Assessing the Cognitive Contributors to Violence and Their Potential to Improve the Predictive Accuracy of Contemporary Risk Assessments

Sarah B Janes

College of Health in Social Science

A THESIS SUBMITTED FOR THE DEGREE OF

Doctor of Philosophy

February 2021

Acknowledgements

At a training for one of my first professional jobs I met a psychiatrist who asked each person why they like psychology, and what their interests are. When I told him about my interests and curiosity around true crime, serial killers, and understanding why people are violent, he looked me in the eyes and said, *you should really re-evaluate your interests, because that is disturbing*. Fast forward several years, and I am completing a PhD examining the neuropsychological risk factors for violence. Not only have I gotten to live and breathe this topic for the last few years, but I also had the opportunity to sit down with and talk to men who have committed violent offences, which has undoubtedly been the best part of this PhD. This group of men opened my eyes up to so much, especially the stigma that follows them, and the hardships that led many of them to where they are today. To the participants in this thesis, thank you so much for your time, this thesis would not have been possible without you. Though, I would not have been able to do this work without the support of my supervisors, Professor Matthias Schwannauer and Dr. Suzanne O'Rourke. Thank you for the sometimes-tough lessons that have made me into the autonomous, resourceful researcher I am today, and for cultivating my interest in this research area.

I also want to thank my mentor and friend, Jon. Thank you for taking my curiosity for neuropsychology, and making into a passion of mine. I have you to thank for all my clinical skills, and perhaps the moments of confidence that I get in infrequent, quick passing waves. You have consistently reminded me that I am smart and capable, and when I did not think I wanted to do this anymore, you encouraged me to remember why I started. Thank you so much for always being in my corner, I am beyond grateful.

This PhD also would have been much more difficult without the support of my friends from home[1], my best friends from Vermont[2], and the friends that I've made here[3]. Also, Sarah B., thank you so much for taking me under your wing in my first year and showing me all the ropes, you have always been and continue to be a huge support in my life, and I am so grateful for you. I will never tire of people confusing us, asking if we are sisters with the same name (?), and being known as "the other Sarah."

Of course, I could never forget Lindsey *Grillz*, who I met on our first day of our masters here in Edinburgh, where we bonded almost immediately over our love for all things

[1] Especially, Hayley & Vanna Gudog
[2] Specifically, Lindsey & Ashley 'GG' Rick
[3] Especially, Elise, Susannah, Claire Ann, Emma, and Mason G. McIntosh

related to forensic psychology. You have been one of my biggest assets during this PhD, you’ve supported me on an academic and personal level, and have repeatedly demonstrated how self-less you are, I am so lucky to call you my friend. Thank you for answering my questions without getting annoyed, listening to my presentations 600 times the night before a conference, being the familiar face in the crowd nodding at me, teaching me not to sweat the small stuff, sharing your dog with me, driving me everywhere, and just for being the friend and support that I needed to get through these last 6 years. I honestly do not know how I would have done it without you.

Saving some of the best for last, I must thank my family. To my Mom, Grandma, Grandpa, Kate, Aunt Jane, and the rest of my family- Thank you for unconditionally loving and supporting me through everything. Especially my Mom who reminds me every single day that if I need something, she can be on a plane in a matter of hours and has proven it several times. I am so lucky to have such a strong female role model in my life, who has not only taught me the importance of getting an education and working hard for what you want, but who has consistently led by example. You have given me the tools to succeed repeatedly, and I am so, so grateful.

Finally (and most importantly), my husband Joshua. Josh, you have been on this journey with me since undergrad, you have moved your entire life for me more than once and you rarely complain. You take every opportunity to turn my negatives into positives, and constantly remind me that my worth is so much more than this degree. Thank you for your love and support, and for putting up with my pranks, my random bouts of loud singing, annoying early morning dancing, and for sympathy laughing at my jokes. You have never let me even entertain the idea of giving up, you have listened to endless practice presentations, plenty of crying, and a ton of anxiety. I am beyond grateful to have you by my side and have loved working from home with you during this pandemic.

Finally, finally, I would like to thank all the NHS and social work sites who allowed me to conduct research in their services, and the psychologists and social workers who went out of their way to help me, especially, Dr. Sarah Harper, Iain Hunter, and Val Lawrie.

Nothing is easier than to denounce the evildoer; nothing is more difficult than to understand him.

Fyodor Dostoyevsky

Abstract

Research on risk assessments has made considerable progress in recent years, with the development of long- and short-term predictive tools, as well as those which are specific to different types of offending including violent, sexual, and general offending. However, while there are many assessments, they differ little in their content and, as a result of the repetitive use of certain risk factors, researchers have suggested that contemporary risk assessments have hit a glass ceiling, as many have only moderate predictive accuracy. Risk assessments are further limited by the use of self-report, collateral information, and file reviews to assess clinical and risk-related factors, rather than validated performance measures, hindering their predictive accuracy. In parallel with these findings, converging evidence from neuropsychology and neurobiology has underlined the brain regions associated with violent behaviour and subsequent research has further demonstrated this through observational studies. The link between neuropsychological functioning and violence risk raises the potential for behaviourally measured neuropsychological risk factors to improve the predictive accuracy of existing risk assessments as well as aiding in the identification and management of these risk factors. The overarching aim of this thesis was to develop a risk assessment tool, for use alongside existing risk assessments, to increase their predictive accuracy, and to aid clinicians in the assessment and management of neuropsychological risk factors for violence. The development of the proposed tool followed the example of the National Institute of Health's development of the MATRICS Consensus Cognitive Battery (MCCB) which sought to identify not only cognitive domains identified by the research literature, but also those that, in a relatively new field, may yet to be examined empirically, through expert consensus.

Following the introductory and background chapters which provide context on the aetiology of cognitive impairments in violent populations, and a critical evaluation of the methodological limitations in violence research and violence risk assessments, are a cohesive set of studies which build upon each other to lay the foundation for the development of a structured risk assessment tool.

The first step in this process was to critically examine the state of the literature and to identify neuropsychological measures which differentiate violent from non-violent offenders, and then to examine their association with violent outcomes in predictive and correlational studies. A systematic review and meta-analysis was undertaken to quantitatively summarise the association between neuropsychological measures and violent outcomes. After acknowledging that existing reviews in this area have largely focused on executive functions

and specific diagnostic groups, the review adopted a broader approach, first examining factors which differentiate violent from non-violent offenders (part 1), followed by separately analysing the neuropsychological correlates of violence (part 2). Forty-two studies were included in the analyses, and 12 individual neuropsychological domains were examined in part 1, and five in part 2. The findings from this study revealed a large range of effect sizes with wide confidence intervals, highlighting significant heterogeneity due to methodological differences between studies, making it difficult to draw definitive conclusions. However, measures of impulsivity, inattention, and lack of insight boasted significant correlations with prospectively measured violent outcomes, revealing their potential to add a small amount of incremental validity to existing risk assessments.

In parallel to the review and meta-analysis, a Delphi study was conducted to achieve an expert consensus on the neuropsychological risk factors thought to be essential for inclusion in a violence risk assessment. The Delphi method consisted of three rounds, in which an international sample of researchers and clinicians participated. A consensus was achieved on a list of 10 cognitive domains, ranked in order of importance, including, poor inhibitory control, behavioural impulsivity, risk taking, affective empathy, response monitoring, affect recognition, cognitive insight, reasoning, cognitive empathy, and clinical insight.

Taking the results from the meta-analysis and the Delphi study, a neuropsychological battery was developed consisting of measures of inattention, response inhibition, reasoning, response monitoring, risk taking, affect recognition, cognitive empathy, and clinical insight, as well as a measure of intelligence. A prospective pilot study was designed to test the feasibility of implementing a neuropsychological battery into routine practice to aid in the assessment and management of risk, and the battery was piloted in n= 32 mentally ill offenders from high, medium, and low secure settings within NHS Scotland, and n=31 violent offenders who are living in the community and are service users of Criminal Justice Social Work in Edinburgh, Scotland. The findings of this study evidenced the pervasive cognitive impairments seen in these populations with a history of violence, regardless of their setting, however, the potential to improve the predictive accuracy of existing risk assessments appeared limited.

The final study was a pilot case-series which sought to re-consent inpatients from the high secure setting to examine neuropsychological measures which explain the variance in short-term risk of aggression. The study was designed with a focus on how the major limitations seen in the violence literature may be addressed, such as small samples and low base rates of observational outcomes. Although an effort was made to tackle these limitations, a low base rate of recorded incidents persisted. However, the study demonstrated how a repeated measures outcome can help to minimally increase power in the presence of small samples, and the use of a novel statistical analysis demonstrated how inferences can be drawn on continuous outcomes even when there is a low base rate of incidents, increasing the specificity of results.

Taken together, the findings of this thesis demonstrate that neuropsychological measures have limited potential to improve the predictive accuracy of risk assessments, however, they bring focus to the often-profound cognitive impairments seen in violent offenders and the necessity of assessing for cognitive impairment in parallel with the assessment of risk. The robust assessment of cognitive strengths and weaknesses has the potential to enhance the formulation of risk and inform its management and treatment, optimising treatment outcomes and the management of violence risk.

Lay Summary

Violence risk assessments are tools used by professionals to assess the level of danger an individual poses to themselves or others. The tools include a list of characteristics and behaviours that are related to violence risk, also known as risk factors. Using past and present information, professionals identify which risk factors the individual has, and then formulates an opinion on how to best treat and manage them. However, one of the problems with violence risk assessments is that they do not always accurately identify individuals who will act violently in the future, suggesting that they could be improved. Research has also found that forensic populations, or individuals who commit crimes, are more likely than the general population to have a history of head injuries, substance misuse, certain diagnoses, and childhood trauma, all of which can cause damage to areas of the brain. When certain areas of the brain become damaged, the abilities (known as cognitive abilities) they hold can become impaired (known as cognitive impairments). Cognitive impairments relate to violence risk as they can increase impulsivity, carelessness, and socially inappropriate behaviours, and can reduce self-control, attention, planning abilities, and goal setting making it more difficult to desist from offending, and to behave in a pro-social manner. In addition to their relationship with violence risk, cognitive impairments may also inhibit an individual from fully benefitting from treatment programmes, understanding simple instructions, and from carrying out daily tasks. Though, despite the relationship with violence risk and poor treatment outcomes, most cognitive abilities are not considered during risk assessments.

This research assessed the cognitive abilities of two samples of violent offenders (e.g., individuals who have committed a violent offence), and then investigated their relationship with violence and their ability to improve the accuracy of risk assessments, focusing on a list of cognitive abilities that were identified in the literature, and by experts in this field. Our findings indicate that most cognitive abilities were not related to violence in these samples, and only a few demonstrated the ability to improve the accuracy of risk assessment tools. However, several participants demonstrated poor performances on the measures of cognitive abilities suggesting that many of them may be functioning at a lower level than normal. The conclusions of this research are that while many cognitive risk factors may not meaningfully improve the accuracy of risk assessments, cognitive abilities should still be assessed as part of a risk assessment to identify the cognitive strengths and weaknesses of offenders to inform individualized treatment plans and support for rehabilitation.

Publications Resulting From this Thesis

One publication resulted from the work within this thesis:

Janes, S., O'Rourke, S., & Schwannauer, M. (2021). Assessing the cognitive contributors to violence: A pilot and feasibility study protocol. *Social Science Protocols*, 1–16. https://doi.org/10.7565/ssp.2021.v4.5213

A copy of the publication is included in Appendix A, reproduced with the permission of the publisher. In addition to the above paper, it is anticipated that the research in this thesis will support a further four publications, and work is ongoing to progress these submissions.

1. **Janes, S.**, McIntosh, L.G., O'Rourke, S., & Schwannauer, M. (*in preparation*). Examining the cognitive contributors to violence in forensic populations: A systematic review and meta-analysis.
2. **Janes, S.**, O'Rourke, S., & Schwannauer, M. (*in preparation*). Achieving an expert consensus on the cognitive domains considered relevant for inclusion in structured violence risk assessments: An international Delphi study.
3. **Janes, S.**, O'Rourke, S., & Schwannauer, M. (*planned*). Assessing the cognitive contributors to violence: A pilot and feasibility study.
4. **Janes, S.**, O'Rourke, S., & Schwannauer, M. (*planned*). Cognitive contributors to the variance in behaviours associated with imminent situational aggression: A pilot case-series.

Table of Contents

List of Tables

List of Figures

1 General Introduction

Chapter Preface

This introductory chapter contextualises the aims of this thesis among some of the long-standing limitations observed in risk assessments. Improving the accuracy of violence risk assessments remains a key priority for future research, as well as the treatment and management of violent offenders. This chapter gives a brief overview of the evolution of risk assessments, and highlights the moderate predictive accuracy seen across risk measures and how they have essentially reached a 'glass ceiling'. This is followed by an introduction to the neuropsychology of violence, its relevance for violence risk assessments and the treatment and management of offenders. Finally, following the rationale and central thesis aims, the structure of the thesis is laid out, outlining a cohesive set of four empirical studies undertaken, with an explanation provided on how each study builds upon the previous one.

1.1 Introduction

The development and validation of violence risk assessments has made considerable progress in recent years transforming from 'prediction' only, into 'prediction and management' tools, with more focus on the individual needs of offenders. In parallel, converging findings from neuropsychological and neurobiological research have identified brain regions associated with violent behaviour and have highlighted a relationship between cognitive impairments and violence, although findings have been inhibited by methodological limitations and poorly defined outcomes. Whilst there are violence risk assessments that encompass neuropsychologically informed risk factors, such as impulsivity and lack of insight, they are often assessed using collateral information, rather than validated performance measures. Moreover, meta-analytic findings have revealed that widely used violence risk assessments have low predictive validities, indicating that, aside from the specific context in which the assessment will be used, there is no risk measure significantly *better* than another (M. A. Campbell et al., 2007; Desmarais et al., 2016; Yang et al., 2010). Further, Monahan & Skeem (2014) postulated that due to existing risk assessments using essentially the same factors and relying on self-report measures, only differing on how the factors are analysed, risk assessments have reached a 'glass ceiling' (Skeem & Monahan, 2014). In parallel, researchers have suggested moving away from only using interviews, self-report measures, or record-based assessments to enhance predictive accuracy (Fazel et al., 2012). Thus, with consideration of the relationship between cognitive deficits and violence risk, it was hypothesised that the addition of cognitive abilities to existing violence risk assessments, measured with validated neuropsychological tools, may have the ability to break the ‘glass ceiling’ and improve predictive accuracy, while subsequently identifying cognitive strengths and weaknesses of the individuals, and informing rehabilitation needs (Haarsma et al., 2020).

1.2 Brief Background on the Evolution of Risk Assessments

Before the prediction of violence was well researched, the judicial system relied on sentencing, and the notion of *incapacitating* offenders from committing any further offences with hopes of improvements through treatment, spontaneous recovery, or age-related complications (Quinsey et al., 1998). Sentencing was meant to serve several purposes, ranging from punishment and general deterrence to rehabilitation, however, it is still often debated whether sentencing practices accomplish these goals. Historically, a focus of sentencing was also to examine the likelihood that an offender was going to commit a new offence, though, this was only a useful practice for high-risk offenders, and it was viewed as a waste of resources to incapacitate, treat, and rehabilitate offenders who essentially had little to no risk of re-offending. Thus, it was posited that the measurement and prediction of risk in offenders would help to categorize those who were at a high and low risk of re-offending.

Historically, researchers looked at the *prediction of dangerousness* primarily in mentally ill offenders (Quinsey et al., 1998). Though, *dangerousness* was an ambiguous term which led to more questions than answers, and significantly impeded research efforts. Through the 1960’s to the early 1980’s, several attempts were made through follow-up studies to predict violent recidivism (dangerousness), but with little success, which led some professionals to believe that *dangerousness* could not be predicted. As it came to light that predicting future *dangerousness* or *violence* required three main components: *risk factors* (factors which are empirically associated with the likelihood that violence will occur), *harm* (the type, severity, and amount of violence being predicted), and *risk level* (the likelihood that violence will actually occur) (National Research Council, 1989 as cited in Heilbrun, 2009), the term *risk assessment* was quickly used in place of the term *dangerousness* to better and more specifically describe violence prediction (Monahan & Steadman, 1994). Although this was an improvement, several limitations remained which were generally found to stem from the use of unreliable methods, overreliance on historical (static) predictor variables, and the inability to identify the antecedents to recidivism. The main source of information for these predictions was the use of case files and criminal records, giving ‘predictors of convenience’ which were also shown to be unreliable (Quinsey et al., 1998).

As an alternative to traditional approaches of assessing risk, in 1990, Andrews et al, formalised the *Risk-Needs-Responsivity* (RNR) approach which added a new component to risk assessment, turning the focus from simply predicting risk, to managing and treating it. Within this approach, *Risk* refers to the idea that individuals who are most likely to engage in future antisocial behaviour, should be receiving the most intensive treatment and management services; *Needs* (also called *criminogenic needs*) refers to deficits or impairments the individual has which may be related to a violent outcome or risk of reoffending, and *Responsivity* refers to the probability of the individual responding to treatments designed to reduce risk of reoffending (Andrews et al., 1990). Following this development, a distinction was made in the field between *prediction of risk* and *risk management*, where static risk factors (factors that cannot be changed through intervention) were used primarily in risk prediction, and dynamic (factors that have the potential to change through intervention) were primarily used in risk appraisal and risk reduction. The distinction between *prediction only* and *both prediction and management* tools gave a clearer picture to those conducting the assessments and allowed a more informed selection of which risk assessment tools to use (Heilbrun, 2009).

1.3 Approaches to Violence Risk Assessment

1.3.1 Unstructured Professional Judgement and Actuarial Approaches

The earliest risk assessments utilised unstructured professional judgement, which involved decisions and evaluations based on a person’s presentation but were often influenced by confounding variables. In addition to this method being unreliable (e.g., often predicting no better than chance), the risk factors that were assessed varied depending on setting, time, and the patient (Ægisdóttir et al., 2006). Following Monahan’s work in 1981, where he concluded that psychiatrists and psychologists can predict risk no better than chance, he pinpointed a “second generation” (Monahan, 1984 p. 141) of risk assessments, and with the development and advancement of statistical methods, the research focus shifted to actuarial methods of assessing risk (Singh, 2012). Actuarial risk assessments include specific risk factors that are statistically associated with violence (Monahan, 2000). Examples of actuarial tools are the Violence Risk Appraisal Guide (VRAG; G. T. Harris et al., 1993), and the Sexual Offender Risk Appraisal Guide (SORAG; Rice & Harris, 2016).

Compared to the unstructured professional judgement method, actuarial tools are far superior (Buchanan, 2008), however they do have limitations. Aside from their proven usefulness and predictive accuracy, actuarial tools have been criticized for looking too narrowly at offenders. Hart (1998) stated that this approach tends to only focus on specific factors rather than the whole picture, potentially leaving out vital details that are case specific, and for excluding risk factors that have not been proven empirically, even when they appear to be an important risk factor to explore. Further, Dolan and Doyle (2000) emphasized that actuarial methods are often not useful in clinical settings, arguing that static risk factors do not inform management or the level of intervention needed, and they are not sensitive to individual change, though they can be useful for group change. While actuarial tools have proven to be useful for research purposes, they were not found to be advantageous in clinical settings, which left a need for a more useful tool for clinicians. Thus, the third generation was marked by the development of risk/need/responsivity-based structured instruments that followed the RNR approach to assessing risk and assessed dynamic risk factors. One of the most notable examples of an RNR instrument is the Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995). Whilst the LSI-R is actuarial in nature, it includes dynamic risk factors, and is recommended to be completed every 12-months, or subsequent to major changes in the offender's circumstances, to capture changes in *needs* that will increase or decrease the offender's risk of general re-offending (Garrington & Boer, 2020).

1.3.2 Structured Professional Judgement Tools

In the 2000's, the field began to shift their focus from risk assessment to risk management, and with that, the idea of reintroducing clinical judgement back into risk assessments resurfaced (Singh, 2012). It was thought that reintroducing clinical judgement back into risk assessments could assist professionals in making more informed decisions (Douglas et al., 2003). With this, Structured Professional Judgement (SPJ) tools, instruments that use clinical judgement to supplement actuarial scales, gained popularity across the field, and marked the fourth generation of risk assessments (Ogloff & Davis, 2020). SPJ tools derive domains from the literature and experience which provide the structure to frame the clinician's thinking. They provide guidelines for assessing risk, ensuring that clinicians consider all relevant domains, and that they have all the pertinent information to judge their relevance. Moreover, they can be viewed as a comprehensive assessment, as they not only provide a predictive element, but they also give a more complete picture of individuals with a focus on understanding their violence risk and how to best manage it (Ogloff & Davis, 2020). An example of one of the most popular SPJ tools is the Historical, Clinical, Risk Management-20 version 3 (HCR-20v3; Douglas et al., 2013). The HCR-20v3 consists of ten historical, five clinical, and five risk management items, and considers empirically validated risk factors, clinical factors that are shown to play a role in violence risk, and items related to the treatment and management of risk, all of which act as a guide for a clinician to make a clinical judgement. However, there is still some debate around the validity of professional judgement. Advocates for actuarial tools argue that adding *clinical judgement* into risk assessment invalidates this method of violence prediction. However, the development of SPJ tools addresses the criticisms that actuarial tools do not take individual variations into account (Sreenivasan et al., 2000). Despite these differences, there is unanimous agreement that actuarial and SPJ tools perform equally well in predicting violence, though they serve different purposes (Singh et al., 2016).

Following the introduction of SPJ tools, and with the rise of the positive psychology movement, instruments which focused only on identifying and managing risk factors were criticised for neglecting the notion that positive actions can reduce the overall risk of future offending, as well as the potential stigmatizing effect they have as a result of focusing only on negative factors (Garrington & Boer, 2020). Shortly after, SPJ tools which included protective factors were developed, such as the Structured Assessment of Protective Factors (SAPROF; De Vogel et al., 2011), a measure of exclusively protective factors, and the Short-Term Assessment of Risk and Treatability (START; Webster et al., 2006), a measure of both risk and protective factors. Criticism of these measures have surrounded the definition of *protective,* and researchers have argued that more appropriate descriptions may be *strengths* and *capabilities*, as even if a protective factor is present, it is only *protective* if the offender demonstrates a willingness to use it (Klepfisz et al., 2017).

1.4 Predictive Ability and Accuracy of Risk Assessments

Over the last few decades, researchers have worked assiduously to improve risk assessments by identifying empirically relevant risk factors for criminality and violence that both accumulatively and accurately predict future behaviour. As a result of these efforts, in 2014, there were more than 400 risk assessments available (Singh et al., 2014), a number that has likely grown in the last six years, and although these instruments cover various types of risk, include slightly different factors, are validated in specific settings or for specific follow-up times, most of them share the same goal or objective, which is to prevent violence (Guy et al., 2015). As there are numerous primary studies looking at the predictive ability of either individual risk assessments or in comparison to other risk assessments, several valuable reviews have helpfully synthesized these findings in addition to shining light on the inconsistencies in findings between studies, as well as the methodological limitations that have overwhelmed the field, of which have consequently contributed to variability in findings and have hindered our ability to increase the accuracy of predictions.

Steadman and Monahan (1996) said it well, when they stated that, "predicting the weather is easy compared with predicting violence" (p. 932), and this is evidenced in the research effort that has gone into examining and attempting to improve the predictive accuracy of various risk instruments. For example, SPJ tools are widely used in forensic psychiatric settings, and whilst their improvement is a priority in this field of research, their predictive accuracy has rarely been found to be more than *modest* (Haque & Webster, 2013). One of the most widely used methods for examining predictive accuracy in the literature is Area Under the Curve (AUC). An AUC score can span from 0 to 1, where .5 would be no better than chance that a violent offender would score high on a risk assessment over a non-violent offender, and 1 being near perfect accuracy. In a widely cited meta-analysis conducted by Singh and colleagues (2011), the predictive validity of nine risk assessments across 68 studies was examined using AUC. In this review, only four risk assessments had a median AUC score greater than .70, which is considered a moderate effect size, namely, the Sexual Violence Risk Assessment- 20 (Boer et al., 1997) with an AUC score of .78, the Sexual Offender Risk Appraisal Guide with an AUC of .75 (G. T. Harris et al., 2003; Rice et al., 2006), VRAG (G. T. Harris et al., 1993; Quinsey et al., 2000) with an AUC of .74, and the Structured Assessment of Violence Risk in Youth (Borum et al., 2006) with an AUC of .71. The remaining AUC sizes ranged from .70 to .66, with the HCR-20, a frequently used risk assessment, being only the fifth best predictor of violence with a pooled effect size of .70 (Sing et al., 2011; Hague & Webster, 2013). Although an AUC score of .75 is interpreted as a moderate to large effect size indicating decent predictive accuracy for two of the measures, these findings suggest that there is room for improvement. In fact, in several studies comparing risk instruments, an AUC value of only 0.70 was consistently achieved (Belfrage et al., 2000; Coid et al., 2009; De Vogel et al., 2004; Douglas et al., 2005; Douglas, Ogloff, et al., 1999; M. Doyle et al., 2002; M. Doyle & Dolan, 2006; Glover et al., 2002; Grann et al., 2000; Gray et al., 2003; Grevatt et al., 2004; Kroner & Loza, 2001; Kroner & Mills, 2001; Mills et al., 2007; Morrissey et al., 2007; Nicholls et al., 2004; Snowden et al., 2007; Stadtland et al., 2005; Tengström, 2001; Warren et al., 2005, as cited in Coid et al., 2011), which Coid et al. (2011) suggested is due to the *glass-ceiling* effect.

