between-subjects variable, Time (pre vs. post) as the within-subjects variable, and CSS-M Total as the dependent variable. A significant multivariate interaction between Program Status and Time was found, \( F(1,117) = 10.95, p = .001, \) partial \( \eta^2 = .09 \). Thus, changes in criminal attitudes, across time, differed for Programmers and Controls. The CSS-M Total stayed relatively consistent between pre assessment (\( M = 23.06, SE = 1.91 \)) and post assessment (\( M = 23.74, SE = 1.74 \)) for Controls, representing a nonsignificant difference, \( t(49) = -0.55, p = .58 \). However, the CSS-M Total decreased between pre assessment (\( M = 27.77, SE = 1.62 \)) and post assessment (\( M = 22.03, SE = 1.48 \)) for Programmers, representing a significant difference, \( t(68) = 4.13, p = .000 \) (see Figure 6). Recall that lower scores represent fewer criminal attitudes. There was also a significant multivariate main effect of Time, \( F(1,117) = 6.80, p = .01, \) partial \( \eta^2 = .06 \). The results indicate that after completion of rehabilitation programs, Programmers reported a significant decrease in criminal attitudes.

To investigate whether criminal attitudes were related to ECFs, correlations between pre CSS-M Total scores and the 10 ECF pre scores were examined, as well as correlations between pre CSS-M Total scores and the 10 ECF post scores. It was proposed that inmates’ initial quantity of criminal attitudes may be negatively related to their performance on ECFs (either pre or post assessment), as a result of the established
Figure 6. CSS-M Total scores for Programmers and Controls pre and post rehabilitation (or equivalent time for Controls).
association of ECF deficits among offenders. There was no theoretical support to indicate that Programmers' criminal attitudes would be related to their ability to improve their ECFs; thus, change scores were not examined. No significant correlations were found, indicating that the extent of criminal attitudes was not related to ECFs before or after completing a correctional rehabilitation program, or equivalent time for Controls (see Table 8).

In summary, participants who completed a rehabilitation program reported a significant increase in their motivation for change, as well as a significant decrease in their criminal attitudes; however, neither change in motivation, nor criminal attitudes, were related to ECF performance.

Discussion

The primary goal of the present dissertation was to longitudinally examine ECFs of various groups of inmates, specifically those who had completed correctional rehabilitation programming, and those who had not. Two sets of analyses were completed to examine ECFs. First, inmates who had completed one of three rehabilitation programs (‘Programmers’) were compared to Controls across time (i.e., pre and post rehabilitation programming). Second, ECF performance of inmates who completed NSAP-M, VPP-M, or who belonged to a comparison group, was also compared across time. As can be gleaned, the first set of analyses was a more broad evaluation of whether rehabilitation programs, in general, affected ECFs, while the second set of analyses provided program specific information regarding changes in ECFs. For ease of comprehension, the results and interpretations of both sets of analyses are summarized together. Following this, the implications of a non-impaired sample will be
Table 8

*Pearson Correlation Coefficients between ECF Scores and Pre CSS-M Total Score, as well as ECF Scores and Post CSS-M Total Score*

<table>
<thead>
<tr>
<th>Variable</th>
<th>CSS-M Total &amp; Pre ECF Score</th>
<th>CSS-M Total &amp; Post ECF Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$p$</td>
</tr>
<tr>
<td>VF Switching</td>
<td>.01</td>
<td>.93</td>
</tr>
<tr>
<td>VF Switching Contrast</td>
<td>-.11</td>
<td>.23</td>
</tr>
<tr>
<td>CW Inhibition</td>
<td>.03</td>
<td>.74</td>
</tr>
<tr>
<td>CW Inhibition/Switching</td>
<td>-.06</td>
<td>.52</td>
</tr>
<tr>
<td>CW Inhibition/Switching Contrast</td>
<td>-.05</td>
<td>.61</td>
</tr>
<tr>
<td>TT Achievement</td>
<td>.11</td>
<td>.23</td>
</tr>
<tr>
<td>MSET Profile</td>
<td>-.10</td>
<td>.28</td>
</tr>
<tr>
<td>IGT Net 4</td>
<td>-.11</td>
<td>.22</td>
</tr>
<tr>
<td>IGT Net 5</td>
<td>-.13</td>
<td>.15</td>
</tr>
<tr>
<td>BRIEF Composite</td>
<td>-.04</td>
<td>.71</td>
</tr>
</tbody>
</table>
discussed, as well as conclusions regarding motivation to change and criminal attitudes, and recommendations regarding future forensic research.

**ECF Differences across Groups**

Across the six measures of ECFs, which encompassed 10 variables of interest, Programmers performed significantly different than Controls on two variables (i.e., BRIEF Composite and IGT Net 4), with no other significant differences uncovered between pre and post assessment across the groups. Although both variables assessed different ECFs, both of the measures represented methods of assessing how ECFs impact an individual’s participation in the real world, following the discussion of Chan et al. (2008). Thus, within this sample, differences were seen between the groups regarding how they were able to participate in everyday tasks requiring ECFs, but not in how impaired or disabled they appeared.

**BRIEF-A.** Across time, Programmers reported improved ECFs on the BRIEF-A, compared to Controls who reported relatively consistent pre and post scores. More specifically, both rehabilitation groups (NSAP-M and VPP-M) self reported significant improvements compared to Controls. Further analyses revealed that following the completion of rehabilitation programming, NSAP-M reported improvements in initiating, working memory, planning and organizing, as well as organization of materials; whereas, VPP-M reported improvements in inhibition, shifting, emotional control, and self-monitoring. However, neither Programmers nor Controls showed impaired mean performance on the BRIEF-A during the baseline assessment indicating normal performance across the groups, with only 13% of Programmers and 10% of Controls reporting initial impairment.
The self-report adult version of the BRIEF is a relatively novel measure, with no published studies utilizing forensic samples for comparison. Looking to studies employing non-forensic samples reveals that compared with adult substance abusers partaking in residential treatment, inmates in the current dissertation reported fewer impaired ECFs (Gunn & Rickwood, 2009). This was the case even among NSAP-M participants, who presumably were most similar to substance abusers seeking treatment. In another study, male community controls obtained scores comparable to the current forensic sample. In comparison, men seeking treatment for hypersexual behaviours obtained higher, yet still not clinically significant scores (i.e., not ≥ 65; Reid, Karim, McCrory & Carpenter, 2010).

**IGT.** Controls showed a significant improvement across time on a segment of the IGT (i.e., IGT Net 4) while Programmers performance remained stable. The IGT is a measure of real-life reward-dependent decision making, particularly involving cognitive flexibility and response monitoring, as well as the regulation of affect and motivation. The IGT Net 4 score is derived from performance on the 61st to 80th card selection (also referred to as ‘Block 4’). Blocks 4 and 5 are considered decision making under risk (Brand et al., 2007) since these selections are made when participants are aware of the risk contingencies due to learning in the first three blocks of trials. If not impaired, participants can utilize learned cognitive strategies to guide advantageous card selections (Bass & Nussbaum, 2010); otherwise, participants continue to make less advantageous and more risky card selections, resulting in poorer performance. Thus, in the current study, Controls displayed a decrease in risky decision making during Block 4, compared to Programmers. When examined by specific program, the IGT Net 4 improvement displayed by Controls differed significantly only from the performance of NSAP-M. However, given that
neither Programmers nor Controls showed impaired functioning at baseline, and no significant group differences were found on the IGT Net 5 (also considered decision making under risk), the improvement seen by Controls is given less clinical importance.

Non-impaired mean performance was unexpected, particularly among NSAP-M. There is a large body of research documenting IGT deficits in substance abusers (e.g., Levi, Nussbaum & Rich, 2010; van der Plas, Crone, van den Wildenberg, Tranel, & Bechara, 2009; Verdejo-Garcia and Perez-Garcia, 2007). Within forensic samples, only psychopaths have been documented to score in the non-impaired range on the IGT (Losel & Schmucker, 2004); however, it is unlikely that the composition of the current sample was predominately psychopaths. Although a unanimously non-impaired forensic sample is rare, differentiation of functioning has been observed, indicating that offenders do not as a whole perform in the impaired range. For example, Bass and Nussbaum (2010) used the IGT to distinguish between non-predatory offenders, scoring in the below average range, and predatory offenders, scoring in the impaired range.

**MSET and D-KEFS.** Inmates who completed rehabilitation programs did not show differential improvement in the areas of general planning, strategy formation, or self-monitoring, as assessed by the MSET. However, mean baseline performance of VPP-M and Controls, but not NSAP-M, showed clinical impairment. Yet overall, only 26% of Programmers and 32% of Controls showed initial impairment. All post assessment scores were in the normal range, indicating that both VPP-M and Controls improved; yet this improvement is likely representative of learning effects.

Programmers did not differ from Controls across time on the variables from the VFT, CWIT, TT, and MSET measures. Thus, inmates who completed rehabilitation programs did not show differential improvement in the areas of cognitive flexibility (also referred
to as set shifting), inhibition, spatial planning, rule learning, or establishing and maintaining a cognitive set, as measured by the three D-KEFS subtests. However, neither Programmers nor Controls displayed impaired mean baseline performance on the subtests of the D-KEFS, with only 22% to 33% of Programmers and 12% to 28% of Controls showing initial impairment.