To explore this perception further, in 2011, Coid and colleagues, conducted a prospective cohort study on n=1,353 male prisoners released from prison. As the majority of studies examining predictive accuracy only look at composite scores on measures, the study aimed to examine individual items from the HCR-20, the Psychopathy Checklist-Revised (PCL-R) and the VRAG for violent and criminal recidivism, and then to investigate the possibility for independent items from the tools that were shown to be predictive, to increase predictive accuracy (Coid et al., 2011). Results for the three instruments demonstrated that they each significantly predicted violent recidivism; however, their AUC values did not exceed a moderate level of predictive accuracy, with the highest value being 0.70 for the VRAG.

When investigating whether removing items from the scales that did not significantly predict the outcome would improve predictive ability, the authors found that aside from the PCL-R, which demonstrated a small improvement in predictive accuracy, only small, and non-significant gains were found for the remaining instruments. After looking at the sub-scales independently, the authors combined them to create a *super instrument* to measure the overall predictive accuracy, which resulted in a similar AUC score of only 0.72. Based on these results, the authors suggested that there may be a *glass ceiling* effect on measures which incorporate static risk factors, and the ability to increase their predictive ability may not be possible (Coid et al., 2011), though they offered other explanations for this, such as various protective factors gained during the follow-up period. Further, the authors posed the question of whether non-predictive items should be removed from these risk instruments, and ultimately concluded that due to the variability in populations and outcomes, what demonstrates predictive ability in one sample, may not in another, especially given their sample was a prison population only. Evidence of this more individualistic examination of risk assessments has been supported by meta-analytic findings where the moderation effects of outcomes were examined in relation to risk assessments, though the results have not been completely consistent.

In a 2007 meta-analysis, Campbell and colleagues found that in papers which operationalised recidivism as *institutional violence* rather than *violent recidivism* found more accurate predictions using the PCL and PCL-R (M. A. Campbell et al., 2007). A 2008 meta-analysis examining the predictive validity of risk assessment tools for juveniles, found that significantly larger effect sizes were produced in studies which used *new arrests* or *new referrals* as their operationalisation of recidivism compared to papers which used *new adjudication* as their operationalisation (Schwalbe, 2008). In 2012, Fazel and colleagues compared various risk assessment instruments, and found differences in predictive accuracy depending on the type of outcome it was meant to predict (e.g., violence, sexual, or any criminal offending). Their findings demonstrated that *violence risk assessments* performed the best, and instruments predicting violent and sexual offending produced high sensitivities and negative predictive values, both outperforming instruments for general offending, further indicating that specificity in the outcome is necessary for increased predictive accuracy. However, instruments specifically designed for predicting violent outcomes in this review still had a median AUC score of 0.72 (Fazel et al., 2012), and similarly, two additional meta-analyses (Leistico et al., 2008; Walters, 2003) found no moderation effects between using *institutional infractions* versus *recidivism* to operationalise outcomes.

Whilst there is value in acknowledging these criticisms, it is beyond the scope of this thesis to attempt to *fix* existing risk assessments or to create a brand-new measure, but it is instead an aim to examine one option that may improve their predictive accuracy. As numerous studies have examined the association between cognitive functions and violence risk in both forensic and civil populations and have found somewhat consistent results to encourage further investigation, a portion of recent literature has begun to examine this exclusively. However, before examining the relevance of neuropsychology to risk assessments, it is first necessary to understand the neurocognitive mechanisms which are thought to underly violent behaviour.

1.5 The Neuropsychology of Violence

Whilst there are numerous biological theories of violence and aggression suggesting neurological functioning as a cause for these behaviours, in 1992, Jones developed a theory of violence implicating neuropsychological substrates in an attempt to conceptualise the link between neuropathology and violence (Jones, 1992 as cited in Golden et al., 1996). Jones' theory was broken into four pathways; (1) an increased activation of the nervous system; (2) a decreased inhibitory ability; (3) deficits in attention, memory, higher mental processes, and concentration; (4) misinterpreting external stimuli and events (Jones, 1992 as cited in Golden et al., 1996). Golden et al. (1996), related this theory to the brain regions where these pathways would likely occur after they have been damaged or impaired, including the frontal and prefrontal areas and their interactions with subcortical areas of the brain, temporal lobes and their interaction with limbic structures, and generalised damage to the entire brain. Although most of these brain regions and functions overlap with one another, it is important to understand how each structure, when injured, can contribute to violence risk. The prefrontal cortex (PFC), for example, plays a vital part in cognitive functioning and pro-social behaviour, and has an important role in executive functions (Dolan, 2012; Raine, 2002; Sapolsky, 2004; Teichner & Golden, 2000; Van Der Gronde et al., 2014). Executive functions, one of the most crucial functions of the human brain, includes attention, planning, working memory, self-awareness, flexibility in behaviours, abstract thinking and decision making (Van Der Gronde et al., 2014). Structural and functional abnormalities of the PFC have been found through PET scans in offender populations (Raine, 2002; Stein, 2000; L. C. Wilson & Scarpa, 2012), EEGs in violent populations (Brower & Price, 2001), and MRIs in antisocial patients (Brower & Price, 2001; Hoptman, 2003). Similarly, damage to the orbitofrontal cortex (OFC) can result in behaviour that is disinhibited and impulsive (Raine & Yang, 2006), suggesting that patients with a lesion in this area of the brain often show reactive aggression when feeling frustrated or threatened. Two additional frontal areas that have been identified as contributors to violence risk are the dorsolateral prefrontal cortex (DLPFC) and the ventromedial prefrontal cortex (VMPFC). When injured, the DLPFC, whose main function is in working memory (Pochon et al., 2001), can lead to impairments in planning, attention shifting, decision-making and perseveration (Raine & Yang, 2006), whereas damage to the VMPFC, which stores knowledge of social norms and rules, may make this information less accessible, often resulting in inappropriate social behaviour, such as physical threats and intimidation in situations where it is provoked or unprovoked (Grafman et al., 1996). The VMPFC has also been associated with poor control of reactive violence (Van Der Gronde et al., 2014).

A damaged amygdala, whose main functions involve emotional responses and emotional memories (Mitchell & Beech, 2011), may present an individual with difficulties in recognising distress or fear cues in others, which can also result in impaired conditioned fear responses (Pardini et al., 2014). Likewise, functional imaging studies have evidenced abnormal functioning of the hippocampus in violent offenders (Hoptman, 2003; Liu & Wuerker, 2005; Raine, 2002), which plays an important role in memory (Van Der Gronde et al., 2014), and is a part of the limbic system, playing a large part in fear conditioning and impulsive behaviour (Cardinal, 2006; van Goozen & Fairchild, 2006). Brain imaging studies have also shown that psychopaths have reduced hippocampal volume (Raine et al., 2004), as well as violent offenders with antisocial personality disorder (Laakso et al., 2002), and violent psychiatric patients (Kumari et al., 2006). Furthermore, the temporal lobes, which contain the hippocampus, and have an important role in forming long-term memories (Van Der Gronde et al., 2014), have been linked to violent behaviour. For example, a 20% volume reduction was seen in patients with aggressive psychopathy (Dolan & Park, 2002), and significant volume loss was observed in the right and left superior temporal gyri in psychopathy (Müller et al., 2008). Finally, studies examining the anterior cingulate cortex (ACC), a limbic region, involved in response selection, regulation of behaviour, response inhibition and empathy (Van Der Gronde et al., 2014), have shown that there is less activation in the ACC in violent populations during various memory and fear-conditioning tasks, and tasks which involve viewing negative images (Raine & Yang, 2006).

1.6 Relevance of Neuropsychological Assessment in Forensic Populations

Neuropsychological assessment can be conceptualised as a method of examining the brain through the study of behaviour by means of standardised measures that provide indices of brain behaviour relationships (Heilbrun et al., 2003). Forensic neuropsychological assessments examine disabilities, capacity to make decisions, competency issues related to civil and criminal cases, and criminal responsibilities, all of which demand a high level of expertise (Larrabee, 2018). Neuropsychological assessment in forensic populations is highly relevant, as individuals involved in the justice system often have higher rates of learning difficulties (She & Stapleton, 2006), ADHD (Young et al., 2015), TBI (Shiroma et al., 2010), trauma (Wolff & Shi, 2012), and problems related to substance abuse (e.g., Fazel et al., 2006; Fishbein, 2000) when compared to normative populations (LaDuke et al., 2017). However, due to the required training and administration time needed for neuropsychological assessments, their use is often rare in institutional settings (LaDuke, 2015; Marceau et al., 2008; Vanderhoff et al., 2011), where they are arguably needed most, and are further constrained by cultural and clinical factors, education, and response bias or malingering (Howieson, 2019). Furthermore, although most criminal and violent populations are outperformed on neuropsychological measures by normative samples, normative data may not be an appropriate comparison to draw conclusions on in these populations. As individuals involved in the justice system often have learning difficulties, TBI, and trauma, for example, they may represent a neuropsychological population that is complex and distinct and is poorly described by current normative data (LaDuke et al., 2017), perhaps leading to descriptions of strengths and weaknesses which are inaccurate, poor conceptualizations of impairments, and misdiagnoses (Cagigas & Manly, 2014). Thus, the use of relevant control groups and norms based on justice-involved populations are warranted in research to further the neuropsychological understanding of these populations. For example, a 2018 study, and the first of its kind, compared executive functions of a sample of forensic psychiatric patients (n= 42) to a sample of prisoners (n= 77), and found that the while the forensic psychiatric patients performed more poorly on executive function measures compared to prisoners, a large proportion of each sample (forensic patients: 9.5-35.7%; prisoners: 5.2-27.3%) displayed clinically significant deficits on all measures compared to what would be expected from a normal population (2.5%) (Shumlich et al., 2019). Although the main focus of this study was not violence, it highlights the degree of the impairments often seen in offending populations and demonstrates how comparing different offending groups to each other may help to demarcate the impairments that are specific to each group, rather than the magnitude of impairments when compared to normal controls. However, there remains a role for comparison to typical populations, though, the typical population and the offending population should have similar demographics.

Neuropsychological measures also hold significant relevance for the treatment benefits gained by justice-involved populations. Violent behaviour is often accompanied by cognitive impairments, many of which indirectly relate to violence risk, often impacting treatment responsivity, such as theory of mind, verbal learning, memory, and intelligence (O'Rourke, 2013). Social cognition, which includes theory of mind, serves as a guide for certain behaviours and plays a role in memory, decision-making, attention, motivation, and emotions (Adolphs, 2001), and is considered to be the foundation of the ability to have healthy emotional regulation and social development in relationships (Oatley, 2004), thus, when social cognition is impaired, these functions become less available to the individual. Likewise, impaired verbal learning and memory may make it more difficult for individuals to process verbal information and remember it. Level of intelligence is also a major factor in an offenders' ability to take part in treatment (Andrews et al., 2011), which may partially explain the association between low IQ and criminal and violent outcomes. For example, Moffitt (1993), suggested that the association between low IQ and violent behaviour may be conflated by neuropsychological deficits elsewhere that are often seen in these populations, such as in language processing, impaired auditory memory, and poor social judgement, similar to those identified by O'Rourke (2013) above. Further, more recently, high intelligence has been established as a protective factor against violence risk, as it can help to override problems often seen in aggressive and violent individuals, such as compensating for having a disadvantaged background in education and job attainment, overcoming social processing biases related to aggression, and having intact executive functions (Ttofi et al., 2016).

Given the presented evidence, it can be assumed that offenders commonly present with cognitive weaknesses which may hinder their ability to process, understand, and recall verbal information. This notion is highly relevant to treatment responsivity and outcomes as treatments for forensic populations are often delivered orally and depend on the participants to have the ability to learn the information and to recall it later (O'Rourke, 2013). Thus, neuropsychological assessment should be considered an essential component of offender treatment programmes, as without knowledge of their cognitive strengths and weaknesses, individuals (who are considered to be heterogeneous) are expected to benefit from *one-size-fits-all* treatments, and as a result, the treatments may be viewed as an inefficient use of time and resources. Moreover, the identification of cognitive strengths and weaknesses can aid in the development of treatment plans that can be individually tailored to the needs of offenders to ensure they receive the necessary support, and their strengths can be leveraged to optimise treatments, a notion which is in line with the RNR model where responsivity encourages the use of cognitive-behavioural treatments, using an approach which is tailored to the offender's ability and learning style (Bonta & Andrews, 2017). Researchers have cautioned that failing to efficiently identify impairments in executive functioning prior to intervention may inhibit treatment progress (Cheng et al., 2019). Further, cognitive impairments can affect an individual's ability to make informed decisions (Palmer et al., 2012), including informed consent (Marons, 2001), as well as their fitness to stand trial, as cognitive abilities play a crucial role in a defendant's ability to understand, comprehend, and assist counsel (Douds & Haut, 2015). Furthermore, the inclusion of this population in treatment evaluations, without prior knowledge of their cognitive abilities, may contribute to invalid and unreliable results regarding treatment effectiveness.

In addition to the utility of neuropsychological assessments for identifying neurocognitive strengths and weaknesses in forensic populations, numerous neuropsychological measures have significantly differentiated violent and non-violent populations and have shown predictive utility in studies of recidivism and inpatient violence, however, the research surrounding the usefulness of cognitive risk factors in the assessment of risk is underrepresented in the literature. Despite this, researchers have begun suggesting the incorporation of cognitive variables in violence risk assessments (e.g., Brugman et al., 2016; Caggiano, 2000; Hancock et al., 2010), while others have developed mobile neurocognitive assessment software to predict reoffending among correctional community probation populations (Haarsma et al., 2020), indicating a belief that neuropsychological measures may also be relevant to violence risk.

1.7 Neuropsychological Assessments with Relevance to Violence Risk

Numerous studies have examined the use of neuropsychological assessments to either differentiate violent from non-violent populations or to predict violent and aggressive behaviours, thus, to best describe these findings, meta-analytic reviews will first be discussed. In 2000, Morgan and Lilienfeld conducted a meta-analytic review investigating the association between antisocial behavior and neuropsychological measures of executive functions, by examining the differences between effect sizes in groups with antisocial behavior compared to non-antisocial control groups. Using both measures of executive functioning and non-executive functioning tasks, they located 39 studies, yielding n= 4,589 participants, and found an average mean weighted effect size of d= 0.62, indicating 0.62 standard deviations between the two groups, however the results were significantly heterogeneous due to methodological differences between studies. Individual neuropsychological measures of executive functions all resulted in significant weighted effect sizes between groups, such as Category Tests (d= 0.37), Mazes (d= 0.74), Stroop Test (d= 0.43), Trails B (d= 0.33), WCST-Perseverative Errors (*d*= 0.24), and Word Fluency (*d*= 0.33). Measures that were operationalised as non-executive functioning measures also resulted in significant weight effect sizes, including Trails A (*d*= 0.34) and WCST-Categories Achieved (*d*= 0.37). When the authors examined IQ, age, sex, and ethnicity, as moderators, they found no significant effect. When broken into subgroups which operationalised antisocial behaviour using different criteria (antisocial personality disorder, conduct disorder, psychopathy, criminality, and delinquency), effect sizes remained significant, and heterogeneity decreased, indicating a moderation effect. Participants who fell in the criminality and delinquency groups, both generated large effect sizes, whereas those in the conduct disorder and psychopathy groups both had small to medium effect sizes (Morgan & Lilienfeld, 2000).

In 2011, Ogilvie and colleagues replicated and updated the Morgan & Lilienfeld review, locating 126 studies involving *n*= 14,786 participants. Whilst they obtained an increased sample size, their results were quite similar to the original review, demonstrating that antisocial groups performed significantly more poorly on measures of executive functioning with a pooled effect size of *d* = 0.44. Results remained heterogeneous, however, when broken into type of executive function measure, this decreased slightly, but tests of heterogeneity remained significant. Effect sizes for specific measures ranged from *d* = -.13 to 3.05, including tasks such as the self-order pointing task (*d* = 0.83), Porteus Maze Task (*d* = 0.71), delayed matching task (*d* = 0.55) and the spatial working memory task (*d* = 0.54). Further, IQ yielded an effect size of *d* = 0.57, using a random effects model, and using meta-regression, the authors found a significant association between IQ effect sizes and the pooled effect sizes for executive functioning across studies, supporting the notion that executive functions and IQ overlap. This review also found significant moderator variables, indicating that populations recruited from correctional settings, non-antisocial comparison groups, and comorbid ADHD, were associated with larger effect sizes. However, due to the significant variation among methodologies used in the included studies, the authors encouraged that caution be used when interpreting the results (Ogilvie et al., 2011).

Meta-analytic findings (e.g., Reinharth et al., 2014) and individual studies have also demonstrated significant, prospective associations between neuropsychological measures and violent and aggressive outcomes, including Wechshler Intelligence Scales (Beggs & Grace, 2008; Dejong et al., 1992; Fullam & Dolan, 2008; Howard et al., 2014; Nazmie et al., 2013), Stroop tasks (Nazmie et al., 2013), MATRICS social cognition (O’Reilly et al., 2015), and the Iowa Gambling Task (Bass & Nussbaum, 2010). Likewise, neuropsychological measures of disaggregated cognitive abilities have demonstrated the capability to significantly differentiate violent offenders from non-violent offenders, including the Seashore Rhythm test, Line

Judgment tests (Ullman, 1989); Luria Nebraska Neuropsychological Battery (LNNB) (Bryant et al., 1984; Ullman, 1989); Wechsler Intelligence Scales (Brimigion, 2014; Yeomans, 1996); Peabody Picture Vocabulary Test (Kennedy et al., 2011); Stroop Colour and Word Test, and Trail Making Tests A and B (Nazmie et al., 2013). However, the state and quality of the current literature presents limitations in drawing definitive conclusions on this.

1.8 The present research

The evidence presented above indicates that (1) existing structured professional judgement risk assessments have room for improvement, namely in predictive accuracy, and (2) the addition of neuropsychological assessments may have the ability to facilitate some of these improvements. While there appears to be ample evidence to hypothesize that there is indeed a significant association between cognitive impairments (particularly in executive functions), and violence risk, the variability seen across study findings may be an indication of the same predicament that risk assessments are in, where as a result of the numerous neuropsychological assessments used in various forensic populations, individual study findings may be specific to that particular setting, population, assessment, and operationalization of the outcome. For example, a quote from Douglas (1999),

> The old question—can violence be predicted? —is not a good one. More interesting and practical is the somewhat more convoluted question: Which groups of subjects, with which particular characteristics, followed over what periods of time, are likely to exhibit precisely defined kinds of violent behaviour? (p. 161).

Albeit, neuropsychologically informed risk factors can be found in some existing risk assessments, such as impulsivity and lack of insight, found on the HCR-20, impulsivity on the LS/CMI, and risk taking/impulsivity and attention deficit/hyperactivity difficulties both found on the SAVRY, though, these items are often not behaviourally measured, and are instead taken from self-report measures and collateral information, likely introducing validity and bias issues (LaDuke, 2015; Podsakoff et al., 1990). Thus, it is hypothesised that behaviourally measuring cognitive risk factors that are already included in risk assessments, and the addition of those which demonstrate an empirical relationship with violence, may improve the predictive accuracy of existing risk measures or at the very least, will apprise professionals on the cognitive strengths and weaknesses in forensic populations, which will subsequently inform formulations of risk and allow for treatment plans to be tailored to the offender's needs, as well as leveraging their strengths to optimize treatments. With this in mind, and considering how broad risk assessment is, the assessment of cognitive functioning alongside of risk assessments may add an additional dimension to risk prediction by identifying those who are less likely to benefit from treatments, and therefore, may be more likely to recidivate, or those who may need extra support in their care and decision-making, perhaps in lieu of directly predicting violence risk. However, in order to determine which abilities demonstrate an empirical relationship, more meta-analyses examining the association between neurocognition and violence risk in a variety of forensic populations are needed, in addition to more high quality, replicable studies. There is also a need for a consensus on the factors considered to be most important in relation to violence risk, as well as the neurocognitive assessments which produce the largest and most accurate effects, to be reached in this field to robustly examine their individual and cumulative value.

This thesis examines the ability of neuropsychological measures to explain violent outcomes, and to add incremental validity to contemporary risk assessments. The overarching aim is to identify neuropsychological domains relevant for inclusion in a structured professional judgement (SPJ) tool that can be used in parallel to existing tools to further increase both their accuracy and utility in treatment planning. It was hypothesised that the addition of these domains in violence risk assessments will increase the accuracy of clinical judgements and formulations of risk, improving our ability to accurately target treatments; and adapt them to maximise their efficacy. The decision to eventually develop a tool rather than making a series of recommendations based on the thesis findings, was founded on guidance from the Risk Management Authority which recommends using SPJ tools for the assessment of violence risk (Risk Management Authority, 2011), and the evidence that suggests that an unstructured approach to risk assessments is not reliable. SPJ tools provide a framework for professional decision making by identifying factors which the literature indicates should inform the risk assessment process, providing guidance on the strength of the evidence and how the factor should be assessed. Therefore, it was presumed that a SPJ tool would provide the necessary structure and guidance for decision making and assessing risk factors, while also fostering reliability and validity.

1.9 Thesis Overview

This thesis is comprised of 11 chapters, with chapters 4 through 10 presenting novel research findings. Each chapter begins with a preface which links the work within the chapter to issues and questions raised elsewhere in the thesis and contextualizes it within the broader aims of the PhD. In addition to the present chapter, Chapters 2 and 3 provide further information and background on the common aetiological contributors to cognitive impairments in forensic populations and the methodological limitations in violence research and the implications for risk assessments.

To address the overarching aim, the thesis followed the example of the National Institute of Mental Health's MATRICS initiative's development of the MATRICS Consensus Cognitive Battery (MCCB) which sought to identify not only cognitive domains identified by the research literature, but also those that, in a relatively new field, may have yet to be studied sufficiently through expert consensus (Nuechterlein et al., 2008). This was completed in three phases, the first two phases, which were completed in parallel to one another, included a critical systematic review of the literature and a series of meta-analyses examining the association between neuropsychological measures and violent outcomes (Chapter 4), and an international Delphi study which sought to achieve an expert consensus on the neuropsychological functions believed to be important for inclusion in a violence risk assessment (Chapter 5). Taking the findings from both studies, a neuropsychological battery was developed, and a feasibility and pilot study was designed to a.) test the feasibility of implementing a neuropsychological battery into forensic settings, and b.) pilot the battery in two distinct samples of violent offenders, which made up phase three of the methodology (Chapters 6, 7, 8 and 9).

Chapter 10 describes a small case-series pilot study which took a different approach. Using the same neuropsychological battery, this study examined its utility in explaining the variance in behaviour associated with imminent aggression over a 14-day period. Further, the study was designed with a secondary aim of overcoming common methodological and conceptual limitations in violence research, which were largely identified through conducting the systematic review for this thesis, as well as drawing from the personal experiences and lessons learned during the completion of the pilot and feasibility study.

The final chapter presents a general discussion and synthesis of the findings from the empirical studies, linking the findings back to the literature and to the central aim of the thesis. The findings are then used to reflect on the implications and recommendations for future practice and research. The thesis closes with concluding remarks on the predictive utility of neuropsychological measures for violent outcomes and emphasizes the importance of conducting routine neuropsychological assessments in forensic populations to enhance treatment effectiveness and benefits for the individual.

2 Aetiological Factors for Neurocognitive Dysfunction in Forensic Populations and their Association with Violence Risk

Chapter Preface

Chapter 1 provided an overview of the research literature, contextualised the aims of the thesis, evidenced the cognitive impairments in forensic populations, and why they are hypothesised to be related to violence risk. This chapter takes a step back and briefly reviews the literature on the common aetiological factors (e.g., head injuries, alcohol and substance misuse, psychotic illness, and adverse childhood experiences) that contribute to cognitive impairments seen in forensic populations, and how these are posited to relate to violence risk. This chapter fits into the aims of the thesis as it provides further evidence to support the notion that the assessment of neuropsychological abilities in forensic populations may aid in the assessment of violence risk.