Although authors have reported deficits in offender populations on the D-KEFS or similar tasks, studies also exist where differences were not found between offenders (or similarly antisocial individuals) and controls on the Stroop (White et al., 1994) on which the CWIT is based, the Tower of London task (Cauffman, Steinberg, & Piquero, 2005) on which the TT is based, and verbal fluency type tasks (Oosterlaan, Scheres, & Sergeant, 2005; Toupin, De´ry, Pauze, Mercier, & Fortin, 2000). In addition, Mullin and Simpson (2007) reported D-KEFS scaled scores in the normal range for 45 maximum security incarcerated males, consistent with scores obtained by inmates in the current dissertation. These authors concluded that the range of scores represented a relative weakness compared with the general population. Using this standard, D-KEFS scores from the current sample also represented a ‘relative weakness’, in that scores were below the mean; however, they did not represent a clinically significant weakness. Broomhall (2005) reported VFT and CWIT scores for 25 incarcerated violent males whose violence was categorized as either reactive in nature (i.e., impulsive with a lack of planning) or instrumental in nature (i.e., high degree of planning involved). Results revealed instrumental offenders scored in the normal range; whereas, reactive offenders showed significant impairments on several variables. Both studies showed that offenders typically performed more poorly than the normative sample, but still within the normal range on the D-KEFS. However, there is some evidence that different types of offenders
(e.g., those engaging in reactive violence, a construct not measured in the current dissertation) may perform in the clinically significant range.

Irrespective of groups, significant improvements between baseline and post assessment were displayed for VFT, CWIT, TT, and MSET variables. In all cases the difference represented improvements across time. However, in the absence of interaction effects, the significant effects appear to speak more to repeated exposure to the testing materials, also referred to as learning effects, opposed to improvement due to treatment or other effects. Possible causes of practice or learning effects include remembering previous responses, reduced anxiety, development of abilities, and improved strategies for test-taking (Hausknecht, Halpert, Di Paolo, & Moriarty Gerrard, 2007). The presence of learning effects is not isolated to the present study. In the absence of intervention, SET profile scores of an adult psychiatric sample improved between initial assessment and three weeks later (Jelicic, Henquet, Derix & Jolles, 2001). As with most tasks measuring ECFs, the SET relies heavily on novelty for accurate assessment and may be especially vulnerable to practice effects. The TT may also be susceptible to learning effects due to the nature of the task. Participants may gain strategies for solving the towers that could carry over to additional administrations. Learning effects among adults have been observed on Tower of Hanoi performance across several years (Davis & Klebe, 2001; Ronnlund, Lovden, & Nilsson, 2008). Similarly, the CWIT requires participants to learn inhibition to one visual cue while attending to another. It is possible that increased exposure to the task could make discrimination of the visual cues easier. Learning effects have been observed in adults on successive administrations of the Stroop (Davidson, Zacks, & Williams, 2003). The VFT is somewhat unique in that it provides an alternative version to be used during re-administrations which asks participants to name words
corresponding to novel letters and categories. Thus, in the case of the VFT, learning effects are more difficult to interpret.

The original empirical question asked: Does an inmate’s ECFs change over the length of a correctional rehabilitation program, beyond what is likely as a result of learning effects? The current data does not appear to provide supportive evidence towards this question. After completion of rehabilitation programming, only 2 of 10 ECF variables revealed a significant difference, one behavioural measure, and one self-report measure; yet the behavioural measure revealed improvement only in Controls, which represented improvement in the opposite direction of the hypothesis. In addition, 7 of the 10 ECF variables showed what are likely learning effects. The second empirical question inquired whether programs differentially affected ECFs. In both cases of significant interaction effects, further examination of the specific program provided little additional information. Thus, specific rehabilitation programs do not appear to differentially affect ECFs. However, unexpectedly, participants represented a relatively non-impaired sample, making it difficult to adequately address the empirical questions at hand. Implications of this are discussed below.

**The Paradox of a Non-Impaired Sample**

The lack of ECF impairment displayed by the current participants is briefly explored. Neither Controls nor Programmers displayed clinically significant mean impairments across the battery of ECF measures during the initial assessment. Thus, even those inmates targeted for rehabilitation did not display ECF impairments on the measures selected for the current dissertation. Yet extant research suggests that inmates demonstrate ECF deficits. One reason for the discrepant results in the current dissertation could relate to the specific prison environment used for participant selection. Important to
this discussion, Fenbrook and Bath Institution are not typical among national or international medium security facilities. They boast a more open, supportive environment where the underlying intention is improved life skills (e.g., learning to cooperate, resolve disputes independently, maintain leisure and work schedules). Because all inmates, even those not enrolled in programming, are still receiving some type of intervention, such as work programs, access to schooling, and appointments with therapists, it may be more difficult to uncover differences. In fact, as Losel pointed out in a 2001 paper, “it may seem ironic, but the more rehabilitative elements we find in regular prisons, the harder it becomes to confirm effects of a treatment program” (p. 73). The milieu effect may also help to explain why this sample failed to display clinically significant ECF deficits.

In addition to the milieu effect that may have been at play, several other reasons explaining why inmates were not initially impaired are possible. One possibility is that self-selection of participants may have influenced the results. As mentioned in more detail later, less than 60% of initially recruited participants completed the study, potentially affecting the outcome. In addition, previous studies citing offenders as having deficits have made these conclusions based on comparisons of non-forensic ‘normals’ in which case there could likely be a statistical, but not necessarily a clinically significant, difference. Related to this, clinical significance is not readily considered when examining ECF deficits in offenders. Instead, ‘relative’ deficits compared to non-forensic controls often predominate. In addition, most measures of ECFs are designed to be sensitive to clinical impairments in functioning. Although useful in clinical settings, these measures may not be sensitive enough to capture subclinical ECF problems or individual differences in ECFs that may impact everyday functioning. It is possible that the
problems in ECFs associated with some groups of offenders result from such subclinical or individual differences in ECFs (Olgivie et al., 2011).

**Motivation for Change and Criminal Attitudes**

Inmates who completed a rehabilitation program reported a significant increase in motivation for change after programming, compared to Controls. Previous research reported similar URICA Composite baseline scores within general forensic samples (McMurran et al., 1998), but lower motivation for change pre and post scores among sex offenders engaging in rehabilitation (Serin & Kennedy, 1997). In contrast to the current results, others have reported a decline in URICA scores after correctional rehabilitation programming (e.g., McMurran, Theodosi, & Sellen, 2006). However, the current sample exhibited relatively high levels of motivation both initially and following programming. Friendship et al. (2003) asserts that completing treatment involves some degree of voluntary participation; thus, offenders who complete treatment are more motivated to change their offending behaviour than comparison offenders. However, some have reported that self evaluation of the related concept of treatment readiness is a poor predictor of treatment outcome (Fishbein et al., 2009).

After completing a correctional rehabilitation program, inmates also reported a significant decrease in criminal attitudes, compared to Controls, which is consistent with decreases found in other studies examining forensic rehabilitation program completion in adult males (e.g., Ashord, Wong, & Sternbach, 2008; Keeling, Rose & Beech, 2006). The post assessment results of Programmers showed an increased support for lawful behaviour and increased acceptance of the need for the justice system, and decreased support for a criminal lifestyle. It is commonly accepted that, to be effective, offender rehabilitation must target criminal thoughts and behaviours, as well as the risk factors and
criminogenic needs associated with these behaviours (Andrews & Bonta, 2006; Andrews et al., 1990). In one study, a clinical cut-off score of 19 was used to categorize whether criminal attitudes were of significant concern, with scores 19 and above representing more significant criminal attitude (Morgan et al., 2010). Baseline scores among the current sample were greater than this cut-off. Baseline scores were somewhat higher than those observed from another incarcerated Canadian male sample (Simourd & Olver, 2002). In fact, CSS-M baseline scores were more in keeping with those observed from a sample of incarcerated women with mental illness (Morgan et al., 2010). Mental illness has been associated with increases in criminal attitudes (see Morgan et al., 2010), an aspect not assessed in the current dissertation. Although post assessment scores for NSAP-M and Controls remained above the cut-off, indicating continued criminal attitudes, VPP-M obtained post assessment scores just at the cut-off score, suggesting a positive shift towards non-criminal attitudes. The positive relationship between antisocial or criminal attitudes, and criminal behaviour is well documented in the literature (e.g., Stevenson, Hall, & Innes, 2003), with the presence of criminal attitudes explaining up to 40% of the total variance in three year recidivism rates (Andrews & Bonta, 1998).

In summary, Programmers reported significant increases in their motivation for change, as well as decreases in their perception of criminal attitudes, compared to Controls. Both motivation and criminal attitudes are identified correctional programming targets. In this regard, rehabilitation appeared to be successful, as indicated by self-report of participants.

**Future Forensic Research Recommendations Inspired by the Current Dissertation**

The current dissertation did not reveal significant ECF improvements as hypothesized; however, other important findings or implications can be discussed. Several issues
existing within the current forensic program rehabilitation literature were initially reviewed and addressed within the conceptualization and methodology of the current dissertation. Some of the methods used to overcome these critiques may be important to consider for future forensic research. As well, additional recommendations have come to light as a result of the completion of the current dissertation. Some of these issues have been voiced elsewhere (e.g., Ogilvie et al., 2011). What follows is a description of prospective suggestions for correctional researchers outlining the issues of participant selection, comparison groups, measurement selection, learning effects, occasions of data collection, and substance use screening. These recommendations are reviewed both as suggestions for future research, but also in many cases, as limitations within the present dissertation.

A discussion of participant selection begins the dialogue of recommendations gathered from the current dissertation. All inmates awaiting programming were invited to participate in the current study. Fewer than 3% were excluded, yet only 59% completed the research. Some declined outright, while a smaller group did not complete the study. Although comparison of demographic data suggested that those who participated and those who dropped out were similar, there was no method of evaluating whether differences in ECFs, either initially or after programming, existed between those who self-selected to participate, and those who declined. It is possible that inmates who volunteered or followed through with the research were more motivated, organized, less impulsive, or possessed better memory capacity; in other words, possessed superior ECFs. Future forensic research should strive to avoid self-selection of inmates. One recommendation would be to make the research component an integrated part of the rehabilitation program curriculum in order to obtain data from a larger percentage of
program participants. It is recognized that ethical issues may arise regarding inmates’ right to refuse research participation without consequence; thus, certain precautions would also need to be implemented. Integrating research elements regarding ECFs into the program curriculum may allow for increased access to valuable data, decreased research costs and required resources, and may assist program reform.