2.1 Introduction

Recent research has made substantial progress in understanding the link between cognitive impairments and violent behaviour (Morgan & Lilienfeld, 2000; Ogilvie et al., 2011; e.g., Reinharth et al., 2014), and as a result of the impact that cognitive impairments may have on offending and violent behaviour, it is essential to understand their aetiology. As read in the previous chapter, damage to certain areas of the brain may manifest as behaviours which are conducive to offending and may disinhibit the functions which allow individuals to behave pro-socially and desist from offending. Common causes of cognitive impairments in offending populations with relevance to this thesis are head injuries, alcohol and substance misuse, a diagnosis of a psychotic illness, and adverse childhood experiences. Whilst all of these can individually contribute to cognitive impairment, they commonly co-occur in forensic populations, further exacerbating the negative effects. For instance, violent psychiatric patients are approximately four times more likely to have a closed head injury when compared to non-violent psychiatric patients, and are more likely to be diagnosed with co-occurring substance abuse (Dinn et al., 2009). Further, cognitive deficits related to certain medical and psychiatric diagnoses are often worsened by head injuries and alcohol/substance misuse and may result in damage to similar underlying mechanisms which increase violence risk. For example, the misuse of substances may lead to head injuries that cause behaviours shown to increase the likelihood of aggression, namely, irritability, emotional instability, and impaired inhibitory control (Dinn et al., 2009). It is therefore the aim of this chapter to briefly describe these common causes for cognitive impairment and how they relate to violence risk, to further contextualize the aims of this thesis.

2.2 Head Injury

Head injury, defined as any trauma to the head, not including superficial facial injuries, is the most common cause of death and disability in people aged 1-40 years in the United Kingdom (UK) (NICE Clinical Guideline 176, 2014). The most common form of head injuries are traumatic brain injuries (TBI), often being the most severe, and can be described as a structural injury and/or physiological disruption of brain function, which has been traumatically induced as a result of an external force (NICE Clinical Guideline 176, 2014), and can also result in a direct disturbance of cognitive and neurological functioning (Wortzel et al., 2013). The main difference between head injuries, such as concussions, and TBI, are the severity of the injury. Severity is often operationalised by the presence or absence of a loss of consciousness (LOC), and the duration of this, as well as symptoms such as amnesia. For example, a LOC of 0-30 minutes is often considered moderate, 30 minutes or more is often considered to be moderate to severe, depending on the injury, and a mild injury, typically referred to as a concussion, rarely leads to LOC or permanent brain damage, however, these differ across measures and studies. Conversely, moderate to severe TBI will often lead to long-term cognitive, behavioural and psychiatric disturbance (Williams et al., 2018). Along with the undoubted impact that TBI has on individuals, are the lasting impairments which interfere with one's ability to function daily, and often, to engage in pro-social behaviours.

TBI is considered an ongoing health condition and has been linked to a variety of behavioural, emotional, cognitive, and somatic symptoms (Ponsford et al., 2008). The neuropsychological effects of TBI most commonly result in impairments characterised by impulsivity and deficient social cognition, as well as executive dysfunction, impacting abilities such as memory, attention, concentration, and planning, as well as emotional regulation (Williams et al., 2018). While milder TBIs can lead to issues with inhibitory control and attention, frontal lobe injuries can increase an individual's risk of impulsive aggression, impair decision making, and increase inappropriate social behaviours (Azouvi et al., 2016). TBI is also associated with slowed information processing, impaired social cognition, personality and behavioural changes, and impaired insight (Arciniegas & Wortzel, 2014; McDonald, 2013; Spikman et al., 2000; Willmott et al., 2009). Further, there is evidence to support the significance of TBI over other common risk factors for cognitive and behaviour dysfunction.

For example, using the Ohio State University TBI Identification (OSU TBI-ID) questionnaire to collect retrospective information in a state prison sample of n= 210, Bogner and Corrigan (2009) found that the extent of the prisoners' exposure to TBI remained significantly associated with TBI consequences (e.g., cognitive and behavioural functioning) over and above other potential influencers (e.g., learning disability, special education, psychiatric treatment history, drug and alcohol history, and demographic characteristics).

2.2.1 TBI and Offending

The association between TBI and offending is hypothesised to be related to damage to the frontal lobes of the brain, affecting social perception, self-control and judgement, emotions and mood (Turkstra et al., 2003). The frontal system of the brain is also the one most often damaged when a TBI occurs, frequently resulting in the dysfunction and impairment of the cognitive abilities it holds, increasing the risk of violent and criminal behavior, namely, impulsive aggression (Azouvi et al., 2016; Williams, Mewse, et al., 2010). As a result of these impairments, individuals with a TBI may be more prone to misperceive situations, behave inappropriately in public, overreact to stimuli, lack communication skills, or act out impulsively, which directly relates to violence risk (Turkstra et al., 2003). Deficits resulting from TBI may also inhibit an individual's ability to comply with correctional authorities, in addition to tending to their activities of daily living (Slaughter et al., 2003), consequently affecting their treatment as well as their ability to be released. Moreover, Williams and colleagues (2010), found an association between having three or more self-reported TBIs and increased violence in offences in young offenders (Williams, Cordan, et al., 2010). The research in this area draws a clear link between TBI and offending, by highlighting the *domino effect* that TBI has on individuals' cognitive abilities, and then the effect that these cognitive abilities have on one's behavior.

2.3 Alcohol and Drug Misuse

Further increasing violence risk, poor health, and cognitive impairments, are alcohol and substance abuse. The World Health Organization (WHO) defines substance abuse as 'the harmful or hazardous use of psychoactive substances, including alcohol and illicit drugs'[4] (World Health Organization, 2019). In addition to substance abuse, psychoactive substances and alcohol can also lead to dependence or 'dependence syndrome' which is a cluster of behavioural, cognitive, and physiological phenomena that develops after repeated substance

[4] From this point, the term 'substance abuse/misuse' will be used to describe both drug and alcohol abuse, unless explicitly noted otherwise.

use, and often leads to several difficulties, including a strong desire to take the drug or alcohol, making the drug or alcohol a higher priority than other obligations, and physical withdrawal (World Health Organization, 2019). Harmful alcohol and substance misuse has also been associated with long-term effects on neuropsychological functioning.

The potential for a decrease in neuropsychological functioning due to alcohol and substance abuse has been well established in the literature (e.g., Fernández-Serrano et al., 2011; José Fernández-Serrano et al., 2010; Van Holst et al., 2011). Long-term drug misuse and dependence have been associated with dysfunction in the prefrontal cortex (PFC), leading to impairments in several cognitive abilities, including executive functioning (Verdejo-García et al., 2004). The PFC is made up of the dorsolateral prefrontal cortex (DLPC), orbitofrontal cortex (OFC), and the anterior cingulate cortex (ACC) (Crews & Boettiger, 2009). Each domain oversees its own cognitive abilities, therefore when there is damage within an individual area, different cognitive deficits will emerge. For example, damage to the DLPC may result in impaired working memory and cognitive flexibility, whereas when the OFC is damaged, individuals may have difficulties with processing emotions or making decisions (Bechara, 2005). With acknowledgement of this, researchers have found that individuals dependent on substances often present with behavioural problems similar to those seen in patients who have a damaged or impaired PFC and its components, for example, apathy, lack of initiative, low motivation, poor emotional regulation, poor judgement, impulsivity, disorganized behavior and the neglecting of goals (Verdejo-García et al., 2006). Additionally, while studies have shown that recreational use of certain drugs may cause mild executive deficits, heavier use leading to dependence may impair several executive functions, including selective attention, inhibition, and flexibility (Verdejo-García et al., 2004). Lasting executive impairments have also been observed in individuals who have engaged in excessive alcohol use or who have alcohol dependence, namely in, fluency, working memory, inhibition, cognitive flexibility, and decision-making (Chanraud et al., 2007; Loeber et al., 2009; Pitel et al., 2009). Moreover, alcohol abuse has been shown to affect more specific functions such as sustained attention, planning, and flexibility than the co-abuse of other substances such as cocaine (R. Z. Goldstein et al., 2004) or heroin (Fishbein et al., 2007). A more recent meta-analysis of 77 studies examining group differences between alcohol dependent adults (n= 2620) and healthy controls (n= 2576) found moderate and large (respectively), significant effects on neuropsychological measures of planning and problem solving (g= 0.77), and inhibitory abilities (g= 0.80), as well as a moderate pooled effect for executive function measures (g= 0.64) with the alcohol dependent group presenting with more deficits (Stephan et al., 2017).

2.3.1 Alcohol and Drug Misuse and Offending

Compared to the general population, prevalence rates of problems with substance use in offending populations are significantly higher (Fazel et al., 2006), and substance use has been consistently linked with criminal offending, high-risk behaviours, and an increased risk of violent crime (Chandler et al., 2009; Hoaken & Stewart, 2003). Likewise, depending on the type of drug, drug users are three to four time more likely to offend than non-drug users (Bennett et al., 2008), for example, compulsive violence for economic gain is more likely to be seen in cocaine and heroin users (P. J. Goldstein, 1985), though, a more recent study differentiated the two and found that opioid dependence was associated more with property offending, while stimulant dependence was associated with recent violent crime (Sutherland et al., 2015). More generally though, individuals who have a substance use disorder contribute more to violent behavior, than all other psychiatric diagnoses combined (Pulay et al., 2008). Due to the overwhelming impact that substance misuse has on violent behavior, P.J. Goldstein (1985) proposed three theories to help further explain the link between the two: *psychopharmacological violence,* where violent acts are committed as a direct result of the physiological effects of the substance used (Kuhns & Clodfelter, 2009); *systemic violence*, which is linked to the aggressive patterns often seen in drug dealing; and *economic-compulsive violence*, which has to do with violent crimes committed for economic gain to support substance use. Research has supported these theories, for example, using data from the Pittsburg Youth Study (*n*= 850 boys), Farrington and colleagues found that individuals who use and sell drugs were significantly more likely to commit homicide or to be a victim of homicide (Farrington et al., 2012). Also, according to findings from a 2013 national survey of *n*= 3080 12 to 17 year old's on drug use and health, it was discovered that compared to individuals who did not sell drugs in the general population, adolescents who were involved in selling drugs, were also involved in more fights, gang fights, criminal attacks, and handgun usage (Shook et al., 2013). Moreover, a study which measured the correlates or risk factors for property and violent offenders who inject drugs (*n*= 887), found that 73.3% of those who reported committing a violent offence in the previous month were also under the influence of drugs and/or alcohol at the time, and when asked about their motivation for committing the violent offence, 46.7% reported that it was opportunistic, perhaps indicating significant impulsivity within this group of offenders (Sutherland et al., 2015).

Additional theories support the notion that violence co-occurs with substance use due to intrapersonal factors, such as sensation seeking, easy drug access, and recent drug offers (Salas-Wright et al., 2016), however, they found that despite greater sensation seeking being reported in youth who are violent, it was not a strong predictor for substance use morbidity in repeat violent offenders, and was unable to distinguish those who have a substance use disorder from those who do not. Equally, others have suggested that both acute alcohol intake and intake of small quantities can impair executive functions (Kringelbach, 2005; Ridderinkhof et al., 2002), which may lead to careless or inappropriate behaviour, often making a person more vulnerable to violent victimization, or more likely to respond to situations in a violent manner (Hawkins & Trobst, 2000). Moreover, individuals with alcohol dependence, similar to individuals dependent on other drugs, are characterised by poor impulse control, and are more commonly diagnosed with impulse control disorders compared to healthy controls (Lejoyeux et al., 1998; Virkkunen et al., 1994).

2.4 Psychosis

Major mental illness has also been implicated in cognitive dysfunction, namely psychosis. *Psychosis* is the experience of delusions, hallucinations and disordered thinking, and *psychotic illness* refers to the experiencing of these symptoms to the point of distress and/or functional impairment (Sheffield et al., 2018). While not a *symptom* of psychotic disorders, psychotic illnesses such as schizophrenia, bipolar disorder, and schizoaffective disorder are often characterised by cognitive dysfunction, indicating that cognitive dysfunction is a core feature of psychotic illnesses (Sheffield et al., 2018). Individuals who are diagnosed with schizophrenia, for example, have been characterised by generalized impairments in IQ, attention, executive functions, memory, language, perception, visuospatial abilities, motor speed, and social cognition (Green et al., 2019; Reichenberg & Harvey, 2007). In addition to those observed after a diagnosis is made, research has also demonstrated that cognitive dysfunction precedes the onset of psychotic symptoms, further suggesting cognitive impairments are a core feature of schizophrenia (Seidman & Mirsky, 2017). Like schizophrenia, bipolar disorder with psychotic features has also been liked to cognitive dysfunction, such as IQ deficits, impaired processing speed, working memory, executive functioning (Kravariti et al., 2009), and cognitive flexibility (Trisha et al., 2018). A notable difference between schizophrenia and bipolar disorder are the degree of pre-morbid cognitive impairments leading up to a first episode of either disorder, where research has suggested that less severe impairments are observed in pre-morbid bipolar disorder than those observed in pre-morbid schizophrenia (Kendler et al., 2016; Seidman et al., 2013). However, little consistency has been found between the cognitive profiles of individuals diagnosed with bipolar disorder with psychotic features and those without (Ancín et al., 2013), though more severe deficits in verbal learning and executive functioning (Levy & Weiss, 2010) have been associated with psychotic features, as well as more severe illness (Simonsen et al., 2011), and poorer outcomes (Bora et al., 2010). Cognitive impairments have also been associated with psychosis in other disorders, including mood disorders such as major depression. Whilst there is limited research on the pre-morbid functioning of individuals at high-risk for depression with psychosis (Seidman & Mirsky, 2017), deficits in IQ, verbal learning, category fluency, and Trails B of the Trail Making Test have demonstrated small to large effect sizes (e.g., z = -0.2 to -0.9) relative to normal controls, even after controlling for IQ (Zanelli et al., 2010). Though, interestingly, individuals with depression with psychosis have cognitive profiles which are similar to individuals who are diagnosed with first-episode schizophrenia (Zanelli et al., 2010). Last, although there is little research examining cognitive functioning in individuals with psychosis-like experiences, those which have, found that the presence of sub-clinical psychosis was also associated with poor processing speed (Rössler et al., 2015), and that there was a negative association between psychosis-like experiences and general cognitive ability (Sheffield et al., 2016), suggesting that despite severity of mental illness, the presence of psychosis is related to impaired global functioning (Seidman & Mirsky, 2017).

2.4.1 Psychosis and Offending

Schizophrenia and other psychotic illnesses are often the most prevalent diagnoses in forensic psychiatry (Scottish Government, 2016). However, most individuals with a diagnosis of a psychotic illness will not be violent, and although a diagnosis of schizophrenia confers an increased risk of violence relative to the general population, this has been attributed to symptoms, comorbidity, cognitive impairments, and neurobiology (Fazel, Gulati, et al., 2009; Fazel, Langström, et al., 2009; O'Reilly et al., 2015; Soyka et al., 2007). More specifically, dynamic risk factors for violence in psychosis include co-morbid alcohol and substance misuse, nonadherence to treatment, impairments in insight, impulsivity, presence of positive symptoms, and trauma (Ullrich et al., 2014; Volavka & Citrome, 2011; Witt et al., 2013). As a core feature of psychotic illness, the relationship between cognitive impairments and violence risk has been meta-analysed, but with inconsistent findings. For instance, a 2014 meta-analysis of 4,764 participants found that global cognitive impairments (r= -.13) and lack of insight (r= .09) were significantly associated with aggressive outcomes in schizophrenia (Reinharth et al., 2014), and other studies have reported increased aggression and violence in schizophrenia in the presence of additional factors. For example, the co-occurrence of personality disorders, substance use disorders, and schizophrenia, have been shown to increase aggressiveness in some individuals (Cuffel et al., 1994; Fazel, Langström, et al., 2009; Tiihonen et al., 1997), and two meta-analyses have supported this notion, demonstrating that psychosis in particular is more associated with violence risk when it is co-morbid with substance use disorders (Douglas et al., 2009; Fazel, Gulati, et al., 2009). An acute diagnosis of bipolar disorder has also been implicated in increased violence risk (Feldmann, 2001), though, it also has the highest comorbidity with substance use disorders (Pettinati et al., 2013), and is often associated with treatment noncompliance, higher rates of hospitalization, worse functional outcomes, and a lower quality of life (Jaworski et al., 2011; Lagerberg et al., 2010; Mazza et al., 2009).

2.5 Adverse Childhood Experiences

The experience of early life adversity has been associated with negative physical and mental health outcomes in adulthood (Pechtel & Pizzagalli, 2011). Adverse childhood events (ACEs) include verbal, physical and sexual abuse, neglect, household mental illness, household criminality, and household substance abuse, to name a few (Hughes et al., 2017). When compared to the general population, individuals with a diagnosis of schizophrenia and bipolar disorders have reported higher rates of early life stress and adversity (Etain et al., 2010; H. L. Fisher et al., 2010, 2011), and trauma and early life stress have also been implicated in impaired cognitive functioning (Koenen et al., 2003; Perez & Widom, 1994). Adverse experiences in children have been linked to impairments in general cognition, and on domains of attention, abstract reasoning, and executive function (Beers & De Bellis, 2002; Mezzacappa et al., 2001), while early life stress in adults has been associated with impaired intellectual functioning, memory, and working memory (Navalta et al., 2006). Moreover, clear links have been made between early life stress and reduced cognitive functioning in psychosis, such as working memory and executive function in first episode (Aas et al., 2011), chronic psychosis (Shannon et al., 2011) and schizophrenia spectrum and bipolar disorders (Aas et al., 2012). Neurobiological studies have also drawn links between early life stress and adversity on the development of the prefrontal cortex, suggesting these experiences negatively impact the prefrontal cortex and the functions it subserves (McEwen, 2012; Underwood et al., 2019). Poor performances on measures of IQ, language, memory, attention, and executive functions have been seen in groups of neglected children (De Bellis et al., 2009), as well as children who spent part of their life in an institutionalized environment (van den Dries et al., 2010). Also, both narrative and quantitative reviews have evidenced that children who have been abused often have reduced gray matter in regions associated with emotions and self-regulation (Lim et al., 2014; Teicher & Samson, 2016). Some studies have also focused on subgroups of abuse, such as emotional, sexual, and physical, with a history of emotional abuse being more common in bipolar disorders, and physical abuse in schizophrenia (Etain et al., 2010), suggesting a clear overlap between mental illness, trauma, and cognitive impairments. Moreover, alcohol and substance misuse may be used as maladaptive coping strategies for symptoms of mental illness, or the trauma associated with childhood, adding a further layer of complexity.

2.5.1 Adverse Childhood Experience and Offending

In a meta-analysis of over 250,000 participants, the pooled odds ratio for individuals with four or more adverse childhood experiences for violence perpetration was 7.51 (95% CI 5.87-11.18), with moderate heterogeneity (Hughes et al., 2017). Moreover, pooled odds ratios for problematic drug use in these individuals was 10.22 (95% CI 7.62-13.71) with low heterogeneity, and problems with alcohol use had odds of 5.84 (95% CI 3.99-8.56), though the heterogeneity was significant at 75%. In addition to the loss of important functions as previously outlined, adverse childhood experiences may also lead to adult violence due to the normalization of violence in childhood (Denson, 2021), or the "cycle of violence" where individuals who were abused as children, grow to be violent adults (Braga et al., 2018). Although adverse childhood experiences were not a main focus of this thesis, their relevance underscores the complex and diverse factors that contribute to violence risk.

2.5 Co-Occurring Head injury, Substance Misuse, & Psychosis

Head injury and substance misuse (often in conjunction with mental illness) often contribute to a decrease in neuropsychological functioning, and an increased risk for violent and offending behavior. The relationship between head injury and substance misuse appears to be circular, where it is difficult to determine which came first, further complicating the ability to investigate specific risk factors for each. While research indicates that damage to the orbitofrontal cortex due to a TBI can cause an organic personality disorder which is conducive to substance abuse (Bjork & Grant, 2009), it also suggests that substance use disorder and drug and alcohol intoxication increases the risk of obtaining a TBI, and the link between the two is most commonly in the direction of substance misuse causing a TBI (Cherpitel, 2007; Taylor et al., 2003). This was further supported by conclusions of a review by Bjork and Grant (2009) which suggests that while substance misuse before TBI may increase the risk of suffering a TBI, TBI also increases the chances of relapsing or becoming dependent on drugs or alcohol post-injury, and can make recovery more difficult. Additionally, pre-injury acute substance misuse increases the likelihood of suffering a more severe TBI (Parry-Jones et al., 2006). Additional research has also supported this, arguing that the risk of developing a substance related disorder was 3.62 times higher in individuals with a TBI compared to those without (Wu et al., 2016). Similarly, a systematic review of 29 individual studies evidenced that in a criminal justice population, rates of drug abuse were higher in individuals with a TBI at any age, when compared to controls (McKinlay & Albicini, 2016). Although methodological weaknesses of individual studies (e.g., self-reports, missing information on timing and severity of TBI, as well as confounding factors) made it difficult for the authors to definitively draw conclusions, this review further exemplifies the overlapping nature of these aetiological factors. Aside from the questionable directionality, the adverse outcomes that overlap with the misuse of substances both before and after a TBI are particularly concerning. Heavy substance misuse following a TBI is linked to poorer medical, social and vocational outcomes (Ponsford et al., 2007), as well as a greater risk of suffering repeated head injuries (Weil et al., 2016). Alternatively, individuals with a history of substance misuse before a TBI, have an increased risk of substance use disorder, mood disorders, and attempted suicide after a TBI injury (Weil et al., 2016).

Research has also suggested that symptoms of psychosis will be experienced in up to 10% of individuals who suffer a TBI (see Batty et al., 2013 for a review), and the emerging cognitive impairments in these cases show similar patterns to those with a diagnosis of schizophrenia. For example, in a review of case studies of individuals with psychosis after TBI, 59% reported memory deficits, and 41% reported executive and spatial deficits (Fujii & Ahmed, 2002), and in a study comparing *n*= 45 individuals who experienced psychosis following a TBI to *n*= 45 controls matched by age and gender who did not experience psychosis after a TBI, found that the group experiencing psychosis was more likely to demonstrate impaired performances on measures of executive functions (77% impaired), verbal memory (83% impaired), and language (54% impaired) (Sachdev et al., 2001), perhaps suggesting an additive effect between TBI and psychosis. Additionally, a study comparing cognitive functioning of *n*= 24 individuals who developed a psychotic disorder following TBI (PDFTBI), *n*= 21 individuals with a TBI without psychosis (TBIWP), and *n*= 24 individuals with schizophrenia and no TBI, found significantly lower scores in the PDFTBI group on measures of intelligence, vocabulary, verbal memory, and executive functions compared to normal controls, and likewise, the schizophrenia groups had significantly lower scores on measures of intelligence , working memory, verbal spatial abilities, and executive functions, however, there were no significant differences observed between the TBIWP group and healthy controls after Bonferroni correction (Fujii et al., 2004). Unfortunately, the authors did not report effect sizes to support their findings, though, these results suggest the emergence of similar cognitive impairments in psychosis even in the presence of different aetiologies, though notably, the schizophrenia group exhibited the most impairments out of all the groups.

Research has also linked co-occurring psychosis and alcohol and substance misuse with deficient cognitive functioning, however, the additive effect from some substances may be more subtle. For example, a 1999 study found that the cognitive deficits related to schizophrenia with and without the presence of alcohol use disorder, were more severe than alcohol use disorder on its own, showing deficits in working memory, abstract reasoning, and social comprehension and perception (D. N. Allen et al., 1999). However, the literature in this area seems to be split where some have found greater deficits in substance users diagnosed with schizophrenia compared to those who are not (Serper et al., 2000; Sevy et al., 1990), and others have found no differences (e.g., Addington & Addington, 1997; L. Cooper et al., 1999), though this appears to be related to the substance under investigation. Studies of this nature involving cocaine-users have found more differences between groups, whereas studies involving alcohol use and cannabis have not. Another study examining the effects of alcohol, cannabis, hallucinogens, cocaine, and stimulants in n=266 patients with psychosis, found no significant associations cross-sectionally or longitudinally between substances and cognitive functioning, however, they did find that substance use was associated with higher positive symptoms of psychosis (r= 0.18) (Pencer & Addington, 2003).

There appears to be significant overlap between impairments seen in TBI and psychosis, and psychosis and substance use, though the degree of impairments may have more to do with the type of substance used. Due to the general deleterious effects that these factors have on neurocognition and mental health, and the prevalence of these overlapping factors seen in forensic populations, assessments of TBI, substance use, and cognitive functioning should be a priority when working with these populations.