In addition to participant selection, this dissertation has highlighted the importance of including an appropriate comparison group when examining correctional programming. As previously discussed, some forensic researchers have examined the benefits of completing a rehabilitation program without the inclusion of a comparison group (e.g., Hodel & West, 2003; Mullin & Simpson, 2007). These results ignore the possibility of other influences. Others have recognized the importance of including a control group, but have failed to fully appreciate the purpose of such a group. For example, Hollin et al. (2008) compared offenders who engaged in several types of rehabilitation programming, to offenders who were not mandated to programming. Thus, these two groups may have differed in important aspects. Initial study results revealed that the reconviction rate of programmers was 67%, while the reconviction rate of controls was 65%, suggesting that no positive impact was made by programming. However, no mention was made that those mandated to programming may have been at increased risk of re-offending initially, compared to those not mandated. In non-forensic settings, randomized control trials are often the gold standard; however, when behaviour change related to public safety is the focus, quasi experimental designs have been supported (Munro, 2005; Slade & Priebe, 2001). As such, waitlist control groups play an important role within forensic research, and are strongly recommended for future forensic studies examining treatment. A waitlist control group assists in excluding the presence of other variables responsible for change,
PRISON REHABILITATION AND EXECUTIVE COGNITIVE FUNCTIONS

takes into consideration the relative risk of the offenders, as well as assists in addressing
the presence of learning effects in re-administration of tests. This latter issue represents
an important finding, and subsequent recommendation; however, before this can be
addressed, further attention should be paid to measurement selection in general.

Within the present discussion, measurement selection can be thought of in terms of
what constructs should be measured, and how should those identified constructs be
measured. To address the ‘what’, let’s return to the idea of intermediary treatment targets.
Earlier in this dissertation, Friendship et al.’s concept of intermediary treatment targets
was briefly introduced (2003). Intermediary treatment targets refer to the short-term
goals and outcomes that come as a result of correctional programming. Published
forensic research typically does not focus on short-term goals and outcomes; instead,
recidivism appears to be the measurement of choice. Recidivism represents a long-term
goal or outcome. When forensic program research focuses exclusively on recidivism as
the outcome, the mechanisms of change go unobserved (Pawson & Tilley, 1994). ECFs
were proposed as a set of possible intermediary treatment targets that, if ameliorated, may
impact recidivism, or in the least, contribute to the impact. The current results do not
appear to provide strong support for ECFs as a short-term goal or outcome; however, the
limitations of the study suggest that further exploration is required before ECFs can be
completely ruled out. There is much research insinuating that ECFs play an important
role in the goal towards reducing recidivism; thus, consistent measurement of ECFs as
possible intermediate treatment goals within forensic treatment research is recommended.

Not surprising given the topic of this dissertation, ECFs were identified as important
variables to continue to measure in future forensic research. However, suggestions
regarding how they should be measured also warrant discussion. Many previous studies
limit their assessment of ECFs to one type of measure. Including various categories of measurement allows researchers to speak to the various levels of functioning impacted. The BRIEF-A and IGT were selected to represent a functional measure of participation, or how an individual appears to interact or participate in the real world, as discussed by Chan et al. (2008). Subsets of the D-KEFS were selected to represent measurable impairment in functioning, and the MSET was selected to represent disability, or how disabled an individuals’ functioning appears. Participation appeared to be the only domain affected within the study. Although there was not a straightforward interpretation of the results of the functional measures of participation (i.e., on the IGT Controls improved, while on the BRIEF-A Programmers improved), including measures from multiple domains assisted in isolating the functional level impacted, and thus is recommended for future research. The importance of including multiple categories of measures can be seen in studies that find discrepant results between various categories of ECF measures. In a group of adults with verified cognitive complaints, such as diagnosis of mild cognitive impairment, ECF difficulties were reported on a measure of participation (i.e., BRIEF-A), whereas no clinically meaningful difficulties were found on measures that focused on impairment (i.e., neuropsychological measures of ECFs including a subtest of the D-KEFS; Rabin et al., 2006). There are disadvantages to each category which can be overcome with the inclusion of multiple categories of measures. For example, one drawback to some measures of participation is that they rely on self-report which may be more susceptible to demand characteristics or social desirability bias. Thus, in the present dissertation, inmates may have altered their responses in an attempt to present what they perceived as the desired outcome (i.e., fewer ECF impairments after completion of rehabilitation programming), or to be perceived in a
positive light. The inclusion of multiple categories of measures helped to identify whether differences were present across a specific, or different, functionality.