2.6 Co-Occurring Aetiological Factors and Violence

Further compounding the personal effects resulting from TBI, alcohol/substance misuse, psychosis, and ACEs, is their potential to increase violence risk. As the literature highlights, TBI and substance misuse contribute to cognitive impairments which may increase violence risk, however, there are inconsistencies in the literature as to whether individuals with no history of violence prior to cognitive impairments will then become offenders. For instance, an increase in aggressive behavior has been observed in individuals after suffering a TBI, but higher levels of aggression before a TBI are also a risk factor for aggression after a TBI (Williams et al., 2018). Though, a recent study on n= 432 forensic psychiatric inpatients found that head injury significantly predicted violence proneness and violence during admission *over and above* a history of drug and alcohol misuse, and having a diagnosis of schizophrenia (relative risk= 1.45) (Brown et al., 2019). Although directionality could not be considered in this study due to its cross-sectional design, and it was limited by methodological weaknesses (i.e., the use of head injury as a binary outcome and record linkage data), these results were in the direction to further demonstrate the magnitude of the impact that head injury has on violence risk, over and above other major risk factors, including schizophrenia. However, it is not known whether the individuals in this study had a co-occurring diagnosis of schizophrenia with a historical TBI, or if TBI on its own was significant. Given the findings presented above which suggested that patients with schizophrenia demonstrated more global impairments than groups who developed psychosis following a TBI and TBI with no psychosis (e.g., Fujii et al., 2004), the co-occurrence of schizophrenia and TBI would suggest greater cognitive impairments, and likely, an increased risk for violence.

Similar research has also evidenced that the relationship between violence and head injury can be explained by co-occurring substance abuse (Dinn et al., 2009). Whilst there may be traits which provoke individuals to act in an antisocial manner even without substance use and head injury, the combination of substance misuse and TBI appear to be both disinhibitors and destabilizers when it comes to violent behavior. For example, they can be considered disinhibitors, as damage to the PFC which subserves executive functions may result in deficient inhibitory control contributing to more impulsive behaviours, and disinhibiting functions which encourage pro-social behaviour. Likewise, they can be considered destabilizers due to the loss of cognitive functions which may contribute to a failure to care for oneself, mental health problems, and dysregulation of emotions and behaviours. Additionally, dependence on substances may also serve as a motivator for violence when individuals commit crimes to support their addiction. As there seems to be more agreement on the association between substance misuse and violence, it is likely that the numerous confounding variables seen in offending populations have made the association between TBI and violence seem less straight forward, for example, individuals who have a TBI also often have lifetime histories of abuse, trauma, mental illness, and drug and alcohol problems, compared to individuals without a TBI (Williams et al., 2018). Further compounding this may be the clear overlap in cognitive impairments seen in psychotic illness, TBI, and substance use, creating difficulty in drawing clear links between impairments and their true aetiology.

It is also important to note, as mentioned previously, that individuals with a severe TBI may be more prone to misperceive situations, behave inappropriately in public, overreact to stimuli, lack communication skills, or act out impulsively (Turkstra et al., 2003), and perhaps unintentionally engage in aggressive and violent behaviour. This is similar to some substance dependent individuals who also present with comparable behavioural problems, namely, poor emotion regulation, poor judgement, impulsivity, and disorganized behavior, among others (Verdejo-García et al., 2006), as well as those who are diagnosed with a psychotic disorder (e.g., Green et al., 2019; Reichenberg & Harvey, 2007) and have a history of childhood trauma (e.g., Beers & De Bellis, 2002; Mezzacappa et al., 2001). Based on the evidence presented, it is further apparent that the co-occurrence of even just one or two of these factors may have an additive effect, essentially exacerbating dysfunction and behavioural manifestations. Though, when looking at violence risk independently, it remains difficult to understand the causal processes, as these may be attributed to one factor or several, and adding an additional complication is the possible reciprocal relationship between these factors and violence. This notion also highlights the heterogeneity observed in violent offenders, cognitive abilities, or impairments, and in violence research.

Nonetheless, due to the harmful impact that substance misuse, psychosis, and childhood trauma have on the functionality of individuals, they are more likely to be considered in violence risk appraisals, yet, considering the evidence provided and the significant overlap, it is astounding that TBI is not yet a *formal* risk factor for violence and its relevance is often overlooked in the assessment of violence risk. Williams et al. (2018) argued that if nothing else, TBI may be an essential indicator for a range of issues which increase an individual's risk for committing crime, further signifying its importance as an independent risk factor for offending. In other words, although TBI does not often result in offending, it has the potential to increase the risk in certain individuals due to the cognitive impairments and behavioural problems that it can lead to. Though, taking all the presented evidence together, it is further astounding that it is not routine practice for a comprehensive neuropsychological evaluation to be completed on offenders prior to treatment or risk assessment, seeing as much of this evidence implicates the underlying cognitive impairments, that are a result of these aetiological factors, as the reason for increased violence and aggression risk. Moreover, as TBI, psychosis, and substance use can be considered dynamic in nature, they do not directly increase violence risk, instead they represent several underlying mechanisms that theoretically relate to violence risk. Thus, it seems that without knowledge of the location or magnitude of impairments, due to the absence of neuropsychological evaluations in many cases- treatments, clinical judgements, and formulations of risk may not be completely accurate or effective.

2.7 Summary

In summary, due to the resulting cognitive impairments from TBI, alcohol and drug misuse, psychosis, and adverse childhood experiences, the presented evidence suggests that these factors contribute to cognitive impairments that increase the risk of engaging in violent behaviour. Although the contributing mechanisms vary, as well as the directionality of which occurs first, all factors contribute to a decrease in the neuropsychological functions which promote pro-social behaviours. Currently, substance use, traumatic experiences, and major mental disorder, including psychotic disorders are considered risk factors for violence and are included in the HCR-20 version 3 (Douglas et al., 2013), however TBI, and many of the cognitive impairments it is known to cause, are not included as formal risk factors in any risk assessments for violence. This may be due to the many limitations which often emerge in TBI research, namely, the failure to converge on a single definition and the various types of measures used, which has likely contributed to somewhat mixed results. Nonetheless, the topic of head injury and violence has increasingly emerged in the literature and will hopefully continue to be looked at as a true marker of future violence. Further, to reiterate what Williams et al. (2018) argued, because we are now more aware of the neuropsychological consequences of TBI, and how the impairment of cognitive abilities may contribute to violence, if nothing else, in offending populations, TBI should be used as an indicator of someone who may go on to be violent, and more so when it coincides with substance abuse. In addition to the propensity for violence, the prevalence of these factors in forensic populations present important implications for clinicians. Notably, aetiological risk factors for cognitive impairments may also hinder an offender's ability to fully benefit from treatment programmes (Williams et al., 2018). Thus, to aid in the identification of specific treatment targets, to optimize treatment gains, and to improve the accuracy of clinical judgement, it is recommended that cognitive abilities of offenders, especially those with these risk factors, are comprehensively neuropsychologically assessed to understand their cognitive strengths and weaknesses.

3 Methodological Limitations of Violence Research and the Implications for Contemporary Risk Assessments and Clinical Practice

Chapter Preface

Following the structure of Monahan's (1988) criticisms of risk assessment research, this chapter outlines the major methodological weaknesses in current risk assessment research, and the implications they have on contemporary risk assessments and clinical practice. Each section is followed by a list of recommendations on how to improve future research. Acknowledgement of the long-lasting limitations observed in risk assessment research is relevant to the current thesis, as it provides additional context and a more in-depth rationale for the central aim of the thesis; that is, to improve the predictive accuracy of structured risk assessments. In addition to our hypothesis that the addition of neuropsychological risk factors for violence will improve the predictive accuracy of risk assessments, it is also postulated that improving the quality of research will further contribute to this aim. Thus, this chapter looks at the issues in risk assessments from a bottom-up perspective with the central recommendation that to improve the performance of risk assessments, the research which informs them must improve first. As will be described in subsequent chapters, the methodological weaknesses identified in this chapter were critically and carefully considered during the developmental stages of the two final studies in this thesis.

3.1 Introduction

Over three decades ago, Monahan highlighted four main methodological weaknesses that he identified whilst reviewing research on the assessment of violence risk in mentally disordered populations, which included: the use of "weak criterion variables", "impoverished predictor variables", "constricted validation samples", and "unsynchronized research efforts" (Monahan, 1988; pp. 251-255). With the addition of one more criticism of risk assessment research (e.g., a lack of transparency in reporting findings) and a discussion on a common limitation observed in this research (e.g., low base rates), the following sections will briefly elaborate on these limitations, and their relevance for current research, followed by recommendations for future research. Subsequently, the implications that these weaknesses have on current risk assessments and clinical practice will be outlined.

3.2 Weak Criterion Variables

In 1988, Monahan highlighted the use of weak criterion variables as one of the most persistent criticisms of risk assessments. For decades, the terms *dangerousness* and *violence* have been used to describe individuals who pose a substantial risk to society or to themselves, however, for the same amount of time, there has been little consensus on what exactly is meant by the term violence. Notably, this limitation has been emphasized is several papers throughout the years (S. T. Harris et al., 2013; Hart, 2011; Yang et al., 2010), and in the HCR-20 V3 manual, where Douglas et al. (2013) conceded this ambiguity in definitions of violence, and proclaimed that they complicate all aspects of risk assessment and management. Further Hart (2011) shared this view, arguing that without a precise definition of violence, there cannot be a precise understanding of what causes it. In conjunction with a vague understanding of violence, ambiguous definitions complicate the way data is interpreted (Bjørkly, 1995), impacts the synthesis and comparison of studies in meta-analyses, and results in predictions of unspecified types and severity of violence. Moreover, they contribute to less accurate predictions, inhibit the dissemination of replicable studies, and essentially impacts the overall quality of research, all of which eventually informs clinical practice. While imprecise and unclear definitions of violence impact individual studies, the problem becomes much more apparent in reviews where similar outcomes are meant to be synthesized for analysis. Bjørkly (1995) highlighted this as a limitation in a review he conducted looking at prediction of aggression in psychiatric patients, where he found that the way aggression was operationalised as a criterion measure, ranging from very specific to the inclusion of 'questionable events', confounded the interpretation of the data. More recently, S.T. Harris and colleagues (2013) completed a review on how violence is quantified in violence research. While this review focused more on tools used to assess violence in a more rigorous way, predictably, the various ways in which violence is defined in this field was also highlighted. Their findings demonstrated varied definitions and dichotomized categorizations of individuals, while only a small number of studies implemented objective measures to gain a deeper understanding of violent behavior. Additionally, in 1983, Harrè and Lamb (1984; as cited in Parrott & Giancola, 2007) reported that there were over 200 different definitions of aggression, and with that being over 30 years ago, there are likely even more today.

A contributor to poorly and vaguely operationalised outcomes, is their lack of specificity. For example, violence is technically a subtype of aggression, and aggression and violence should be thought of as a continuum of severity, where aggression may include behaviors such as shoving, pushing, and yelling, whereas violence suggests the intent to cause severe physical harm (Dewall et al., 2011). In this context, and according to DeWall and colleagues, aggression is behavior which is intended to harm a person who wants to avoid the harm, and violence is an extreme form of aggression where the goal is to cause severe harm, usually physical injury, or death. One of the issues with specificity in outcomes is that the terms *aggression* and *violence* are frequently used interchangeably, although they are not the same (Dewall et al., 2011; Serper, 2011). Further, violence is often used as an 'umbrella term' to account for all types of behaviours and levels of severity, which can be detrimental to the interpretation of results. Likewise, compiling a broad range of behaviours into one composite score and calling it violence can lead to the overgeneralization of which individuals are violent, and may put individuals arbitrarily into groups which they realistically do not belong in. This assertion was made by Bjørkly (1995), who argued that broad definitions of aggression allow behaviours such as serious physical assaults and verbal aggression to be treated as equal indicators of aggression. Moreover, specifying the type of violence and severity is especially important in inpatient settings where proxies are frequently used as indicators of aggression, as the lack of specificity may impede the generalizability of results to other inpatient settings and contexts.

Bearing this in mind, it seems negligent to ignore the severity of violence and to not explicitly differentiate it in research, whenever possible. While it is acknowledged that a common limitation in violence research is low base rates of violence, which quite often constricts the ability to robustly analyse certain outcomes (e.g., Quinsey, 1981), it is still crucial to report individual incidents and how they were rated, otherwise, as Bjørkly (1995) highlighted, researchers lose the ability to determine specific risk factors for serious violence as compared to those for verbal aggression, for example. In addition to being specific about severity of incidents, types of aggression and violence should be explicitly stated whenever possible. Geen (1998) suggested that there are seven different types of aggression; reactive, proactive, instrumental, predatory, irritable, impulsive, and psychotic, and while these will not be described in detail here, the list further demonstrates that not all aggressive behaviors are equal. Though it may not be a priority of all research to measure specific types of aggression, at this time, reporting significant findings for predictive accuracy of non-specific types of aggression and violence is arguably less meaningful for both future research and clinical practice, and will not help the field to advance methodologically.

3.2.1 Recommendations for Future Research:

- Define outcomes thoroughly and specifically.

- Measure and report severity, frequency, and subtypes of violence when the information is available.
- Be explicit in how violence is operationalised and everything that is included in it.
- Use an objective rating scale to quantify violence (see S.T. Harris et al., 2013 for a review).

3.3 Impoverished Predictor Variables

Researchers have intimated that there are three main aspects influencing an individual's propensity to use violence, namely; motivators, disinhibitors, and destabilizers (Hart et al., 2016), all of which can be conceptualised as *risk factors*. When considering risk factors, often in forensic research, investigators need to rely on variables of convenience, such as those held in institutional files, which may contain incomplete and inaccurate information (G. T. Harris et al., 2015). Moreover, file information is often based on self-report with little or no evidence of corroboration to ensure accuracy of the information. Additionally, in 1994, Monahan and Steadman pointed out that although violence is a construct which is multifaceted, many risk assessments not only encompass only a narrow range of predictor variables, but they are often chosen for inclusion without a theoretical basis. For example, the 20 predictors chosen for inclusion in the HCR-20 were based on expert opinion, rather than systematic review or multivariate analyses (Webster et al., 1997 as cited in Fazel, 2019). To improve this, Monahan and Steadman suggested that patients be assessed on several domains, including psychological, social, and biological, and begin to develop measures which are theoretically relevant to violent behaviour. However, a caveat to this, which remains relevant today, is that information on many theoretically derived predictors may not be readily available for collection from records and files (G. T. Harris et al., 2015), perhaps contributing to an increased need for resources which may not be available (e.g., an assessment or evaluation may need to take place, rather than simply recording information from a file). This difficulty may be more pertinent for dynamic risk factors rather than static, for example, due to the changing properties of the dynamic predictors, they often cannot be accurately recorded from a file unless they are frequently or routinely assessed. Further, their changing nature may introduce challenges into data collection, and as G.T. Harris et al (2015) suggested, the frequency of observations of the risk factor needs to match the rate at which it fluctuates, otherwise recidivism cannot be accurately predicted.

Aside from the cited difficulties surrounding the ascertainment of dynamic risk factors, their addition to risk assessments has allowed clinicians to gain a more complete picture of the individual and made treatment and management of behaviours additional priorities of assessing risk. However, the lack of a theoretical basis for these factors, that is, their relationship to the aetiology of offending, has persisted (Mann et al., 2010; Ward & Beech, 2015). Sturmey and McMurran (2011) argued that without knowledge of the theoretical mechanisms that underpin dynamic risk factors to cause violence, their modification and management may be less successful, and may hinder the development of comprehensive case formulations. Ward & Beech (2015) speculated that the lack of theoretical basis in risk assessments may be due to the theory of scientific method accepted by researchers, arguing that researchers in this field are "narrowly empiricist" (Ward & Beech, 2015, p. 101) and are more focused on the formulation of empirical generalisations and predictions, than explanation (Haig, 2014; Ward et al., 2005). Following these critiques, Ward & Beech (2015) warned that without clear concepts and theories, the research will lose direction, a notion supported by the current researcher (PhD student). Given that dynamic risk factors are arguably more complex relative to static risk factors, predominantly due to their ability to change, and their vital role in risk formulations and management, it is the opinion of the current researcher that deriving dynamic risk factors based on their theoretical relevance to violence will only strengthen the accuracy of risk assessments, treatments, formulations of risk, and management of behaviours.

3.3.1 Recommendations for Future Research:

- Risk factors being investigated for the prediction of violence or offending should be theoretically derived, and studies should provide evidence for their relevance.
- Unless dynamic risk factors are routinely measured and can be collected in a timely manner, they should be assessed by the researcher to ensure validity and quality.
- When investigating change in dynamic factors, the researcher should ensure that the frequency of observations aligns with the rate at which the risk factor fluctuates.

3.4 Constricted Validation Samples

Due to the nature of violence research, and the barriers often in place in forensic settings, samples for research are often chosen by convenience (G. T. Harris et al., 2015). While this method of recruitment is common in research with human participants, it presents limitations which contributes to the poor predictive accuracy of current risk assessments. For example, with acknowledgement that violence is a heterogeneous construct, recruiting a convenience sample often consists of examining a group of individuals who may only share the common characteristic of having a violent conviction, that is, there may be individuals who are included who have repeatedly offended, and those who have perpetrated violence only once. G.T. Harris et al (2015) suggested that examining a set of homogeneous predictors of violence in heterogeneous populations, may result in researchers overlooking or missing accurate predictions for subgroups within the sample. Further, they rightly suggested that samples of offenders are often biased, for instance, offenders who are more likely to behave violently are often recruited to allow for a large enough base rate of violence to analyse. In addition to bias related to violent offending, the study of forensic psychiatric patients further promotes bias, as only those who are well enough to participate in the study and are observed as having capacity to consent, can be invited to take part. While researchers in forensics are often left without many options for recruitment, theoretically, choosing a sample based on their forensic history, who are also well enough to take part in research, may be counter-intuitive, as it can be assumed that patients who are well enough to volunteer to take part safely and ethically in research, are also less likely to be violent. Although this is not always the case, it appears to be a common scenario, and a difficult one to work around. Comprehensive routinely collected data that is consistent across jurisdictions, and data linkage between criminal justice and health records may improve this, as consent may not be required in these cases, however, quality and completeness of this data may present additional limitations. Findings based on these biased samples may have implications for future research, and risk assessments, for example, considerable inconsistencies in effect sizes across studies, incomparable base rates, and diminished predictive accuracy for specific subgroups of violence. Low statistical power due to a small sample size is an additional artifact of constricted samples, which will also affect predictive accuracy. Moreover, even in robustly designed studies, low power may contribute to unreliable findings, such as a lower chance of detecting a true effect, low positive predictive values, and an inflated magnitude of the effect (Button et al., 2013). Given the difficulty of attaining a large sample in this field, ways to increase statistical power or contingency plans should be at the forefront of planning in the early stages of writing proposals.

3.4.1 Recommendations for Future Research:

- Consider repeated measure designs, as this will separate the error variance from the subject variance, lowering the overall random error.
- Recruit from more than one setting.
- Recruit more than one type of offending group.
- Collaborate with researchers to include additional settings and populations of interest.

3.5 Low Base Rates

Although low rates of outcomes are not technically a *methodological* weakness, they directly affect the predictive accuracy of risk assessments. The base rate is operationalised as the prevalence of the outcome of interest, for the current thesis that is, violent recidivism. Whilst reported base rates of violence have increased as research methodology has advanced, the problem persists as forensic samples in inpatient environments are often well managed, leaving little room to behave violently, and incidents which take place in the community are only recorded if the individual has been caught (charged or convicted), often leaving smaller scale incidents (which would be counted as violent in inpatient settings), unaccounted for. Moreover, when it comes to the accurate prediction of violence, the consideration of base rates is crucial (Monahan, 1981), as the lower the base rate, the more difficult it becomes to accurately predict behaviour. G.T. Harris and colleagues (2015) reiterated the difficulty in observing associations between predictors and outcomes with a low base rate, suggesting that base rates closest to 50% are optimal for identifying useful predictors, which may require longer follow-up periods. For example, studies which have employed long follow-up periods, for example, approaching 20 years, had base rates exceeding 50% (Rice et al., 2013). Although longer follow-up periods may be more effective in increasing base rates, they carry their own limitations, such as high monetary and resource costs, as well as high attrition rates. Moreover, an argument can be made against the examination of the association between dynamic factors and violence over the long term, particularly if the participant is in a treatment programme focused on diminishing recidivism through improving dynamic risk factors, as this will inevitably affect base rates of incidents. Douglas & Ogloff (2003) also noted that the tools chosen to measure the outcome will certainly affect the base rate as well as the type of violence observed in the sample, which in turn can also affect the predictive power and accuracy of instruments, as well as the statistical strength observed between the predictor and outcome.

In addition to these limitations, a low base rate may also increase the chances of the researcher broadening their definition of violence or dichotomizing their sample, consequently, counting less severe behaviours or proxies as violence, and potentially introducing artificial variance. A low base rate is a limitation that is exceptionally difficult to avoid, for example, researchers will often try to choose a sample who has a higher chance of behaving violently during the follow-up period, such as forensic psychiatric patients or individuals with violent convictions, and while this may help the low base rate problem, it then increases the *constricted samples* weakness discussed above. As a result, generalizability of findings may be lacking, however, generalizability will also be an issue if the definition of violence is not explicitly operationalised, thus there needs to be a balance that fits within the aims and objectives of the study.

3.5.1 Recommendations for Future Research:

- The statistical base rate of violence should always be considered in risk assessment research.
- Be aware of the implications of broadening definitions or adding outcomes/proxies in the presence of a low base rate and be explicit about any outcomes that are not *actual* violence.
- Consider a longer follow-up period, and the use of robust tools to measure the outcome that fits within the aims and objectives of the study.

3.6 A Lack of Transparency in Reporting Findings

In Chapter 1 of this thesis, AUC scores were introduced as the most widely used statistic to determine predictive accuracy of risk assessments, however, their independent use has been criticized in the literature, as they are uninformative, and often misleading when they are reported in the absence of other performance indicators (Fazel, 2019). Thus, it is recommended that various performance indicators are also reported such as *sensitivity* (the proportion of offenders that the instrument correctly classified as being high risk), *specificity* (the proportion of offenders that the instrument incorrectly estimated that would commit a crime or behave violently), *positive predictive values* (PPV; the proportion of people that a tool identified as high risk that actually go on to offend), and *negative predictive values* (NPV; the proportion of offenders that are identified as low risk that do not offend). One of the benefits of reporting AUC scores, however, is that they are not dependent on the base rate of outcomes, and they present measures of discrimination (operationalised as sensitivity and 1-specificity) at all possible cut-off points. Sensitivity, specificity, NPV, and PPV, on the other hand, will change if the tool's cut-off changes, and they are sensitive to base rates, making it more difficult to compare tools across studies. Further, *calibration* (how closely the instrument's predicted risk matches the observed risk) is also an important statistic to report, though only if probability scores are used (Fazel, 2019). The absence of this information hinders the comparison of measures and only provides a piece of the picture, however, the lack of reporting additional performance indicators remains a limitation in risk assessment research.

Individual studies examining the performance of risk instruments have been systematically and meta-analytically reviewed (e.g., Fazel et al., 2012; Singh, Desmarais, et al., 2013; Yang et al., 2010), highlighting the lack of reporting of additional performance indicators. For instance, a review which examined the predictive validity of 19 risk measures used in US corrections found that only summary statistics were reported in the papers, and they were only reported for the outcome of *general recidivism*, as opposed to *violent recidivism* (Desmarais et al., 2016). Further, all 19 instruments that were examined provided probabilistic scores for re-offending, though, calibration statistics were not reported for any. This was a similar finding to the Singh et al (2013) review where in 50 studies, not one reported calibration statistics. Given these findings, Fazel (2019) suggested that instruments which are published with low levels of transparency threaten validity, as they cannot be checked by other researchers.

Related to the lack of transparency, risk assessments have also been criticised for authorship bias. In violence risk assessment research, it is not uncommon to find that researchers who were involved in the development of an instrument will publish more positive findings, than an outside group, a phenomenon referred to as *authorship bias* (Singh, Grann, et al., 2013). After a significant authorship effect was found in a large meta-analysis on risk assessment literature (Singh, Grann, et al., 2013), several explanations were offered as to why this might occur, which largely focused on the notion that tool designers use their instruments in the exact way in which they were developed to be used, they design studies which optimize predictive validity, and they ensure proper training is given to tool administrators. Risk assessment tools should only be used in the way they were designed, which comes down to specificity; tools are designed for a specific outcome, population, and method of use, which is more than likely understood by researchers in the field. The issue with *authorship bias* is however, that that tool designers may be less likely to publish studies where their tool performed poorly, which contributes to a bias in peer-reviewed research (Singh, Grann, et al., 2013). Moreover, a decrease in accuracy of prediction is often seen when a risk instrument is applied to a population which it was not originally standardised for (Coid et al., 2011), without complete transparency about the specific way in which the tool is meant to be used, independent groups using the tool are unlikely to replicate the same accuracy as the original development group (Singh, Grann, et al. 2013). Authorship bias is important to acknowledge as it has the potential to threaten validity of research and further contributes to the heterogeneity of the predictive validity of risk assessment tools.

3.6.1 Recommendations for Future Research:

- Provide appropriate performance indicators to ensure full transparency (e.g., sensitivity and specificity).

- Normalise publishing negative results and be explicit when reporting findings.

3.7 Unsynchronized Research Efforts

Monahan declared unsynchronized research efforts as a problem in violence research with mentally ill offenders in 1988, referencing the lack of coordination among research programmes, and subsequently pointing out the rarity of finding two studies which define predictor variables in the same manner (Monahan, 1988). He further highlighted the disparities observed between follow-up periods, and the use of retrospective data with an over-reliance on records and files, opining that these limitations have "seriously hindered the development of the field," (Monahan, 1988, pp. 255). In 2014, this notion was reiterated by Ward & Beech who linked the lack of attention to theoretical issues associated with risk assessments, with the lack of communication between researchers working within the risk assessment field, and in similar areas. Ward and Beech starkly claimed that researchers examining risk of offending and risk assessments appear to only read papers within their own specialisation, alleging that there is a belief that researchers do not find value in learning from each other. Whilst this conjecture has not been observed by the current researcher, there are indications that there is a general lack of communication between researchers in the current literature. For example, following the completion of the systematic review and meta-analysis for this thesis (Chapter 4), several of the limitations identified in the literature matched the consequences of unsynchronized research efforts outlined by Monahan. For instance, ill-defined explanatory and outcome variables, wide variations of research methods across studies, a lack of replication of methods and measures, and an almost complete lack of consistency in findings. This lack of synchronization, as Monahan (1988) put it, "drastically reduces the confidence with which findings can be generalised and impedes the cumulative development of knowledge" (pp. 225).

3.7.1 Recommendations for Future Research:

- Replicate methods of similar research to observe whether this brings more consistency into the findings across studies (this is more specific to the neuropsychology and violence risk research).
- Reach a consensus on a standard of research that is expected in this research area, as well as a standard of quality.
- Carefully consider the operationalisations of key variables, and how they match up with similar studies.

3.8 Implications for Risk Assessments

Taken together, the researcher postulates that the presented criticisms and limitations of risk assessment research are largely contributing to the modest predictive accuracy of risk assessments. This postulation is further evidenced by the notion that a risk assessment is no better than the data that was used to inform it, as the kind of data researchers can accumulate is often met with practical limitations (Haque & Webster, 2019). Whilst it is acknowledged that this is not the first time that these issues have been reflected upon, in the relatively *young* field of research examining the association between neuropsychological functioning and violence risk, many of these criticisms and limitations are blatant and require attention to move the field forward. Although the kind of data researchers can accumulate is often met with practical limitations, the combination of the presented limitations and their natural consequences appear to have a domino effect in research that is not only hindering future research but is also informing clinical practice and future risk assessments. For example, the recruitment of small or constricted samples may be unavoidable, which is widely understood in fields that recruit human subjects, especially vulnerable groups, however, this practical limitation may then affect the outcome and quality of the study. With small samples often comes low base rates of incidents or offences during a follow-up period, and not only does this lower predictive power further, but it also often entices researchers to broaden their definition of violence or aggression. Although broadening the definition of the outcome will likely increase the accuracy of predictions because it is analogous to creating a larger target to hit (Monahan, 1981), it may also come with the introduction of secondary outcomes and proxies for violence. And whilst there is obvious value in investigating secondary and proxy outcomes, the main aim of the study may become less apparent, and researchers may fail to be explicit in the exact behaviours in which they are counting as violence, which inevitably increases heterogeneity across similar studies, and limits comparability and reproducibility. Moreover, introducing secondary and proxy outcomes may diminish the adequacy of predictors, for example, a study examining the predictive validity of risk factors which were shown to have strong theoretical relevance for violence, may not have the same relevance for the proxies of violence, creating additional heterogeneity and inconsistent findings. An alternative scenario observed in the research literature as a result of small sample size and low base rates is the use of binary outcomes to categorize violent and non-violent offenders. Although categorizing offenders is useful for examining predictive performance of risk measures, it can be less meaningful for the general understanding of the severity and frequency of offending. Further, it is evident that innovation is a priority in the research investigating neuropsychological predictors of violence, rather than replicability, which has arguably accentuated these limitations as many appear to cycle through individual studies, and become stark when the findings are synthesised in meta-analyses. And, while meta-analyses are intended to provide robust, quality evidence, their findings are often translated into clinical practice, and used to inform future risk assessments. Collectively, all the limitations listed in this chapter will inevitably decrease the predictive accuracy of risk assessments.

3.9 Implications for Clinical Practice

A major concern of poor-quality research output is its translation into clinical practice. In 1981, Monahan asserted that clinicians were prone to numerous types of *systematic error* when it came to the process of predicting violence. To further clarify this, he explained that there is often a vague understanding of the behaviour that is being predicted, little attention is paid to base rates of violent behaviour, predictors may be based on illusory correlations (e.g., when a correlation is claimed or reported, but an actual statistical correlation does not exist), and the failure to account for environmental factors (Monahan, 1981), all of which have the potential to bias decision-making in assessing risk and decrease their accuracy. Following the general framework of the scientist-practitioner model, which suggests that clinical practice should be informed by scientific research, it may be safe to assume that these systematic errors are at least partially passed down from researchers to clinicians. Albeit, regardless of whether clinicians have a research background, they should have a working knowledge of the key aspects that contribute to the accurate assessment of violence risk and management. Decisions should be made based on the standardised assessment procedures, and they should have a role in the development and testing of emerging and existing risk assessments. To encourage this, researchers should convey empirical findings and literature in a way that is practical for clinicians (Douglas, Cox, et al., 1999). It is posited that a more collaborative effort from both researchers and clinicians may decrease bias in decision making, and perhaps increase predictive accuracy.

3.10 Chapter Summary

In summary, it is clear from the presented evidence that the limitations that Monahan identified in 1988, remain persistent in the violence risk assessment research. Whilst this is not an exhaustive list of limitations, they appear to be the most common in contemporary research (namely, the literature investigating neuropsychological risk factors for violence), particularly, small sample sizes, low base rates, and poor operationalisations of violence. It is also acknowledged that not all the above limitations are easily avoided, and create a balancing act for the research team, however, some are more easily solved. In 2014 following a critical appraisal of measurement in risk assessment research, Grych and Hamby offered an explanation for these weaknesses, suggesting that in light of the technological and statistical advances observed in the literature, it may be the case that researchers have lost sight of some of the most trivial, but important elements in science. They followed this up by saying that complex statistical models are of little value when the most important aspects of the study, such as the measures generating the data, have not been carefully considered. Sharing the same view as the current researcher, they suggested that more attention needs to be paid to the small details, including how violence is conceptualized and measured, and how the central constructs are defined, assessed, and interpreted, for the field to advance (Grych & Hamby, 2014). As will be seen in later chapters, these limitations were critically and carefully considered in the planning stages of the studies that involved violent offending participants, especially, the final study in this thesis.

4 Examining the Cognitive Contributors to Violence in Forensic Populations: A Systematic Review and Meta-Analysis

Chapter Preface

While valuable, existing meta-analytic reviews examining the association between cognitive functioning and violence risk have focused on executive functions, or specific diagnostic groups, thus failing to provide a comprehensive understanding of this relationship. This chapter presents a review that takes a broader approach, first examining neuropsychological factors which differentiate violent from non-violent offenders (part one), before separately analyzing the neuropsychological correlates of violence (part two). Forty-two studies were included, and methodological quality was assessed. This chapter presents the first of the three cohesive studies that are central to the aim of this thesis, which is to assess the predictive utility of neuropsychological measures for violence. The findings from this study were combined with the findings from the Delphi study (Chapter 5) to create a neuropsychological battery of measures which was piloted in two forensic samples in the study presented in Chapters 6 to 9.

4.1 Background

Several studies have found an association between cognitive impairments and violence; however, many reviews have only focused on executive functioning in relation to violence, or have included only populations with a diagnosis of a major mental illness (e.g. schizophrenia, antisocial personality disorder). These meta-analytic reviews have found contrasting results, for example, two recent reviews found that several cognitive impairments were significantly associated with violence in schizophrenia and antisocial personality disorder (Reinharth et al., 2014; Sedgwick et al., 2017), while a third found that lower scores on neuropsychological measures in individuals with psychosis were not significantly associated with violent outcomes (Witt et al., 2013). In contrast, reviews looking specifically at how measures of *executive functions* differentiate antisocial and non-antisocial populations, have found fairly consistent results, showing that antisocial offenders score lower on measures of executive functions when compared to antisocial offenders (Morgan & Lilienfeld, 2000; Ogilvie et al., 2011); albeit, the grand mean effect sizes were quite heterogeneous due to the variation in methodologies between studies.

Meta-analytic reviews on this topic have primarily focused on composite cognitive domains (e.g., executive functions), measures of executive functions, or specific diagnostic groups, leaving gaps in the literature. To address this, the current review aimed to take a broader approach in looking at cognitive abilities associated with violence, by including violent offenders from various settings (irrespective of diagnosis), and by looking at disaggregated cognitive abilities rather than composite abilities, such as executive functioning, and full-scale IQ. This was based on evidence from the three-factor model of executive functions proposed by Miyake et al. (2000), which proclaimed that executive functions are both diverse and unitary. Through latent factor analysis of nine executive function measures, they found that a three-factor model (e.g., inhibition, shifting, and working memory) produced a significantly better fit than one or two factor models, and they suggested that while an individual can be impaired on one executive function domain, they will not necessarily be impaired on another, suggesting that they are separable. Likewise, as there is an overlap between executive functions and intelligence (Miyake et al., 2000), some individuals with impairments in executive functions demonstrate normal intelligence scores (e.g., Damasio, 1994), suggesting that they are also separate constructs. Further, composite measures of intelligence are a composite measure of cognitive abilities which are also separable, and a common distinction made is between fluid intelligence (the ability to solve problems, regardless of previously acquired knowledge) and crystallized intelligence (the use of previously acquired knowledge and skills to solve problems; Carroll, 1993). As fluid intelligence is more sensitive to frontal damage than crystallized intelligence, individuals with frontal lobe damage may show more deficits on measures of fluid intelligence, thus, intelligence tests that use composite scores, such as the WAIS, are less sensitive to frontal lobe damage (Duncan et al., 1995), and composite scores of full-scale IQ may be misleading. Moreover, it is necessary to clarify the strength of individual neuropsychological risk factors for violence and those which differentiate violent subgroups from the wider population to assist in the development of evidence-based risk assessments, to inform targeted treatments for individuals with these risk factors, and to further understand the mechanisms which place individuals at risk for being violent toward others (Witt et al., 2013).

Thus, a systematic review and meta-analysis of 42 studies is presented to examine the range of cognitive factors which are associated with violence toward others. To complete this, the cognitive factors which differentiate violent from non-violent offenders were investigated first, followed by a separate analysis to look at the cognitive factors which are correlated with violent outcomes. Subgroups of offenders which were included in the review (where there were sufficient data) were then investigated to examine the magnitude and significance of cognitive impairments in relation to violence in these specific groups. Congruent with much of the available literature, it was expected that individuals who had a history of violent offending would perform significantly poorer on measures of neuropsychological functioning, and that poorer performance on measures or higher cognitive impairments would significantly correlate with violence toward others.

4.1.1 Review Aims

The present review focuses on the association between cognitive impairments and violent offending toward others in forensic populations (e.g., those who have been either charged or convicted of a violent offence or are in the care of a secure forensic psychiatric hospital).

This review aimed to (a) identify the neuropsychological differences between violent and non-violent offenders, (b) identify cognitive correlates of violent offending in violent, forensic populations and (c) identify cognitive profiles of offenders by conducting subgroup analyses broken down by type of offender (where there was sufficient data), and (d) determine gaps in the literature and areas in need of further investigation.

4.2 Method

4.2.1 Protocol Registration

This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2010), and the protocol was published on the PROSPERO website on 15 August 2016 (study registration number: CRD42016043925).

4.2.2 Search Strategy

Systematic literature searches of electronic databases Web of Science, PsycINFO (1987-May week 1, 2016), Embase (1980-Week 19, 2016), Medline (1946-April week 4, 2016), and CINAHL (1937-August 3, 2016) (all searches were updated on April 19, 2020) were performed (an example of a full search can be found in Appendix E). An attempt was made to include all published and unpublished studies relevant to understanding the effect that cognitive abilities have on violent offending in forensic populations. Search terms related to potential independent variables (cognitive abilities), forensic and mentally ill populations, and outcome terms (violence) were combined. A combination of search terms and search features (e.g., explode) were corroborated with a librarian to ensure the return of as many relevant papers as possible. Terms were combined using a Boolean search strategy: cognitive abilities AND violence AND offenders OR mentally ill offenders AND risk. Manual searches were conducted by inspecting reference lists of articles and reviews to find relevant studies that were not yet included (a list of concepts and keywords can be found in Appendix D).

All identified papers were transported into Mendeley for further investigation. Titles and abstracts were first analysed against the inclusion/exclusion criteria, from which a list of full text papers was compiled for further examination against the same criteria. After full-text articles were examined for eligibility by the first author (SJ), a second reviewer (LGM) went through a random sample of 25% of the papers, examined them for eligibility, and coded them based on which synthesis they would be included in.

4.2.3 Eligibility Criteria

Inclusion and exclusion criteria were defined as follows, using the PI(E)CO framework:

Population: Studies were included if the sample had a recorded history of violence and had been (a) convicted or charged with a violent offence (forensic) or (b) were in the care of the forensic mental health system as a result of their past offences or perceived risk of violence. This included samples of all ages, and with all diagnoses. Studies were excluded from the review if they focused primarily on fire-setting, as there is research to show that these populations may have different risk factors, included non-forensic sample(s), if not all members of the violent group or sample had committed violent contact offences *unless* they were in a secure forensic environment.

Exposure: Studies were included if they examined a relationship (association, prediction, comparison) between cognitive abilities and violence and physical aggression, and if neuropsychological assessment outcomes and cognitive components from risk assessments were included as predictor, comparison, or outcome variables. Studies were excluded if individual cognitive domains were not reported, for example, a composite score encompassing several domains.

Comparison: Studies were included if they compared a sample of violent offenders to a sample of non-violent offenders. Studies were excluded if they compared violent offenders to healthy controls, non-forensic samples, or other violent offenders.

Outcomes: Studies were included if outcome variables related to violence and physical aggression. Studies were excluded from this review if they did not examine the relationship between neuropsychological variables and violence or physical aggression, did not provide a clear definition and measurement of violence, aggression, or type of violence, and when a composite score of violence or aggression was used, and included self-injury or verbal aggression with no differentiation. Studies were also excluded if they did not provide quantitative outcomes.

Published and unpublished studies (including theses and dissertations) were also included in this review. Studies which were not published in English were excluded.

4.2.4 Definitions of Violence

There are various ways in which violence is defined within the literature in this field, and currently, there is no consensus on the elements that should be included, or what makes up a comprehensive explanation for such a complex construct. While the term ‘violence’ can encompass a wide range of behaviours ranging from shouting angrily at someone to murder, it is imperative that discretion is used when choosing a definition to fit the specifics of a study design. Further, definitions often vary according to which measurement technique is used, level of severity, and the context in which it is being measured (inpatient vs. community). Due to the many ways in which ‘violence’ can be operationalised, quantitative synthesis is often difficult, and can lead to ambiguity in our interpretation of the results. For the purposes of this review, violence has been described as, the purposeful use of physical strength or force to harm another individual/individuals, and committing sexual acts against another person without their consent, or against those unable to give consent, and more concretely defined as, behaviour involving “an intentional act of physical aggression against another individual that is likely to cause physical injury” (Meloy, 2006, p. 536). In an attempt to avoid heterogenic outcomes, acts of physical violence and aggression stood as the main focus of this review.

4.2.5 Study Selection and Data Extraction

Abstracts and titles were reviewed by the first author (SJ), and those which were considered as irrelevant were excluded. Full-text papers were reviewed against the inclusion criteria by the first author and a random sample of 25% of the full-text papers were reviewed by a second reviewer (author LGM). Papers included in this review were categorised based upon type of study (neuropsychological differences between violent and non-violent groups and neuropsychological correlates of violence). For each study, the following variables were extracted and coded (where available) and were recorded on a data extraction form: Author and year of publication or submission, sample size, country, setting, gender, age, study design, cognitive variables examined, measures used and type, and type of violence/outcome variable.

4.2.6 Quality Assessment

Due to the variety of research designs included in this review, two separate quality assessment measures were used, The Newcastle-Ottawa Scale (NOS; Wells et al., 2015), for cohort and cross-sectional studies, and the National Institute of Health (NIH) Quality Assessment tool for observational studies with no control group (National Institute of Health, 2014). Quality assessment was completed independently by one reviewer for all included studies, and a second reviewer for 25% of the papers. Disagreements were reviewed and discussed until a consensus was reached. In line with best practice, no papers were excluded based on quality.

4.2.7 Statistical Analyses

Meta-analyses were conducted, and forest plots were created using the m*etafor* (Viechtbauer, 2010) and *meta* (Balduzzi et al., 2019) packages in R: A language and environment for statistical Computing, Version 3.3.6 (R Core Team, 2020). Studies were grouped for meta-analysis according to the cognitive domain(s) that were reported. Cognitive domains that were reported were analysed separately to allow for identification of which cognitive domains might evidence greater impairments. Studies did not need to use the same assessment tool to be grouped, as they all had violent outcomes, and studies which measured the same cognitive ability with more than one tool (e.g., two impulsivity measures) were pooled to obtain an average effect size for the cognitive domain. Individual test scores were extracted in addition to indices scores where they were available. To improve the validity of results, analyses were only conducted on cognitive domains examined in three or more studies (Witt et al., 2013). Significance of pooled effects was determined by examining the 95% confidence intervals, where an effect was significant if the confidence intervals did not include 0.

All meta-analyses were conducted using random-effects models as there was likely to be heterogeneity due to the broad inclusion criteria. Heterogeneity of variance among studies was identified by Cochran's Q test for heterogeneity (Cochran, 1954) and its magnitude was assessed using the I^2 statistic, which describes the percentage of variance due to among-study factors. Although there is no universal rule of thumb for the interpretation of heterogeneity, Higgins and colleagues (2003), suggest that tentative values of low, moderate and high heterogeneity correspond to I^2 scores of 25%, 50% and 75% respectively. When the presence of heterogeneity was identified by the Q test (p = <.05), additional analyses were conducted to identify the source of heterogeneity by identifying outliers and studies influencing the pooled effect and heterogeneity using the *influence* function (Viechtbauer & Cheung, 2010) in the *metafor* programme in R (Viechtbauer, 2010) and subsequently performing sensitivity analyses (Higgins & Green, 2011). According to Higgins and Green (2011), it is best to run the analysis with and without potential outlying/influencing studies, reporting results for both (sensitivity analysis), however, it is rarely informative to produce forest plots for each, thus, forest plots were only created for all studies included and not to reflect the removal of outlying/influencing studies. Further, Higgins and Green (2011) advised that the exclusion of studies after sensitivity analysis is likely to introduce bias, therefore, no studies were excluded on this basis. Where studies were removed but effects remained similar to the original, results were considered robust, whereas if the effects differed, results were interpreted with caution (Aromataris & Munn, 2018).

Publication bias was measured when there were $k \geq 10$ studies for a domain, by generating funnel plots and running a regression test for funnel plot asymmetry where the model is a weighted regression with multiplicative dispersion and the predictor is the standard error (Higgins & Green, 2011). Models with a p-value <.05 suggest the presence of publication bias. The *trim.fill.rma* function in the *metafor* programme in R was used to estimate the number of missing studies and the true effect if the missing studies were included in the original analysis (Viechtbauer, 2010).

To examine cognitive profiles by type of offender, subgroup analyses were planned where there were sufficient data. However, it is suggested that subgroup analyses only be conducted in meta-analyses where there are $k = \geq 10$ studies per group (Higgins & Green, 2011). As there were not enough data to examine type of offender as a categorical moderator within each meta-analysis, papers were collated into subgroups and run as their own analysis where there were $k = \geq 3$ studies for each subgroup within a domain. Although this approach limits statistical interpretation within the wider analysis, it provides an estimate of the magnitude of the pooled effect for individual groups who may have differential cognitive profiles. Subgroups were based on the setting in which they were recruited (e.g., mentally ill offenders (MIO) from a forensic psychiatric hospital, court referrals recruited from court settings, and prisoners recruited from prisons). An additional subgroup analysis was performed on correlation and predictions studies which used a prospective design.

For between-group studies, data were presented in a variety of manners, including means and standard deviations, t-tests, and ANOVAS. As a central interest was looking at the differences between the groups, all data were transformed into Cohen's *d* for comparison. All variables were transformed into effect size *d* using an effect size calculator provided by the Campbell Collaboration (D. Wilson, n.d.). The following quantitative descriptors are used to define effect ranges for Cohen's *d* and *g*: small (.20); medium (.50); large (.80) (Cohen, 1988). In studies which examined prediction or association of violence, data were presented as correlations, odds ratios, means and standard deviations, and area under the curve (AUC).

Given the observational nature of most studies, all data were transformed into correlation coefficients for comparison. When data were presented as a correlation, they were not transformed, when data were presented as odds ratios, an online effect size calculator using excel was utilized and data were transformed into *d* using methods from Borenstein et al. (2009), and then transformed from *d* to *r* using the method of Rosenthal (1994). If they were reported as means or an AUC score, they were transformed into *d* using methods from Ruscio (2008) and then from *d* to *r* using Rosenthal (1994). The following qualitative descriptors are used to define effect ranges: minimal (<.10); small (.10-.29); medium (.30-.49); large (.50-.69); very large (.70+) (Cohen, 1988). Meta-analyses were conducted using *ZCOR* in *metaphor* as the measure in models, which allowed for raw correlations to be transformed into Fisher's *Z*-scores, and then transformed back into correlations for interpretation (Viechtbauer, 2010).

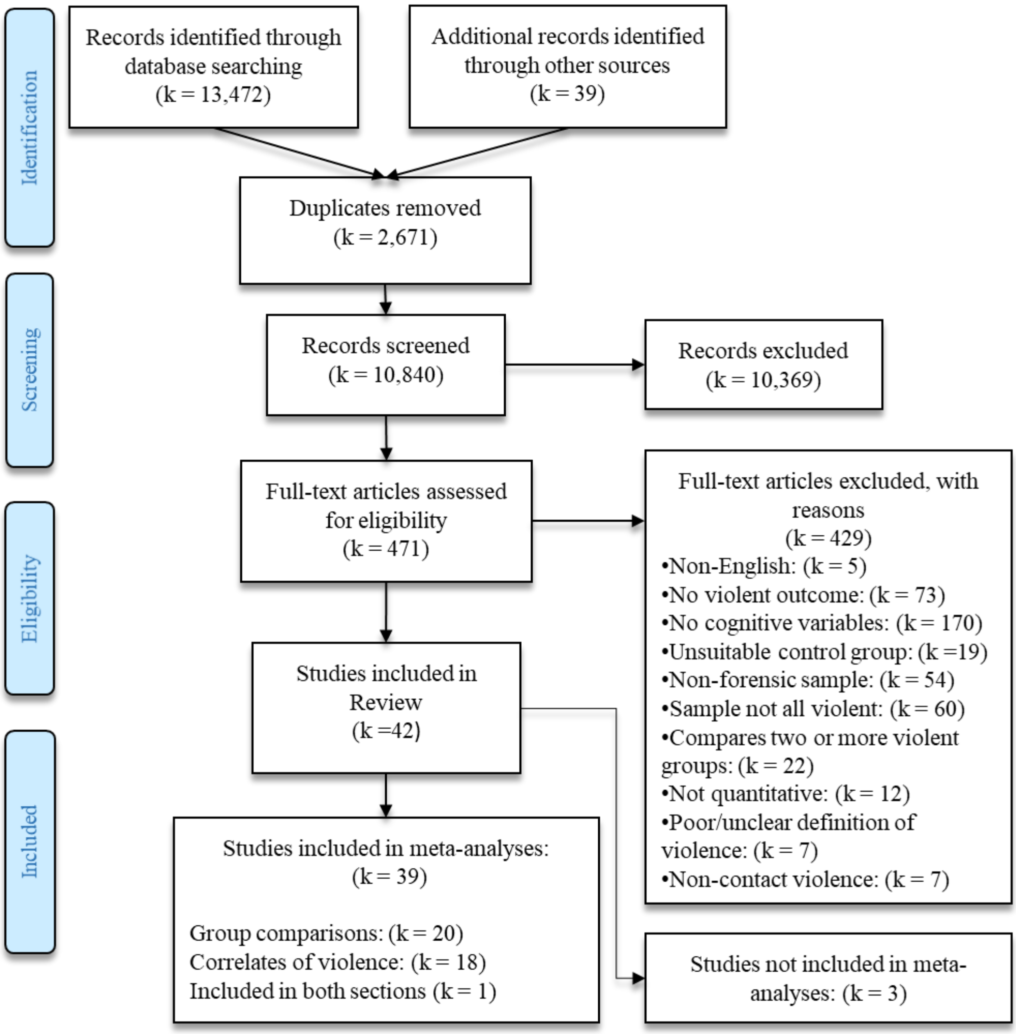
4.3 Results

4.3.1 Study Selection

A total of 13,511 titles and abstracts were screened based on the above inclusion and exclusion criteria, of which 13,040 (96.5%) were removed; leaving 471 full text papers to be screened. Following the removal of duplicates, 10,840 papers were removed based on their title and abstracts demonstrating no relevance to this review, primary reasons for exclusion for the remaining papers are reported in Figure 4.1. Forty-two studies were included in the review, 39 were included in the meta-analyses, and of the remaining three papers, one did not provide sufficient information to calculate an effect, and the remaining two reported on cognitive domains which did not have enough individual papers to include in a meta-analysis. There was 96% agreement (Kappa = 0.83, $p < .001$) on the inclusion and exclusion of papers between the primary and secondary reviewers. All discrepancies were discussed and resolved.