A further issue related to measurement selection relates to the likelihood of learning effects with the re-administration of measures. Results of the current dissertation revealed significant learning effects among many of the ECF variables. As noted by Chan et al. (2008), theoretically, only novel tasks can evaluate deficits of ECFs. Clearly, tasks can only be truly novel during their first administration. Repeated administration of measures can result in improvements due to learning. As such, some argue that ECF measures cannot be reliably re-administered (Lowe & Rabbitt, 1998), while others suggest that ECFs “can be measured reliably across at least two time points, even after relatively short delays (4 to 8 weeks)” (Ettenhofer, Hambrick, & Abeles, 2006, p. 610). However, few forensic researchers discuss, or attempt to overcome, this problem. In the present dissertation an alternative form of the VFT was used in anticipation of possible learning effects; but several measures selected did not have a known alternative form to be used at re-administration (e.g., MSET, CWIT, TT). Future forensic research involving ECFs should acknowledge the presence of these effects. Certain techniques, used alone or in conjunction, may decrease the likelihood of learning effects. Solutions such as, employing measures that offer alternative forms for repeat administration and/or randomly assigning alternative forms so that answers or techniques can not be as easily shared with new participants, may decrease undesired learning effects.

As implied in the discussion of learning effects, re-administration of measures plays an important role in detecting change. Assessment needs to routinely include baseline and post measurements in order to effectively examine change, as well as uncover the mechanisms responsible for change. Establishing a baseline of functioning and repeating
assessment following treatment, as was completed for the current dissertation, examines immediate change. However, follow up data, which goes beyond the completion of the intervention and assesses change outside of the confines of treatment, is another important aspect of change investigation. The present study measured ECFs an average of two weeks after the completion of programming. It is possible that more time was required to consolidate learning of new skills. Although immediate improvements following rehabilitation of ECFs have been reported within non-forensic populations; thus far, evidence is not readily available to provide examples of immediate or later ECF improvements of offenders after correctional rehabilitation. The current dissertation helped to address part of this question, but follow up data (e.g., 6 or 12 months after programming) examining ECFs of those who completed rehabilitation may provide further insights. This recommendation, although reasonable, can be extremely difficult in practice. Obtaining follow up data even among inmates who continue to be incarcerated can be challenging. To illustrate, in the current dissertation 27 participants did not provide post data due to transfer to another institution, release to the community, or refusal, despite the short period between baseline and post assessment.

In addition to the above recommendations that have come out of the present dissertation, one final recommendation is proposed. This suggestion may represent the most novel, as it is one not previously encountered in the literature. Anecdotal evidence encountered during the current dissertation suggests that screening for substance intoxication on the day of testing may be warranted. Beyond the fact that one of the groups (i.e., NSAP-M) represented previous substance abusers, it is possible, and perhaps even likely, that some participants were also current substance abusers. Anecdotally, participants shared narratives of current substance using, as well as producing (i.e.,
brewing alcohol in their cells), and obtaining substances (i.e., receiving drugs in the mail). In addition, several participants failed to attend scheduled testing sessions, later citing ‘hangover’ as the reason.

Routine substance use by offenders has been documented in Canadian, American, and European incarcerated populations, with cannabis and alcohol being the most commonly abused substances, followed by cocaine, heroin, and amphetamines (Plourde & Brochu, 2002; Simpler & Langhinrichsen-Rohling, 2005; Strang et al., 2006). Substance abuse or dependence is significantly associated with both antisocial behaviour and ECF deficits (Ogilvie et al., 2011). Current evidence suggests that there is a complex and reciprocal relationship between ECFs and substance use, with poorer ECFs increasing the risk of engaging in substance use, as well as resulting from prolonged substance use (Blume & Marlatt, 2009; Clark, Thatcher, & Taper, 2008; Verdejo-Garcia, Lopez-Torrecillas, de Arcos, & Perez-Garcia, 2005). Substance use has the potential to temporarily impair ECFs. Substances that have been linked to impairments include alcohol (Bates, Bowden, & Barry, 2002), cannabis (Verdejo-Garcia et al., 2005), cocaine (Aharonovish, Nunes, and Hasin, 2003), and methamphetamine (Kalechstein, Newton, & Freen, 2003).

Although participants did not display ECF impairment, the anecdotal and previous research evidence suggests substance abuse could pose problems in future research; thus, substance screening on the day of testing may be justifiable. An important protection standard within correctional research is that information obtained during a research project can in no way be used to affect incarceration, case management plans, or decisions of release, either positively or negatively. This protection would need to be retained with respect to substance screening in order to protect participants.
In summary, several recommendations have been made for use in future research. Together, these represent knowledge gathered from decades of forensic research focused on improving offender outcomes. Many examples of previous research have incorporated some or even many of these recommendations; however, practical limitations such as resource limitations may restrict researchers’ ability to incorporate all recommendations. The list is intended to represent the benchmark in methodology of forensic ECF programming research. The present dissertation gives an example of how some of the recommendations can be incorporated into one study. Specifically, in the current methodology the recommendations of including an appropriate comparison group, measuring ECFs using various categories of measures, and using a pre/post design, were employed. Others were not addressed and in some cases represent limitations of the present dissertation, including self-selection of participants, measurement selection focused on avoiding learning effects, incorporating a follow up period beyond treatment completion, and substance use screening.

**Additional Limitations**

In addition to the limitations discussed within the recommendations section, three additional limitations are worth noting including low sample size, impact of aboriginal ancestry, and use of non-forensically validated measures. Firstly, sample size is identified as a study limitation. One of the significant challenges of conducting research involving forensic rehabilitation programming, specifically within Canada, is the reality that programming occurs relatively infrequently and includes few participants at a given time. During the 12 months of participant recruitment, 17 targeted rehabilitation groups were run across the three targeted prisons. This totaled only approximately 200 possible research participants, including those who were wait-listed for a program. Other
researchers have encountered similar issues. When targeting male inmates who had completed a violence rehabilitation program at one of two Canadian prisons, only 70 participants were recruited over a 5 year period, with only a pool of approximately 150 additional participants (who met criteria for the program) to draw from (Serin et al., 2009). Also identified as a sample issue, it is important to point out that the current data are based on a restricted sample of relatively non-impaired incarcerated males. Therefore, the results may not be generalizable to other incarcerated samples.

Another sampling issue relates to the ancestry of the current sample. Indigenous people make up 17% of the Canadian correctional population (CSC, 2007). The current sample is reflective of this, with 13% of participants self-identifying as either Aboriginal or Inuit. CSC has dedicated rehabilitation programs focusing on violence and substance abuse which incorporate traditional native teachings. Aboriginal inmates are permitted to enroll in either the traditional teaching rehabilitation program or general rehabilitation program (e.g., NSAP-M, VPP-M, MIFVPP), or a combination of both; therefore, some Inuit or Aboriginal participants may have been simultaneously enrolled in specific programs not included in the study. Losel (2001) argued that program evaluations are complicated by offenders receiving multiple interventions, such as multiple programming. General rehabilitation programs have been shown to be differentially effective for aboriginal offenders, compared to non-aboriginal offenders (Nafekh et al., 2009). It is unclear how exposure to components of other programs may have affected ECF performance.

A final limitation to the current dissertation is that the measures used had not previously been validated for use with populations of inmates or offenders. No formal normative data was available for offenders. Unfortunately this was even the case for the
one measure, the CSS-M, specifically designed for use with a forensic population. Thus, instead of comparing the results with normative forensic data to assist with interpretation, scores were compared with other published studies that utilized the measures with various small forensic samples. It is conceivable that unaccounted for biases present in the studies used for comparison impacted the interpretation of results of the current dissertation.

In summary, several limitations have been identified that may have affected the outcome and interpretation of the current research, as well as areas for attention in future research. These limitations include low sample size, self-selection of participants, disregarding aboriginal ancestry, use of measures that were not validated for use with forensic populations and that did not control for learning effects, as well as not incorporating a follow up period beyond treatment completion.

**Final Thoughts**

Traditional evaluation research has concentrated and continues to focus on the impact that rehabilitation programs has on recidivism. Extant research establishes that rehabilitation works. As suggested by Lipsey and Cullen (2007), the most important research task is to address the questions of when, why, how, and for whom rehabilitation works best. It is important to identify the pathways through which treatment has its effects. The current dissertation examined ECFs and motivation as a possible bridge between treatment and recidivism effects. This dissertation attempted to address both the why and how. The current dissertation attempted to overcome some of the common criticisms of forensic rehabilitation research, namely by including an appropriate comparison group, focusing on measurement selection, and assessing baseline and post treatment performance. The results revealed several noteworthy findings. Perhaps most
surprisingly, inmates did not reveal ECF deficits prior to engaging in rehabilitation programming. This, as well as several other findings, guided recommendations for future forensic researchers. With continued pursuit, better developed theories of change will help to explain the effects of current rehabilitation, and guide efforts to create improved interventions.
References


model: Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behavior, 38*, 735-755.


Assessment Resources, Inc.


Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J. (1999). Evaluating the use


associated with neurocognitive impairment in the initial phases of abstinence.


(Eds.), *Delinquency: Causes, Reduction and Prevention* (pp. 71-100). Nova Science Publishers.


Persistence of drug use during imprisonment: relationship of drug type, recency of
use and severity of dependence to use of heroin, cocaine and amphetamine in prison. *Addiction, 101.8*, 1125-1132.


van der Plas, E.A.A., Crone, E.A., van den Wildenberg, W.P.M., Tranel, D. & Bechara,


Endnote

Data Reduction Procedure

As a result of having 10 ECF dependent variables of interest, many of which included overlap in the description of what they measured, a variable reduction procedure was also investigated. Principal Component Analysis (PCA) often provides a relatively small number of components (made up of several variables) that account for much of the variance in a set of observed variables (Fabrigar, Wegener, MacCallum, & Srahan, 1999). PCA is appropriate when a researcher suspects that variables have shared variance. There was some theoretical support for completing a PCA with the current data; however, the results illustrated this procedure to be inappropriate for two reasons. First, sample size was a concern. Typically, 200 to 300 participants are the minimum used in appropriate and meaningful PCA (Tabachnick & Fidell, 2007; Comrey & Lee, 1992). In addition, the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO; Kaiser, 1970), which represents the ratio of the squared correlation between variables to the squared partial correlation between variables, indicated that PCA would not be appropriate. KMO values range from 0 to 1, with higher values indicating that PCA should yield distinct and reliable factors. Kaiser (1974) recommended accepting values greater than 0.5 as barely acceptable, but the KMO calculated from the current analysis was only 0.34. Thus, it was for these reasons, as well as the fact that repeated-measures data includes two data points which adds to the complexity of obtaining meaningful PCA outcomes, that PCA was not performed.
Appendix A