The meta-analytic results section of this review is broken into two parts. Part one will examine cognitive domains which differentiate violent and non-violent populations, and part two will highlight cognitive correlates of violence. In addition to this, as FSIQ is a composite score, the analysis of it was not reported as a main finding in this review, but rather, individual measures and index scores were reported. However, individual studies that fit the inclusion criteria, but only measured FSIQ were not excluded and are reviewed in an exploratory analysis before the discussion. A list of measures used in individual studies to ascertain cognitive functioning can be found in Appendix B.

Figure 4. 1 PRISMA Flow Chart



4.3.2 Cognitive Differences Between Violent and Non-Violent Offenders

Study Characteristics. Of the 42 included papers, there were 22 that met inclusion criteria for the comparison of violent and non-violent offenders. Characteristics of the included studies are presented in Table 4.1. Six of the papers were unpublished theses, and the remaining 16 were peer-reviewed. The publication or submission dates ranged from 1978 to 2019, where nearly half (45%) were published from 2000 onwards. Fourteen (64%) studies examined samples from the United States, and the remaining samples were from Canada, China, Germany, and the Netherlands. A total of 1,657 violent offenders were included in this analysis. Violent group sample sizes ranged from 20 to 343 participants, where the majority (81%) had n = <100, with an average sample size of n= 75. There was a total of 1,433 non-violent offenders, with sample sizes ranging from 14 to 369, and an average sample size of

n= 65. Participants' age ranged from 10 to 61 years old with mean age reported in only seven papers (violent and non-violent), where the average age of the violent offenders was 25.99 (7.74), and the non-violent offenders was 27.53 (8.71). Six studies reported the mean age of both groups combined, making the overall mean age of all violent and non-violent offenders 24.26 (8.56). Women made up approximately 5% of the participants for both violent and non-violent offenders, 12 studies had all male samples, six included a small number of women, and four did not report this. The majority of studies measured cognitive abilities using behavioural or performance measures (72%), 19% used self-report measures, and one study used a combination of self-report and behavioural measures.

Methodological Quality. The results of the quality assessments showed that 18 (81%) papers were considered to be of 'fair' quality, one was 'good' quality, and three were 'low quality'. A second rater assessed a random 25% of the papers, and there was a 100% agreement (Kappa= 1.00, p = .01) on the quality of the included papers. As the NOS utilizes stars as the scoring system, qualitative descriptors were utilized to categorize papers for ease of interpretation and explanation (e.g., high quality= 8-7 stars; fair quality= 6-5 stars; low quality= 4-3 stars; unacceptable quality= ≤ 2 stars). The most common items missing from papers were an adequate description of the sampling strategy, representativeness of the sample and controlling for confounding variables. All discrepancies between the two raters were discussed, and an agreement was reached for all.

Meta-analysis of Between Group Studies. For each domain, a positive pooled effect size denotes poorer performance of violent offenders relative to non-violent offenders. The presence of heterogeneity was explored further when there were enough papers to do so, and only one domain had a sufficient number of papers to measure publication bias. Results are reported below for each domain. Forest plots presenting pooled effect sizes and heterogeneity are shown in Figure 4.2 for all domains.

Verbal Comprehension. Violent offenders presented significantly lower scores on measures of verbal comprehension, with a small effect size, relative to non-violent offenders (k = 7; violent n= 347; non-violent n= 344, g= 0.35 [0.04, 0.66], I^2=72%). There was heterogeneity within this group of papers which was explored further by identifying outliers and subsequently performing sensitivity analyses (Higgins & Green, 2011). Analyses revealed that there were no statistical outliers or influential studies, thus three studies which did not use WAIS or WISC scales to measure verbal comprehension were removed (e.g., Bryant et al., 1984; Kennedy et al., 2011; Ullman, 1989), resulting in a large decrease in heterogeneity (I^2=

17%). The effect size decreased and remained significant (*k* = 4; *n* violent = 186; *n* non-violent = 229; *g*= 0.26, [0.02, 0.50]).

Examination of the forest plot revealed that only Bryant et al. (1989), Kennedy et al. (2011), and Brimigion (2014) had significant and moderate to large effect sizes. Bryant et al. (1984) found the largest effect size (*g*= 0.84) in a sample of adult prisoners, with no exclusionary criteria, though a proportion of the violent group and of the non-violent group had learning disabilities. They also assessed the participants in their sample using the entire Luria Nebraska Neuropsychological Battery (LNNB), as opposed to using archived data. Like Bryant et al, Brimigion (2014) also recruited prisoners, though they were juveniles rather than adults, and the study was based on archived data, had no exclusion criteria, and Wechsler scales were administered to measure the construct (e.g., verbal comprehension index; VCI), whereas Kennedy et al. (2011) also recruited juveniles, used archived data, only excluded individuals without completed testing profiles, and administered the Peabody Picture Vocabulary Test (PPVT-III) to assess the domain. A limitation of all these studies is their small sample sizes, and both the Kennedy et al. and the Brimigion papers had vastly different between group sizes which can affect statistical power and Type 1 error rates (Rusticus & Lovato, 2014). For example, Kennedy et al, included *n*= 64 violent offenders, and less than half non-violent offenders (*n*= 31), and Brimigion included *n*= 23 violent offenders, and almost five times as many non-violent offenders (*n*= 112), whereas Bryant et al recruited equal sized groups. Moreover, only Brimigion matched the samples by age, socioeconomic status, and sex, though the remaining papers ran analyses to identify significant characteristic differences between groups. Tarter et al. (1983) recruited juvenile offenders with mental illness through the courts, but excluded individuals with psychosis, brain trauma, and abnormal brain scans, which may explain the small effect sizes reported in their study, as brain injuries, such as traumatic brain injuries (TBI) have been linked to cognitive difficulties (Ponsford et al., 2008), and have an overlapping relationship with mental illness and increased aggression (Brown et al., 2019).

***Reasoning*.** Four included studies with *n* = 186 violent offenders and *n* = 229 non-violent offenders measured reasoning, revealing a significant, small effect size and no heterogeneity (*g*= 0.26 [0.05, 0.48], I^2=0%). All included studies assessed reasoning using a Wechsler scale, and despite the significant pooled effect size, no individual studies demonstrated a significant effect. Moreover, the confidence intervals are wide suggesting uncertainty in the true effect.

Attention. Attention scores were reported in six studies, resulting in a small, significant effect size in the expected direction (*n* violent = 299; *n* non-violent = 187**,** *g*= 0.33 [0.05, 0.62], I^2=52%), however, heterogeneity was evident. Although heterogeneity was likely due to attention being analysed as a single construct in this review, as there was not enough data to look at the different types separately (e.g., sustained, selective, auditory), contributors to heterogeneity were investigated. The removal of two statistical outliers (e.g., Bryant et al., 1984; Tarter et al., 1983) decreased heterogeneity (I^2=0%), and the difference between the performance of violent and non-violent offenders on measures of attention remained significant (*k* = 4; *n* violent = 213; *n* non-violent =104; *g*= 0.30, [0.05, 0.54]), and only decreased slightly, indicating that the true effect is likely within this range. However, Bryant et al. (1984) was the only paper to contribute a large, significant effect size to this analysis, and it was much larger than the remaining papers (*g*= 0.83); additionally, the effect size for Tarter et al. (1983) was trending in the opposite direction. Taking these findings in conjunction with the significant heterogeneity observed before the removal of outliers, and the spread of the 95% confidence intervals, more research is required to determine the true between group effect for attention.

Impulsivity. Violent offenders had significantly higher scores on measures of impulsivity indicating greater impairment relative to non-violent offenders (*k*= 4; *n* violent = 117; *n* non-violent = 104), and the data were homogenous, revealing a small effect size (*g*= 0.28 [0.10, 0.47], I^2=0%). Only Zhou et al (2014) found a significant effect, however, the only notable differences between this and the other studies were the inclusion of juvenile participants aged 15 to 17 years, and the magnitude of difference between those in the violent group (*n*= 236) and those in the non-violent group (*n*= 87). Comparatively, the remaining studies under this domain had sample sizes ranging from *n*=32 to 43 in violent groups, and *n*= 15 to 75 in the non-violent group, thus, the violent offending group in the Zhou et al had more than double the sample size than the other three studies had combined, which increased the likelihood of finding a significant effect. Three out of the four studies in this domain used the BIS-11 to measure impulsivity in prisoners resulting in effects of similar magnitude, ranging from *d*= 0.30-0.33, whereas the fourth study measured impulsivity in probationers with the Impulsiveness Scale- Short Form and found a much smaller effect size (*d*= 0.19).

Expressive Speech. Violent offenders performed significantly poorer on measures of expressive speech (*k* = 3; violent *n* = 117; non-violent *n* = 104), resulting in a large effect size and no significant heterogeneity (*g*= 0.85[0.53, 1.18], I^2=22%). Both Ullman (1989) and Bryant (1984) contributed significant, large effects to the forest plot (*g*= 0.88, *g*= 1.05, respectively), whereas Rimmer (1999) found a much smaller, non-significant effect (*g*= 0.42). Notably, both Ullman and Bryant employed the LNNB to measure expressive speech, whereas Rimmer employed the Learning-Verbal Scale, which may explain the difference in magnitude in effect sizes.

***Non-Significant Findings*.** No significant differences were found between violent and non-violent offenders on measures of memory, examined as a construct, as there was not enough data to look at the different types separately (e.g., long/short-term, verbal, episodic, semantic; *k*= 4; violent *n*= 132; non-violent *n*= 126, g= 0.21, [-0.05, 0.47]), working memory (*k*= 5; violent *n*= 264; non-violent *n*= 309, g= 0.35, [-0.03, 0.73]), processing speed (*k*= 4; violent *n*= 196; non-violent *n*= 225, *g*= 0.20, [-0.02, 0.42]), response inhibition (*k*= 3; violent *n*= 186; non-violent *n*= 128, *g*= 0.20, [-0.03, 0.44]), cognitive flexibility (*k*= 4; violent *n*= 184; non-violent *n*= 166, *g*= 0.19, [-0.02, 0.41]), planning (*k*= 6; violent *n*= 264; non-violent *n*= 235, *g*= 0.12, [-0.11, 0.36]), and motor skills (*k*= 4; violent *n*= 157; non-violent *n*= 154, *g*= 0.18, [-0.78, 0.41]).

Table 4. 1 Characteristics of Between Group Studies

Study	Sample Size		Study Characteristics			QA	Cognitive Variables Assessed
First Author (year)	Violent	Non	Country	Setting	Violence		
Bock (2014)	343	369	Germany	Prison	Armed robbery, manslaughter, murder, and bodily injury	********	IQ, empathy
Brimigion (2014)^	23	112	USA	Res.youth centre	Murder, forcible rape, robbery, assault	******	IQ
Bryant (1984)	55	55	USA	Prison	Assaultive crimes against persons	*****	EF
Busch (1990)	71	71	USA	Juvenile Detention	Convicted of homicide	******	IQ, EF
Chan (2012)	34	75	China	Probation	Robbery, serious assaults, wounding	******	Impulsivity
Cornell (1992)	72	77	USA	Prison	Convicted of homicide; serious assault	****	IQ
Duwors (1998)^	32	15	USA	Prison	Incarcerated for a violent offence	*****	Impulsivity
Edwards (2003)	43	40	USA	Prison	Conviction of spouse abuse	*****	Impulsivity
Feichtinger (2007)^	42	56	USA	Prison	Assault, use of weapons, sexual assault, murder, manslaughter, and impaired driving causing harm or death	******	IQ, EF
Goldstein (2001)	66	112	USA	Prison	Past conviction of violent crime & charges of a violent crime	*****	Empathy
Greenfield (2007)	20	19	Canada	Prison	Violent offenders (against person)	*****	IQ, planning
Gretton (1998)^	107	50	Canada	Youth Court/Inpatient Unit	Murder, manslaughter, attempted murder, assault, sexual crimes, robbery, kidnapping, possession of a weapon, and arson	******	IQ
Hays (1978)	25	39	USA	Juvenile Court/Probation	Committing one more murder	***	IQ
Kennedy (2011)	64	31	USA	Juvenile Court	At least one violent arrest	******	Language, inhibition
Kuin (2017)	71	14	Netherlands	Prison	Convictions for assault and battery, manslaughter, murder, sex offenses or arson with risk for persons.	*****	IQ, cognitive flexibility, drawing conclusions
Meijers (2017)	85	45	Netherlands	Prison	Murder, arson, rape, and serious violence	******	IQ, planning, WM, response memory, RI, set shifting, attention, risk taking, decision making
Rimmer (1998)^	20	20	USA	Youth Detention	Assault, aggravated assault, attempted murder, and/or murder	******	EF, verbal learning, problem solving
Tarter (1983)	31	28	USA	Juvenile Court	Inflicting personal injury of a nonsexual nature.	******	IQ, EF, memory, language, learning
Ullman (1989)^	42	29	USA	Prison	History of two or more assaultive charges; murder, manslaughter, forcible sexual assault, assault, kidnapping and robbery.	*******	CF, expressive, speech, language, visuospatial abilities, memory
Umbach (2019)	114	71	USA	Prison	Violent offences	******	Risky decision making, IQ
Umbrases (2018)	61	18	USA	Military Prison	Violent offences (murder and rape)	*****	IQ, PRI, VCI
Zhou (2014)	236	87	China	Juvenile Detention	Convicted of homicide, assault, rape, robbery, and affray.	*****	Impulsivity

Note. This table presents first authors and characteristics of included studies.; ^indicates a thesis; *IQ*= intelligence; *EF*= executive functions; *WM*= working memory; *VCI*= verbal comprehension index; *PRI*= perceptual reasoning index; *RI*= response inhibition; *CF*= cognitive flexibility; *QA*= Quality Assessment.

Figure 4. 2 Forest Plots for Between Group Studies

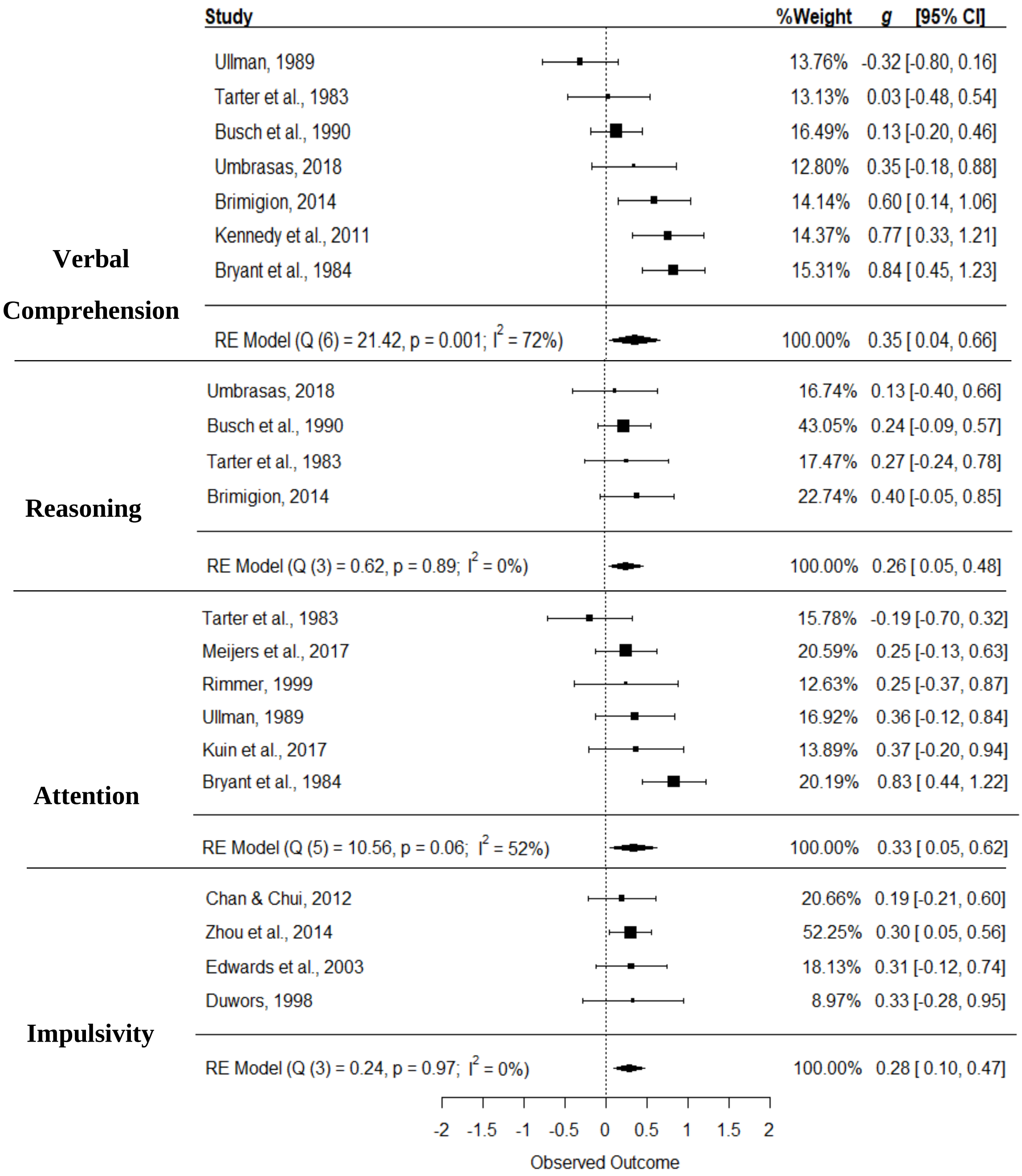


Figure 4.2 continued Forest Plots for Between Group Studies

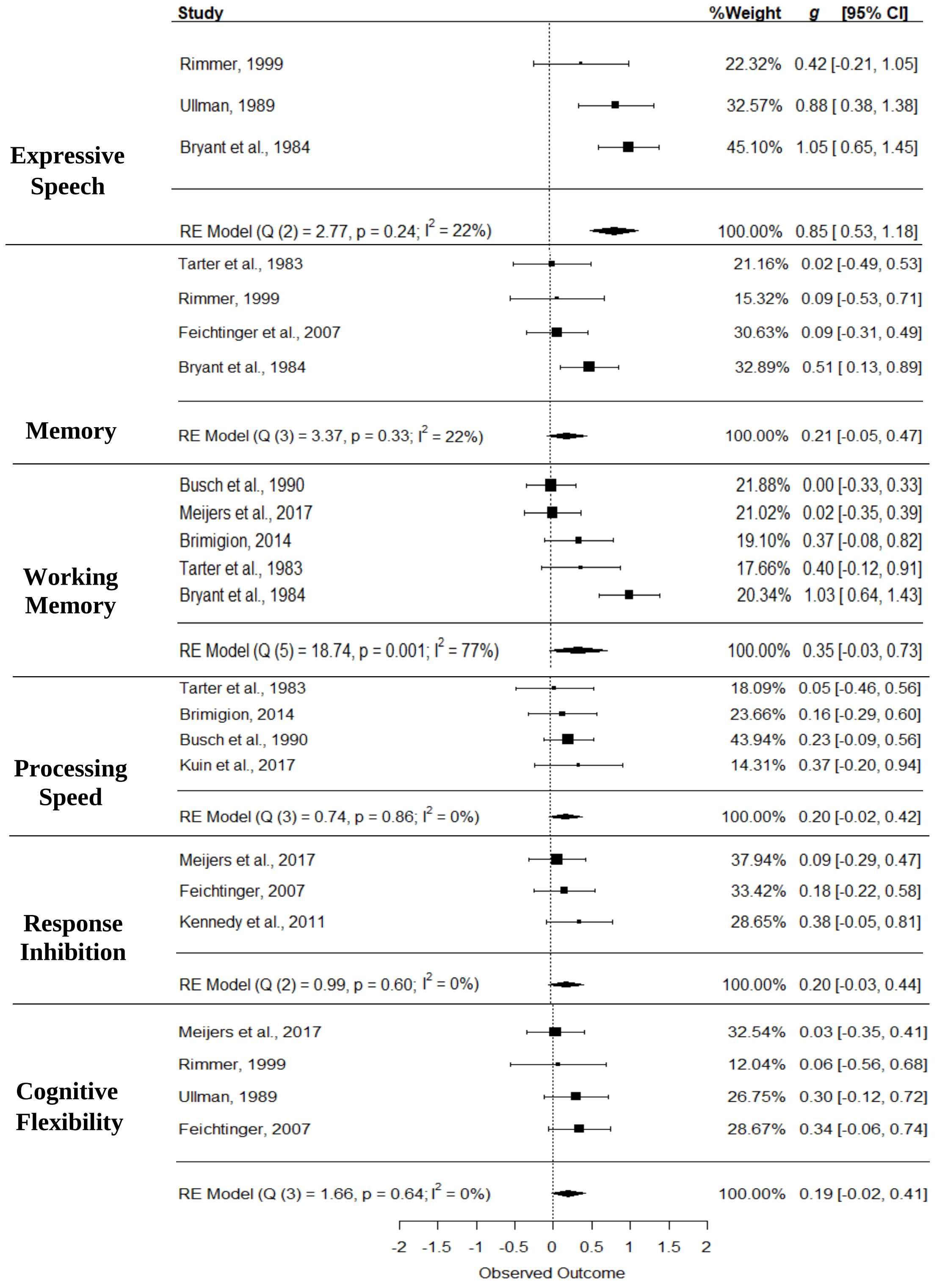
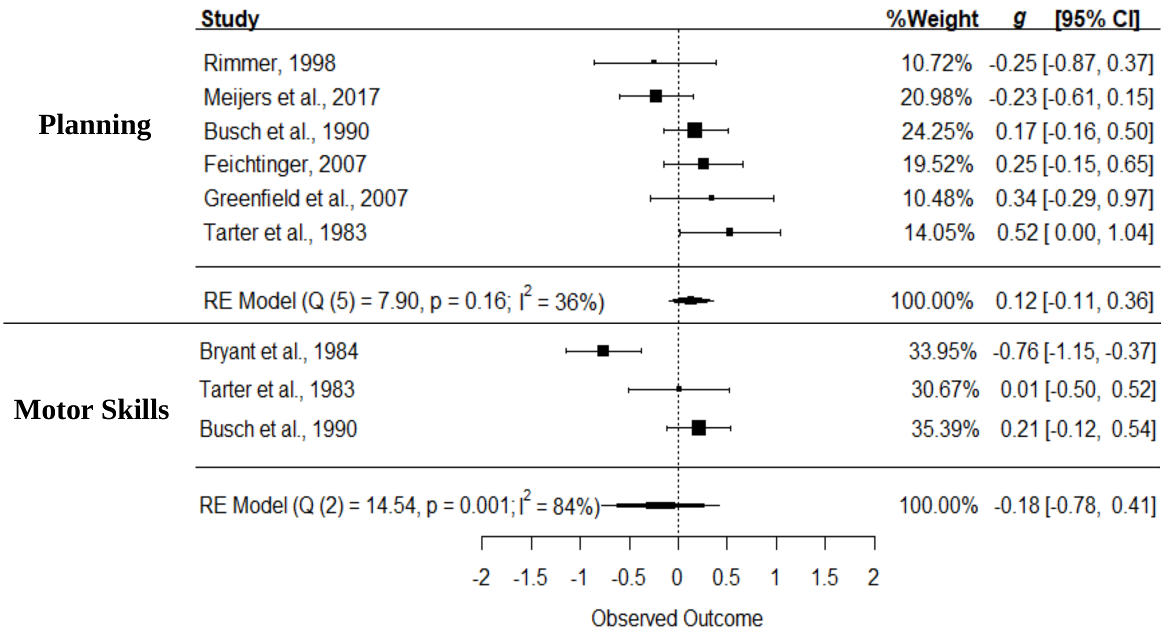


Figure 4.2 continued Forest Plots for Between Group Studies



Note. Positive effect sizes reflect poorer performance of violent offenders on measures; *RE*= Random Effects Model; *Q*= Cochran's *Q* test for heterogeneity; *g* = effect size; *CI* = 95% confidence interval of effect size *g*; I^2 = proportion of dispersion due to variability between studies.