Office of Research Ethics
The University of Western Ontario
Room 5150 Support Services Building, London, ON, Canada N6A 3K7
Telephone: (519) 661-3036 Fax: (519) 850-2466 E-mail: ethics@uwo.ca
Website: www.uwo.ca/researcehtics

Use of Human Subjects - Ethics Approval Notice

Principal Investigator: Dr. P.N.S. Hoaken
Revision Number: 3
Review Number: 16280
Review Date: January 27, 2011
Protocol Title: Car Brief Correctional Rehabilitation Programming Improve Offender's Executive Cognitive Functioning
Department and Institution: Psychology, University of Western Ontario
Ethics Approval Date: January 27, 2011
Sponsor: SSHRC SOCIAL SCIENCE HUMANITIES RESEARCH COUNCIL
Documents Reviewed and Approved: Revised study and data
Documents Received for Information:

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICB Good Clinical Practices: Consolidated Guidelines, and the applicable laws and regulations of Ontario has reviewed and granted approval to the above-referenced revision(s) or amendment(s) on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time, you must request it using the UWO Updated Approval Request Form.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the HSREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of monitor, telephone number). Expected review of minor change(s) in ongoing studies will be considered. Subjects must receive a copy of the revised information/consent documentation.

Investigators must promptly report to the HSREB:

a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;

b) all adverse and unexpected experiences or events that are both serious and unexpected;

c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to this office for approval.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

Chair of HSREB: Dr. Joseph Gilbert
PDA Ref. #: 101B 44000 95335

Ethics Officer to Contact for Further Information

Janet Sutherland  Elizabeth Wambolt  Grace Kelly

This is an official document. Please retain the original in your files.
CURRICULUM VITAE
ERIN H. ROSS

Post-secondary Education and Degrees:
Doctorate, Clinical Psychology,
The University of Western Ontario, In progress – September 2007 to Present

Masters of Science, Clinical Psychology,
The University of Western Ontario, October, 2007

Bachelor of Arts, Honors, Psychology,
The University of Western Ontario, October, 2003

Awards, Scholarships and Research Grants
Graduate Thesis Research Award. The University of Western Ontario (2010-2011).

Finalist: Pillar Innovation Award, presented by Pillar Nonprofit Network, London, Ontario
for my involvement in starting and maintaining “Advocacy through Action” a student run
advocacy group within The University of Western Ontario (November, 2009).

Social Sciences and Humanities Research Council of Canada (SSHRC), Joseph-

Public Education Award. The Ontario Psychological Association (OPA), for my
involvement in starting “Advocacy through Action” within The University of Western
Ontario (2008).

Graduate Thesis Research Award. The University of Western Ontario (2008-2009).

Western Graduate Research Scholarship. The University of Western Ontario (2008-
2009).

Western Graduate Research Scholarship. The University of Western Ontario (2007-
2008).

Best Graduate Student Poster. Second Place. Canadian Psychological Association
(CPA) Convention, Criminal Justice Section, June, 2007, Ottawa, Ontario.

Graduate Thesis Research Award. The University of Western Ontario (2006-2007).

Masters Ontario Women’s Health Scholars Award. Ontario Council on Graduate Studies

Social Sciences and Humanities Research Council of Canada (SSHRC), Canada

Publications
Drug use and abuse and human aggressive behaviour. In J. Verster, K. Brady, E.
Strain, M. Galanter, & P. Conrod (Eds.), Drug Abuse and Addiction in Medical Illness:
Causes, Consequences, and Treatment. Springer/Humana Press USA.


Invited Lectures/Workshops


may Facilitate Improved Offender Programming. Theory review presented at the Canadian Psychological Association Annual Conference, Montreal, Ontario.


Research Experience
2008-2009

2004-2006
Senior Researcher, Peter N.S. Hoaken, Ph.D., C. Psych, The University of Western Ontario.

2001-2005
Research Assistant, Graham Reid, Ph. D., C. Psych, The University of Western Ontario.

Recent Clinical Experience
2010 - 2011  

2007 - 2010  

2006 - 2007  

**APA Approved Internship**  
June 2011 – July 2012  
Consultation Liaison in conjunction with the Psychiatry Department, and Transplant Psychology. Henry Ford Health Sciences, Detroit, Michigan.

**Teaching and Supervision Experience**  
2005 - 2010  
Teaching Assistant, The University of Western Ontario

Honors Thesis Supervisor, The University of Western Ontario