Additional Domains. There were 12 domains that did not have a sufficient number of studies to conduct meta-analyses; language, visuo-spatial abilities, and risk taking were reported in *k*= 2 studies each, and empathy, decision making, drawing conclusions, verbal learning, empathy, facial emotion recognition, and four domains from the Luria-Nebraska Neuropsychological Battery (LNNB) including writing, reading, tactile, and visual were reported in *k*= 1 study.

Profiles of Offenders by Subgroup. Subgroup analyses were only performed on subgroups within domains that had $k = \geq 3$ studies, thus there was only enough data to examine verbal comprehension, working memory, reasoning, attention, impulsivity, and cognitive flexibility in the 'prisoner' subgroup, however, there was not enough data to examine 'MIO' or court referrals as their own subgroups. As a subgroup there were significant differences between violent and non-violent prisoners on measures of reasoning (*k*=3, *n* violent = 155, *n* non-violent= 201, *g*= 0.26 [0.02, 0.50], I^2=0%), attention (*k* =4; *n* violent = 213, *n* non-violent= 104, *g*= 0.30 [0.05, 0.54], I^2=0%), and impulsivity (*k* =3; *n* violent = 311; *n* non-violent= 136, *g*= 0.30 [0.09, 0.51], I^2=0%). All domains reflect higher impairments in the violent groups relative to the non-violent groups. No significant effects were detected on the remaining domains.

Exploratory Analysis

Intelligence. As stated in the methods section of this review, intelligence is often reported as a composite score giving little insight into the specific abilities which contribute to the score, and thus it was not part of the main analyses in this review. However, there have been debates on whether intelligence is a protective or risk factor for violence. For example, historical studies suggested that low IQ was related to conduct disorder and antisocial behaviours and postulated that this may in part be due to associated cognitive impairments (Moffit, 1993), and Maguin and Loeber (1996) also agreed that low IQ played a part in offending behaviour and stated that antisocial behaviour is often accompanied by academic problems. More recent research has conceptualised high intelligence as a protective factor against violence, suggesting that it can help to override problems often seen in aggressive individuals, such as compensating for having a disadvantaged background in education and job attainment (Damian et al., 2015; see Ttofi et al., 2016 for a review).

As a result of these suppositions, studies that reported information on the intelligence of offenders were further investigated. Intelligence was examined in k= 13 studies (violent n= 971, non-violent n= 918) and produced a small significant effect size (g= 0.15 [0.02, 0.27]) with 25% between study heterogeneity. However, after controlling for publication bias, the trim-and-fill analysis (Appendix C) estimated that three studies were missing, resulting in a small, non-significant effect (g= 0.09 [-0.04, 0.22]). Under the intelligence domain, only two papers reported significant differences between groups with moderate effect sizes (e.g., Brimigion, 2014, g= 0.54 [0.09, 0.99] and Hays et al., 1978, g= 0.53 [0.02, 1.04]), both assessing intelligence with a Wechsler scale, however the remaining six studies that also used a Wechsler scale found small, non-significant effects (e.g., g= -0.15 to 0.44). Studies that used other measures to assess intelligence, including the Test of Non-Verbal Intelligence in prisoners (e.g., Greenfield et al., 2007, g= 0.44 [-0.20, 1.08]), Ravens Standard Progressive Matrices in prisoners and juveniles in a detention center (e.g., Kuin et al., 2017, g= 0.07 [-0.50, 0.65]; Rimmer, 1998, g= 0.06 [-0.09, 0.21]), and the Vienna Matrices Test in prisoners (e.g., Bock et al., 2014, g= 0.06 [-0.56, 0.68]) failed to significantly differentiate the groups. one explanation for these findings may be in the methodology of the studies, as the Hays et al. (1978) study compared juvenile murderers to status offenders, operationalised as non-violent offences that only juveniles would be charged for, such as truancy, which likely inflated effect sizes. Brimigion (2014) had the largest effect size, though it was only one point higher than the effect that Hays et al found, but was of better quality, as they matched their groups by age, socioeconomic status, and sex, though both studies included juvenile samples. There were four other studies which used a Wechsler Scale to measure intelligence in juvenile offenders (e.g., Busch et al., 1990, *g*= 0.30 [-0.03, 0.63]; Tarter et al., 1983, *g*= 0.43 [-0.08, 0.94]; Gretton, 1998, *g*= -0.15 [-0.48, 0.18]; and Cornell & Wilson, 1992, *g*= -0.11 [-0.43, 0.21]), however, effect sizes were small, and their findings were not significant. When FSIQ was disaggregated into verbal IQ and performance IQ, neither produced significant pooled effect sizes. Though, when these domains were further broken down into indices and subtests, pooled effects for verbal comprehension and reasoning significantly differentiated the groups (e.g., *g*= 0.35; *g*= 0.26, respectively) though working memory and processing speed tasks did not (Figure 4.2). These findings support the notion that the assessment of disaggregated cognitive functions may be necessary in this population, however, the magnitude of the effect sizes suggests that the intellectual abilities between violent and non-violent offenders is of a negligible magnitude.

4.3.3 Cognitive Correlates of Violent Offending

Characteristics of Studies. There were 21 studies that fit the inclusion criteria for identifying the association between cognitive abilities or impairments and violence. Characteristics of the 21 individual studies are presented in Table 4.2. The publication or submission date of included papers ranged from 1992 to 2018, where over half (76%) were published from 2008 onwards. Two of the included studies were unpublished theses, and the remaining 19 were published, peer reviewed papers. Four (19%) of the included studies examined samples from the United States, and the remaining samples were from the United Kingdom, the Netherlands, Canada, Ireland, New Zealand, Kosovo, Finland, and Germany. These studies provided a total of 2,377 offenders for inclusion in this synthesis. Sample sizes ranged from 41 to 409, with the average sample size being approximately *n* = 113. Within the included studies, the age ranged from 15 years old to 74 years old, only 17 papers reported mean ages, where the average age of the offenders was 35.95 (6.93). Most participants were men, with women making up approximately 7% of offenders. The majority of studies followed a prospective design (55%), and the remaining studies followed a correlational (25%) or retrospective design (20%). Types of measures used in this section varied, 40% of studies used a behavioural or performance measure to ascertain cognitive functioning, 35% used self-report measures, 15% used risk assessments, and 10% used a clinical rating scale.

Table 4. 2 Characteristics of Correlation and Prediction Studies

Study	Sample Size	Study Characteristics			QA	Cognitive domains assessed
First Author (year)		Country	Setting	Violence Measure		
Abidin (2013)	100	Ireland	Forensic Hospital	Actual, attempted, threatened harm to others	Fair	Impulsivity, insight, attention
Alia-Klein (2007)	60	USA	Forensic Hospital	Violence Assessment Scale	Low	Insight
Bass (2010)	45	Canada	Forensic Patients	Frequency of seclusions	Low	Decision making
Beggs (2008)	216	New Zealand	Prison	Violent recidivism	Fair	IQ
Belfrage (2000)	41	Sweden	Prison	Institutional Violence	Fair	Impulsivity, insight
Brugman (2016)	69	Netherlands	Forensic Patients	MOAS (rated for separate levels of severity)	Fair	Attn. bias, emotion recognition, RI
Coid (2015)	409	UK	Med. Secure services/ Prisoners/ community	MacArthur Community Violence Scale	Good	Insight, impulsivity
De Vogel (2006)	127	Netherlands	Forensic Hospital	Physical Violence	Fair	Insight, impulsivity
Dejong (1992)	248	Finland	Court referred	Violent Recidivism	Low	IQ
Edwards (2003)	44	USA	Prison	CTS Severe physical Male Violence	Low	Impulsivity
Fullam (2008)	82	England	Med/High Forensic Hospitals	"The individual was the clear instigator or co-aggressor, and if the incident involved physical aggression to staff, in-patients or property."	Fair	IQ
Howard (2014)	100	England	High Secure Hospitals	Violent Index of modified version of GRS	Fair	Premeditation, urgency, SS, persev.
Lodewijks (2008)	66	Netherlands	Juvenile Corrections	Physical violence against persons (Incident Files)	Fair	Risk Taking/ impulsivity, empathy
McDermott (2008)	108	USA	State Hospital	Inpatient physical aggression	Fair	Impulsivity
McKee (2004)*	111	UK	Forensic Patients	Violent Convictions	Low	Impulsivity
Moulden (2009)*	122	Canada	Prison	Self-reported Aggression Questionnaire	Low	Empathy
Nazmie (2013)	65	Kosovo	High Secure Forensic	Inpatient violent behaviour/ Violent recidivism	Fair	IQ, CF, PS, visual attention
Nigel (2018)	158	Germany	Forensic Patients	Violent offenses with penal consequences	Fair	Empathy
O'Reilly (2015)	89	Ireland	Forensic Hospital	An individual was classified as violent if they were the clear instigator or co-aggressor, and if the incident involved harm to staff or other patients.	Good	Attn., PS, WM, social cognition, reasoning, visual & verbal learning
Smith (2013)	73	USA	Forensic hospital	Inpatient Aggression	Fair	Self-centred impulsivity
Tonnaer (2016)	44	Netherlands	Forensic Psych. Centre	Self- reported Reactive-Proactive Aggression Questionnaire	Fair	Impulsivity, WM, SS, RI, divided & flexible Attn.

Note. The table presents characteristics of included studies; *indicates a thesis; *IQ*= intelligence; *RI*= response inhibition; *WM*= working memory; *QA*= Quality Assessment; *Med*.= Medium; *Persev.*= Perseveration; *Attn.* = attention; *CF*= Cognitive flexibility; *SS*= Sensation seeking; *Psych.*= psychiatric; *MOAS*= Modified Overt Aggression Scale; CTS= Conflict Tactics Scale; GRS= Gunn Robertson Scale

Methodological Quality. The results of the quality assessments showed that 14 were considered to be of 'fair' quality, 2 were 'good' quality, and 5 were 'low quality'. The most common reasons for lower quality assessments were not justifying sample size, the use of retrospective and cross-sectional designs, and not reporting number of loss to follow-up or participation rates. A second rater assessed a random 25% of the papers, and there was a 100% agreement (Kappa= 1.00, p = .02) on the quality of the included papers.

Meta-Analysis of Correlation and Prediction Studies. Forest plots for the pooled effect sizes and heterogeneity are shown in Figure 4.3 for all domains.

Impulsivity. Impulsivity was significantly associated with violent outcomes (k= 11, n = 1,221) with a small, significant correlation and no significant heterogeneity (r = .26 [.19, .33]). There was evidence of publication bias within this set of papers as evidenced by the trim-and-fill funnel plot (Appendix C). The trim-and-fill analysis estimated that there were 4 missing studies, and that the true effect after controlling for publication bias is r = .22 [.14, .28], I^2= 38%, thus the reported effect size of r = .26 may be slightly inflated and should be interpreted cautiously. The forest plot indicated that the studies by McDermott et al (2008) and Smith et al (2013) both had small and non-significant effects (r = .11 [-.08, .29], r = .15 [-.08, .37]), respectively, though they were both rated to be of fair quality, and used self-report measures of impulsivity in mentally ill offenders. The McDermott et al study assessed impulsivity using the BIS-11, a self-report measure, which was similar to one other study (Edwards et al., 2003, r = .33 [.04, .58]), however, McDermott et al did not report the total score for the measure, thus only the cognitive impulsivity score was used in the meta-analysis, which likely reflected a smaller effect than the total would have. Smith et al used the Psychopathic Personality Inventory (PPI), also a self-report measure, which assessed self-centred impulsivity. Additionally, both McDermott et al and Smith et al defined their outcome the same (e.g., physical aggression), and operationalised it using the same coding scheme, which may have also contributed to their non-significant findings, as it was specific to one behaviour. The study with the largest effect (e.g., Belfrage et al., 2000, r = .46 [.18, .67],) had the smallest sample out of the papers, measured impulsivity using the HCR-20, and defined their outcome as *institutional violence*, operationalised as assault of staff or other prisoners, severe damage to prison cells, and severe threats to staff, an operationalisation that is much broader than that employed by the McDermott et al and Smith et al studies, which likely contributed to the significant finding.

Post-hoc exploratory analyses were performed to investigate the use of risk assessments as a measure of impulsivity relative to behavioural and self-report measures.

Interestingly, the use of risk assessments alone produced a moderate effect (*r*= .31 [.21, .61]), but with moderate heterogeneity (I^2= 53.8%), and when this was further broken down to examine the HCR-20 on its own (e.g., independent of other risk assessments), it produced a small effect (*r*= 0.29 [.16, .42]), with moderate heterogeneity (I^2= 62%). However, when self-report (e.g., BIS-11, UPPS Impulsivity Scale, PPI), clinician rated (e.g., PANSS), and behavioural measures (e.g., Matching Familiar Figures Test, and Wisconsin Card Sorting Test-Perseverative errors) were examined as a subgroup, they produced a small effect size (*r*= .24 [.16, .31]) with no heterogeneity.

Attention. For violent offenders, impairments in attention were significantly correlated with violence, and data were homogenous (*k*= 4; *n*= 298, *r* = -.13 [-.24, -.01]). Notedly, the effect size reflects attention as a single construct, as there was insufficient data to analyse types of attention separately. Of the studies that examined attention, the only one to contribute a small, significant effect (e.g., Abidin et al., 2013, *r* = -.22 [.40, -.02]), also had the largest sample size and used the PANSS to measure the domain, which is not a valid measure of attention, but is based on a structured interview. All the studies under this domain included mentally ill offenders, and three out of the four studies measured the outcome prospectively (Tonnaer et al., 2016 employed a correlational design). Moreover, the two studies with the smallest effects (e.g., Tonnaer et al., 2016, *r* = -.10 [-.39, .20], and Nazmie et al., 2013, *r* = .01 [-.23, .25]), were the only studies under this domain to have specific exclusion criteria. For example, the Nazmie et al study excluded any patients with a history of organic brain syndrome, head injury, and intellectual disabilities, and similarly, Tonnaer et al excluded patients with psychotic disorders, and those with an IQ score lower than 80, whereas the patients in the other studies had diagnoses of schizophrenia spectrum disorders, other psychotic disorders, and intellectual disabilities. As it is well known that schizophrenia and psychotic illness are often characterised by cognitive impairments (Sheffield et al., 2018), it is likely that controlling for these factors decreased the overall effect that cognitive impairments had on violent outcomes, like the findings in the between group studies. This finding reiterates the notion that the assessment of cognitive impairments in relation to violence risk may be more important for offenders with risk factors for cognitive impairment.

Insight. Decreased insight in violent offenders was significantly associated with increased violence as evidenced by the random effects model (*k* = 5; *n*= 734, *r* = -.25 [-.32, -.18]). Insight typically involves three parts, "an awareness of having an illness, an attribution of recognisable symptoms of that illness and an appreciation for the need of treatment" (Ekinci & Ekinci, 2013, p. 116). Insight was however examined as a single construct and there were insufficient data to consider dimensions separately, or to conduct subgroup analyses. The forest plot showed that two studies did not to find a significant effect Belfrage et al. 2000, (r = -.20 [-.48, .11]), which was the only study under this domain to recruit prisoners, and Abidin et al., 2013, (r = -0.18 [-.37, .02]), though they used risk assessments and the PANSS to measure the construct. Interestingly, the largest effect was found by Alia-Klein et al (2007) (r = -.37 [-.57, -.13], which was the only study to use a measure designed to assess insight specifically (e.g., SUMD), however, the quality of the paper was low, as the outcome was collected retrospectively.

Non-Significant Findings. No significant correlation was found between empathy scores of violent offenders and violent outcomes (k = 3; n= 346, r = -.15 [-.33, .05]). There are two main types of empathy, affective and cognitive. Affective empathy is the ability to both understand another's mental or emotional state, and to share that emotional experience, while cognitive empathy is the ability to recognise another's feelings (not necessarily taking their perspective, which is how it differs from first order theory of mind) (Rankin et al., 2005). This analysis was conducted combining empathy into a single construct which likely contributed to the presence of significant heterogeneity. Likewise, no significant effects were found between measures of cognitive flexibility and violence (k= 3, n= 185, r = -.08 [-.22, .07]). Under the empathy domain, only Lodewijks et al. (2008) found a significant correlation with violence, and they were the only study out of the three to measure their outcome prospectively, though they used a risk assessment (e.g., SAVRY), whereas the other two studies were correlational and retrospective, and both used self-report measures. Moreover, Lodewijks et al was the only paper within this domain which examined a juvenile sample.

Additional Domains. There were 17 domains where there were insufficient studies to conduct meta-analyses; vocabulary, perseverance, verbal learning, visual learning, social cognition, aspects of memory, planning, disinhibition, attentional bias for aggression and threats, automatic associations with violence, vigilance, emotion recognition, and risky decision making were all reported in k= 1 study each, and sensation seeking, reasoning, working memory, and processing speed were examined in k= 2 studies each.

Cognitive Profiles of Subgroups. Subgroup analyses were only conducted on subgroups within domains that had k= ≥ 3 studies, thus there was only enough data to examine impulsivity and insight in ‘MIO’, and impulsivity in prisoners, further, all the included papers which examined attention included only MIO samples. There was insufficient data to conduct subgroup analyses for court referrals on any domain. A significant effect was found for the MIO group on measures of impulsivity with a small effect size, and low heterogeneity (k = 8;

n = 1,071; r= .23, [.16, .29], I^2= 21%), and insight (k = 3; n = 693; r= -.25, [-.32, -.18], I^2= 0%). There was a significant association between prisoners and violence on measures of impulsivity with a medium effect k = 3; n = 150; r= .38, [.23, .51], I^2= 0%).

Figure 4. 3 Forest Plots for Correlates of Violence

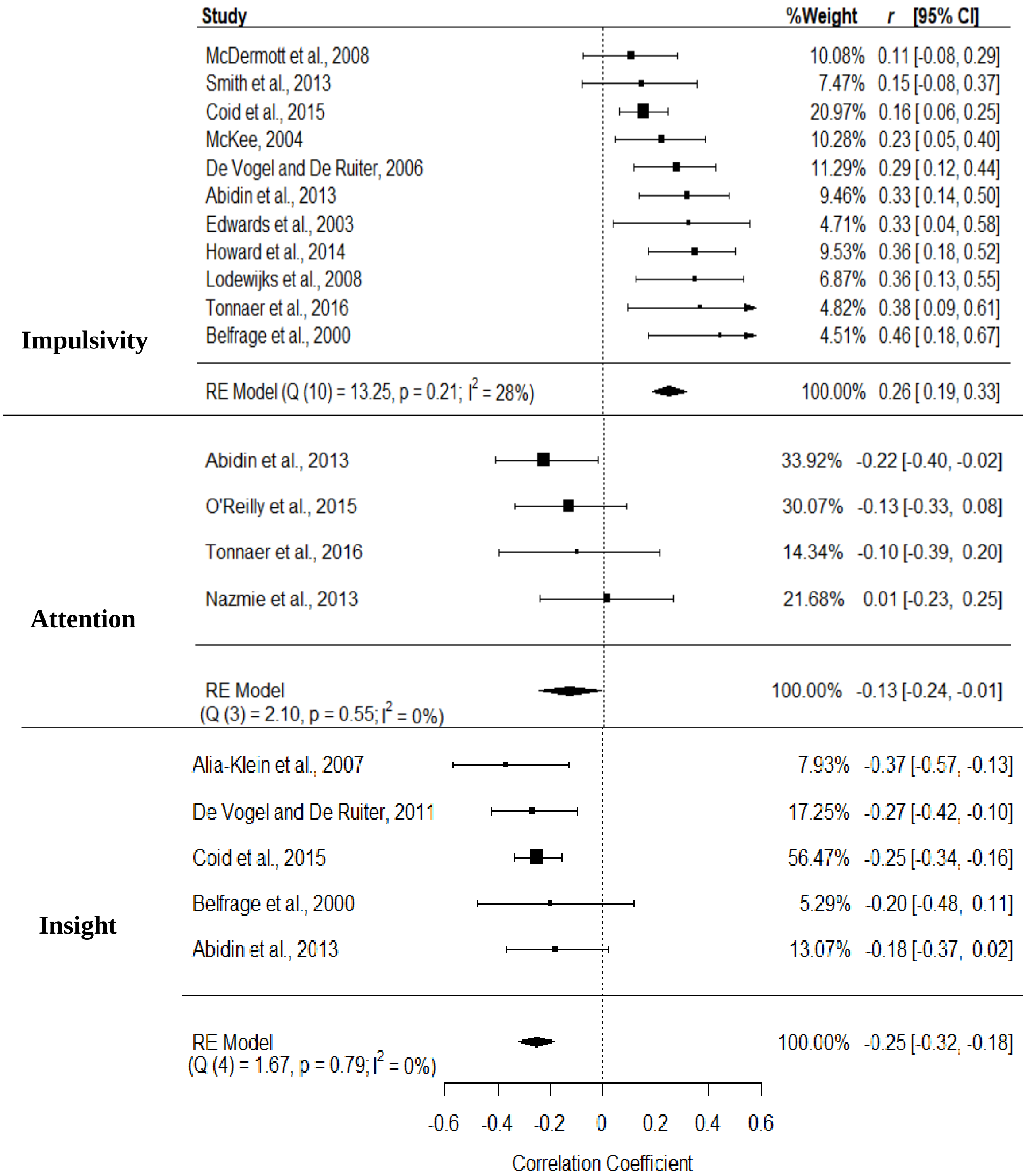
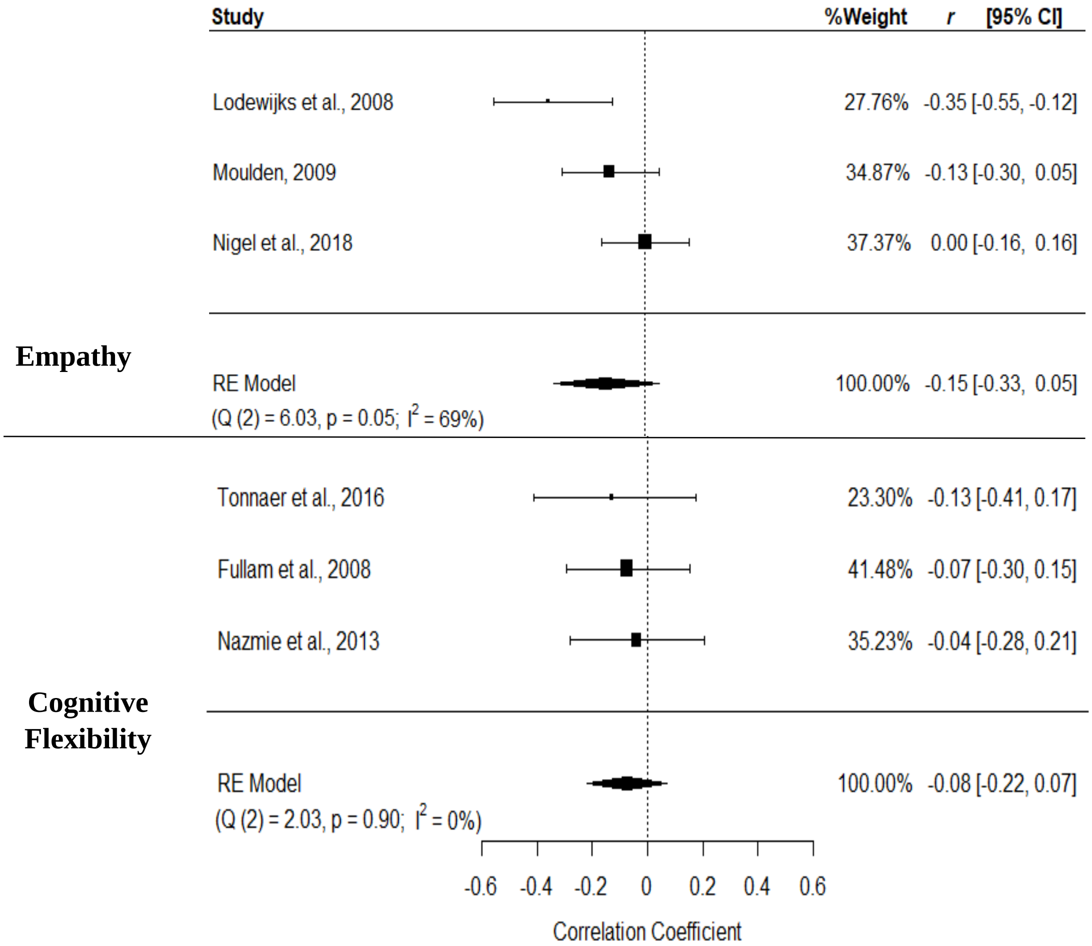


Figure 4.3 Continued. Forest Plots for Correlates of Violence



Note. Negative effect sizes reflect poor performance on measures, except for impulsivity, as higher impulsivity scores indicate increased impairment; *r*= correlation coefficient; 95%CI= 95% confidence interval; *Q*=Cochran's *Q* test for heterogeneity; I^2= Percent of heterogeneity between studies, *RE*= Random Effects Model.

Prospective Studies. An additional subgroup analysis was performed to examine the association between cognitive functioning and violence in studies which used a prospective design. There were only enough studies to examine this in impulsivity, attention, and insight, all of which revealed significant findings. Small effect sizes were found for impulsivity with only 21% heterogeneity (k = 6, n= 882; r= .27, [.20, .25]), attention (k = 4, n= 254, r= -.16, [-.28, -.03], I^2= 17%), and insight (k = 4, n= 633, r= -.17, [-.26, -.07], I^2= 0%). In this review, studies which tested participants at baseline and followed them up prospectively were rated to be of better quality relative to those which used a retrospective or correlational design, however, they still only produced small effect sizes. Also noteworthy is, of the six prospective studies under impulsivity, only one used a measure specifically designed to assess impulsivity (McDermott et al., 2008), while the other studies used risk assessments or clinical rating scales. This was also the case for insight, whereby all the papers used risk assessments or a clinical rating scale to measure the domain.

Exploratory Analysis

Intelligence. Intelligence was also explored in this portion of the review in k= 5 studies (n= 611) and resulted in a small significant effect (r = -.14 [-.25, -.02]), and moderate heterogeneity (I^2= 54%). After the removal of an outlier (Howard et al., 2014, the only correlational study), heterogeneity decreased to 0% and increased the effect size (k= 4, n = 511; r = -.18 [-.25, -.10]). All the papers measured intelligence with a Wechsler Scale, however, only three papers found small, significant effects (e.g., Fullam & Dolan, 2008, r = -.32 [-.50, -.11]; Dejong et al., 1992, r = -.17 [-.29, -.05]; and Beggs and Grace, 2008, r = -.14 [-0.27, -0.01]). The study with the largest effect was Fullam and Dolan, which was also rated as a good quality paper, conversely, the Dejong et al paper was rated as low quality, and Beggs and Grace as fair quality, however, both the Dejong et al and the Beggs and Grace papers had much larger samples relative to the other studies, and both used the same outcome of violent recidivism. There were not enough papers in part 2 of this review to examine subscales of FSIQ. Like intelligence in the between group studies, the effect sizes here are negligible, and the confidence intervals suggest that the true effect would not exceed a small magnitude.

4.4 Discussion

Despite the supporting evidence from reviews that cognitive impairments are indeed associated with violent, aggressive, and antisocial outcomes (Morgan & Lilienfeld, 2000; Ogilvie et al., 2011; Reinharth et al., 2014), they are often not considered in relation to violence risk. Identifying differences in cognitive functioning between violent and non-violent offenders may help to demarcate the impairments which are more common in one population relative to the other and may provide additional information necessary to improve their treatment, care, and management. Likewise, identifying measures of cognitive impairments that correlate with violent outcomes may have the potential to inform risk assessments, and perhaps improve them. This chapter reported a systematic review and meta-analyses of between group differences and correlations of violence in 42 studies involving 5,411 offenders. To the authors' knowledge, this is the first meta-analytic review to focus on cognitive contributors to violence in only forensic populations, beyond executive functioning.

4.4.1 Cognitive Differences Between Violent and Non-Violent Offenders

Of the 12 cognitive domains that were analysed, effect sizes ranged from small to large, and five had significant pooled effects in the hypothesised direction showing that, within these studies, the violent offenders had significantly poorer performances on neuropsychological measures when compared to non-violent offenders. Notably, of these the only significant domains found to be homogenous without further break down included reasoning, impulsivity, and expressive speech, suggesting robust findings. Sensitivity analyses revealed homogenous findings on the remaining domains after removing outliers, however, these findings should be interpreted cautiously. All the included studies in this section of the review were cross-sectionally analysed, used validated neuropsychological measures, all individuals in the violent groups were charged or convicted of a violent offence, and all those in the non-violent groups has no known recorded violent offences, but were charged or convicted with non-violent offences. The primary differences between these studies were age distribution of the populations, the various neuropsychological measures used, whether they assessed participants in their own sample, as opposed to using archived data, their exclusion criteria, and whether their groups were matched for analyses.

Upon examination of the forest plots, the paper by Bryant et al. (1984) was the only study to consistently contribute a significant individual effect to the meta-analyses, and the effect was often much larger compared to those provided by other studies. As previously stated, Bryant et al. (1984) included adult prisoners with no exclusion criteria, and assessed their own participants using the LNNB, which differentiated it from the other included studies, aside from Ullman (1989) who also used a subtest from the LNNB in a sample of adult prisoners, but only to measure attention and expressive speech. Whilst Bryant et al. contributed the largest effect size in most cases to the meta-analyses, under the expressive speech domain, both Bryant et al., and Ullman found comparably large significant effect sizes using the LNNB. Though it should be acknowledged that Bryant et al. assessed for learning disabilities in 60 out of 110 participants and found that a proportion had learning disabilities, operationalised as IQ < 79 (e.g., violent group= 31%, non-violent group= 24%), which may explain the greater magnitude of effects compared to the other studies, as it is presumed that those individuals had more obvious cognitive impairments relative to the rest of the sample. When subsets of papers were examined by type of offender, there were only available data to examine prisoners as a subgroup. Although the effects from the subsets of papers were of a similar magnitude as the pooled effect sizes for all the papers, once broken down, heterogeneity decreased under the attention domain, suggesting that perhaps, type of offender may have a moderation effect, similar to Ogilvie et al (2011) findings, however, there were not enough papers to examine this using robust subgroup analyses.

Based on these findings, the use of the LNNB in prisoners, and samples with minimal exclusion criteria, produced the most robust effects for examining between group differences. Aside from the measures used, three out of the five studies that contributed significant effects assessed juveniles, though the largest effects consistently came from the Bryant et al study. Moreover, Bryant et al. (1984), Kennedy et al. (2011), and Brimigion (2014) studies boasted moderate to large significant effects and had minimal or no exclusion criteria, whereas the papers by Ullman (1989) and Tarter et al (1983) had somewhat smaller effects on certain domains and had stricter exclusion criteria. This finding suggests that there are more significant differences between violent and non-violent offenders who have risk factors for cognitive impairment, relative to those who do not, and may also suggest an additive effect. The possible additive effect may explain the larger effect sizes, and further emphasizes the notion that assessment of cognitive impairments in relation to violence risk may be more important for offenders with risk factors for cognitive impairment. This finding links back to the literature reviewed in chapter 2 of the thesis which suggested that the co-occurrence of risk factors for cognitive impairments such as substance use, psychosis, and head injuries, may exacerbate cognitive impairments (e.g., D. N. Allen et al., 1999; Fujii et al., 2004; Sachdev et al., 2001).

Overall, the variability observed in effect sizes, and the ability for some measures to differentiate similar populations and not others, demonstrates the heterogeneity that is likely to be found in cognitive impairments, both within and between the groups. Thus, to gain a better understanding of the measures and cognitive impairments that differentiate violent from non-violent offenders, there is a need for replication studies using the same neuropsychological measures, in similar samples to those included in this review. Given that several of the studies did not produce meaningful effects, it is hypothesised that the wide range of neuropsychological measures used in the individual studies may have buried the true effects, and contributed, at least partially, to the variability in magnitude of the effect sizes. Therefore, attempting to replicate initial findings may allow for a greater understanding of the neuropsychological measures that are useful for differentiating violent from non-violent offenders, and likewise, may aid in developing a more focused list of measures for researchers to use in this context to decrease heterogeneity.

4.4.2 Cognitive Correlates of Violent Offending

The five cognitive domains that were analysed revealed small effects ranging from *r*= -0.25 to 0.26, with three being significant in the hypothesised direction suggesting that increased cognitive impairment on certain domains is correlated with increased violence. Notably, the only significant domains found to be homogenous without further break down included impulsivity, attention, and insight, suggesting robust findings.

Findings from this portion of the review suggest that effect sizes are related to the operationalization of the outcome and the measure employed in the individual studies. For example, as previously stated, Belfrage et al. (2000) provided the largest effect for impulsivity using the HCR-20, however, in comparison to some of the other papers (e.g., McDermott et al. and Smith et al.), Belfrage et al employed a much broader operationalization for their outcome. This finding demonstrates one of the issues in this type of research, where specifically defined outcomes (e.g., physical aggression) found no significant effects, and a broadly defined one did (e.g., institutional violence, including severe assault, threats, or property damage). To reiterate Monahan's (1981) thoughts presented in chapter 3, the larger the target, the easier it is to hit. While the main focus of this review was on physical aggression and violence, broad operationalisations, such as Belfrage et al's were difficult to disregard, and are considered a limitation in this review. However, it should be highlighted that Belfrage et al's findings may be misleading as it is difficult to decipher whether the significant effect is for actual, realised violence, or for severe threats and destruction of property. Findings such as this one contributes to the uncertainty observed in risk assessment research and adds little to the understanding of risk factors for physical violence, specifically.

Findings from the post-hoc analyses under the impulsivity domain suggest that the use of risk assessments and clinical judgement to measure impulsivity produces a slightly larger correlation to violent outcomes relative to behavioural, self-report, and clinical measures, however, even once the papers were broken into subgroups and the HCR-20 was examined on its own, it produced moderate heterogeneity. Further, it is unclear what information in relation to impulsivity was considered when the risk assessments were completed, which may have added to the overall heterogeneity. This finding also suggests that self-report, clinical, and behavioural measures are more suitable for the assessment of cognitive functioning relative to risk assessments, as aside from their potential for *task impurity*, (e.g., when tasks measure multiple, rather than the intended single, abilities; Burgess, 2004; Miyake & Friedman, 2012), they were developed to measure cognitive abilities specifically, whereas risk assessments were not. Therefore, the use of valid neuropsychological and clinical measures to assess the neuropsychological risk factors on risk measures, such as impulsivity, may reduce heterogeneity and increase accuracy. This notion was further evidenced under the insight domain where Alia-Klein et al (2007) reported the largest effect and was the only study to use a measure specifically designed to measure insight (e.g., SUMD). Additionally, like the between group studies, the results from the attention domain re-emphasized the possibility for an additive effect, as studies which excluded participants with certain characteristics (e.g., psychosis, low IQ, and head injury) found smaller effects relative to those which included participants with a diagnosis of schizophrenia and other psychotic illnesses.

Finally, the analysis of prospective studies suggest that neuropsychological domains found to be significantly associated with violent outcomes (e.g., impulsivity, attention, and insight), would individually add approximately 3-7% incremental variance to contemporary risk assessments, though impulsivity and insight are already included in certain risk measures. However, as it is unclear whether the risk assessment ratings of these domains were based on scores from valid neuropsychological assessments, it cannot be concluded that the assessment of cognitive abilities do not add incremental validity to risk assessments based on these findings alone. In order to move this field of research forward, transparent, high quality, replicable studies using the same neuropsychological measures in various forensic samples, prospectively, is crucial. Thus, like the conclusions Ogilvie et al (2011) drew based on their review findings, due to the heterogeneity in effect sizes resulting from variations in methodologies, the true value of effect sizes for these domains cannot be confidently estimated at this time. However, given the large proportion of small effect sizes and wide confidence intervals, it is postulated that many of the reviewed cognitive abilities may not be meaningful additions to violence risk assessments, though, more research is needed.

4.4.3 Limitations of Review

Congruent with similar reviews, heterogeneity of the meta-analytic data remains the primary limitation of this review. As a result of potentially poor differentiation of violent and non-violent samples, conceptual overlap in cognitive measures and differing definitions of violent outcomes, it remains difficult to confidently interpret the domains which robustly differentiate violent and non-violent groups. The narrow inclusion criteria utilized in part two sought to address these concerns by ensuring the inclusion of only studies measuring actual violence (although this was often difficult to decipher), but this may in turn have limited the overall generalizability of these results. Likewise, as violent behavior is heterogeneous, subtypes of violence, such as impulsive and instrumental, differ in their origins, mechanisms, and management (Volavka, 1999). Thus, when sufficient data are available, the examination of subgroups within violent populations will assist in ascertaining information on specific risk factors for each group, perhaps decreasing the overall heterogeneity and lending to more tailored treatment programs and increased accuracy of risk assessments. Furthermore, along with validated cognitive assessments, violence risk assessments were also included in this review as a measure of cognitive abilities. Although these measures are widely used, they are often scored using clinical judgement, rather than structured assessment.

4.4.4 Considerations for Clinical Utility

Neuropsychological functioning is often not considered in relation to assessing violence risk and according to the findings of this and past reviews (e.g., Reinharth et al., 2014), this is an area that is in need of further examination. Based on their own review findings, Reinharth and colleagues (2014) recommended the use of behavioural cognitive measures to assess cognitive functioning (and clinical rating scales to measure insight), as well as the addition of global cognitive measures or proxy measures to risk assessments for aggression in individuals with psychosis. However, based on the findings from the current review, more replicable studies are needed before this can be implemented into practice. Albeit the results of this and past reviews (e.g., Morgan & Lilienfeld, 2000; Ogilivie et al., 2011; Sedgwick et al., 2017) evidence that cognitive impairments are not only manifested in populations with schizophrenia and other psychoses, but also in prisoners and court referred individuals with and without major mental illness, although they may be more pronounced in individuals with schizophrenia (Sedgwick et al., 2017), and those with TBI. Such populations traditionally have less access to cognitive assessments, but there is potential for this dearth to be addressed through the increasing availability of software measures allowing self-completion. Acknowledgement of the cognitive impairments which may surface in certain populations, may aid in the treatment and management of these individuals, and should be considered in formulations of risk.

4.4.5 Recommendations for Future Research

Due to the multidimensional underlying structure of many neuropsychological measures, with several different cognitive abilities interacting to explain one given performance (Duggan & Garcia-Barrera, 2015; Karr et al., 2018), and their reputation for task impurity, it is recommended that efforts are made to reach a consensus in the field on which cognitive measures are the most robust in measuring cognitive risk factors for violence. It is further recommended that a consensus be reached on the cognitive domains which are most important to violence risk. As it currently stands, the wide variety of differential cognitive measures, and the use of one assessment to measure several different functions has inhibited the ability for findings to be replicated and easily synthesised. Future research should focus on employing valid and reliable measures of dissociable cognitive functions, not rated using clinical judgement, and these should be clinically available rather than designed for research practice only, to aid in application. Individual studies should focus on stating the exact cognitive functions they have set out to measure, thoroughly and specifically operationalising violence, and employing prospective designs to increase the accuracy of results. Further examination of the magnitude of impairments observed in violent offenders with historical risk factors for cognitive impairment relative to violent offenders without these risk factors may serve as a valuable steppingstone to gaining a better understanding of these populations.

4.5 Conclusions

The current review investigated the cognitive impairments which differentiate violent and non-violent offenders, as well as those associated with violent outcomes. In comparison to non-violent offenders, violent offenders had significantly poorer performance on measures of reasoning, impulsivity, and expressive speech, all with small effect sizes, except for expressive speech which had the largest effect size. Domains which significantly and homogenously correlated with violence were impulsivity, attention, and insight, all with small effect sizes. However, based on these findings, it cannot be concluded that valid and reliable measures of disaggregated cognitive functions will significantly improve the predictive accuracy of contemporary risk assessments, though it is postulated that they may not have utility in this way based on the overwhelming proportion of small effect sizes. The findings of this review highlight the need for more high quality, replicable studies using valid and reliable measures of dissociable cognitive functions before a conclusive decision can be made on whether the addition of cognitive functions will add value to violence risk assessments.

5 Achieving an Expert Consensus on the Cognitive Domains Considered Relevant for Inclusion in Structured Violence Risk Assessments: An International Delphi Study

Chapter Preface

Chapter 4 presented a systematic review of the literature, a quantitative synthesis of the findings, and a critical evaluation of the current state of the research in this area. Whilst many of the individual studies did not demonstrate significant effects between cognitive variables and violence, several pooled effects emerged as significant, however, most of them were small in magnitude. The size of the effects in this area of research has casted doubt on the ability for measures of neuropsychological abilities to add significant incremental validity to existing risk assessments, however, this may be the result of the quantity and quality of the literature, which remains limited. Potentially further compounding this is that the measurement of latent variables is not always straightforward, and there is a broad spectrum of neurocognitive measures, most of which can tap into more than one ability, which is likely contributing to the heterogeneity observed across studies. As the quality and quantity of studies examining the association between cognitive functions and violence risk is limited, such that it cannot be solely relied upon, the primary aim of the current study was to utilize the Delphi method to obtain a list of cognitive domains that would be relevant for inclusion on a structured violence risk assessment. The use of the Delphi method would allow these findings to be based on the consensus of an expert panel, made up of both researchers and clinicians, to ensure the capture of domains that are relevant, but have yet to be studied or are underrepresented in the literature. The central research question for the current study was, what are the top 10 cognitive domains relevant for inclusion on a structured violence risk assessment?

5.1 The Delphi Method

The Delphi method is a process used to collect the judgements and opinions of experts in a certain field using a series of questionnaires combined with feedback (Skulmoski et al., 2007). It is a method that is suitable for use when there is incomplete knowledge or research on a topic and allows for an opportunity for experts to give their opinion about a subject, they are knowledgeable about. The Delphi process is conducted in rounds, and includes questionnaires that are each developed based on the results of the former round, with the process stopping when the research question has been answered, or consensus has been met on the questionnaire items (Skulmoski et al., 2007). Although this method was originally created to search for full consensus, more modern applications have defined it as a ‘social research approach’ where the main aim is to gain a reliable group opinion from a group of experts (Landeta, 2006). Strengths of the Delphi method include achieving consensus or answering research questions on topics where there is little empirical evidence or uncertainty, broadening knowledge, as well as the sharing of knowledge and opinions, while maintaining anonymity and avoiding group conflict (Fefer et al., 2016; Gupta & Clarke, 1996; McKenna, 1994; Stokes, 1997). Additionally, the Delphi method allows for an opportunity to gain expert knowledge about a particular topic from an international sample, in a cost-effective manner.

5.2 Method

5.2.1 The Delphi Process for the Current Study

As Delphi studies can vary depending on the question being asked, three rounds were anticipated in the current study, with the intention of evaluating the need for a fourth if the question remained unanswered or consensus was not reached. For each Delphi round, data were generated, analysed, and sent back to the panellists in the form of a new questionnaire.

5.2.2 Literature Review and Questionnaire Development

A review of literature was conducted in December 2015 and January 2016, identifying papers that were published between 2010 and 2016 to identify cognitive domains that had been the subject of research in this field. Electronic databases (Web of Science, PsycInfo, and Medline) were searched using keywords which corresponded with cognitive deficits, mental illness, violence, offenders, and violence risk. The inclusion criteria were purposely kept broad and included, violent samples (violent offenders, delinquency, psychopathy, sexual offenders) of any age and the use of any neuropsychological tools to measure cognitive abilities. Online Surveys (OS; formerly Bristol Online Survey; Online Surveys, 2018) was used as a means of creating, analysing, and distributing the questionnaire for each round, and communication with panellists was conducted via email. Subsequent to its development, the questionnaire was piloted amongst academic colleagues to ensure clarity and the generation of appropriate data.

5.2.3 Participant Selection

To ensure that the proposed structured risk assessment is both based on a rigorous appraisal of the research literature and fully meets the needs of those working in the relevant fields, the Delphi panel was comprised of researchers and clinicians specialising in the fields of forensic and neuropsychology. According to Okoli and Pawlowski (2004), who recommend 10 to 18 experts as a sample size, the Delphi panel size depends more on the group arriving at a consensus than on statistical power. We chose to invite a large international sample of experts, hoping to exceed the recommended sample size, and thus engage a heterogeneous group encompassing diversity and experience from various backgrounds.

Clinicians in the fields of forensic and neuropsychology were contacted, as it was posited that the inclusion of practicing clinicians in this process would identify essential domains that have yet to be thoroughly examined in the literature but are apparent in clinical practice. To obtain robust and diverse expert opinions, clinicians from both the United Kingdom and the United States were contacted. For the United States, the American Psychological Association (APA) website was used to identify practicing clinicians who specialise in both forensic and neuropsychology. Psychologists chartered in both forensic and neuropsychology in the United Kingdom were identified through the British Psychological Society (BPS) website utilising the search function. Only clinicians with a listed email address were invited to participate.

Researchers were identified through the same procedures as the literature review from 2010-2016. To ensure that researchers were likely to remain active in the field, it was decided to only include studies from the last five to six years for this search. Researchers listed as the contact author on studies published between these years, in this or a relevant field or relevant topic (forensic, clinical and neuropsychology, as well as risk assessments and risk factors for violence and offending) were contacted.

5.2.4 Rounds

For each round, a link was generated for the questionnaire, and was sent to all panellists via email. Panellists who completed the round one questionnaire were instructed to enter their email address if they wished to take part in subsequent rounds, which was how the panel was constructed. As individual feedback was not given to panellists in this study, they were given the option to send their email to a separate email account, enabling their answers to remain anonymous. Further, only the research team was able to see listed emails and answers, allowing for confidentiality between panel members. As the first round also served as the initial invitation into the study, the instructions made clear the type of expertise and knowledge sought and participants were asked to participate only if they met these requirements.

In all rounds, panellists were first presented with an information sheet, and by continuing onto the questionnaire, they were considered to have consented to take part in the study. In round one, all panellists were asked questions about their professional backgrounds, area of expertise, and the country they were currently working in. Rounds two and three followed these same procedures, though also asking years of experience. Details on the process of each round are presented in Table 5.1.

5.2.5 Analyses

Demographic information for all rounds, percentages, and mean and median ranks for rounds 2 and 3 were obtained from OS. In round one, domains which were rated as *essential* or *relevant* (combined) by ≥80% of the panellists, were kept for successive rounds (Langlands et al., 2008). To quantify the agreement among panel members, Kendall's Coefficient of Concordance (W) using the *irr* package (Gamer et al., 2012) in R Studio (R Core Team, 2020) was utilised. In this nonparametric statistical approach, a strong agreement or consensus exists for $W \geq 0.7$; a moderate agreement for $W = 0.3\text{-}0.6$; and a weak agreement for $W < 0.3$ (Schmidt, 1997). This statistic has been used in other Delphi studies to quantify consensus (García-Crespo et al., 2010; Schmidt et al., 2001). Standard deviations for each domain from rounds two and three were also used to consider stability and consensus, where a decrease between rounds indicated a move toward consensus (Fefer et al., 2016). Whilst the researcher was interested in the degree of consensus reached, as well as the stability between rounds, the research question was considered to be answered once we had a list of the top 10 ranked cognitive domains.

Ethical permissions for the current study were obtained from the University of Edinburgh and can be found Appendix F.

Table 5.1 Delphi Process for Each Round

Round	Aims of Round	Questionnaire	Analyses	Sub-Analyses
1	**1.** Refine the initial list of cognitive domains identified through the literature review. **2**. Obtain a working list of tools that panellists use to measure specific domains, for potential use alongside the final risk assessment.	**Panellists were asked to:** **1.** Rate domains as *Essential*, *Relevant*, or *Irrelevant* for inclusion on a violence risk assessment. **2.** Suggest in a free text box how best the domain can be operationally defined.	Any domains that were rated as *essential* or *relevant* by ≥80% of the panellists were considered to be the core list of domains to be used for subsequent rounds (Langlands et al., 2008).	Responses from free text boxes were only collated in Round 1, and responses were used to inform the researchers of the measures being used by both researchers and clinicians.
2	**1.** To further refine the list of domains, and to identify those which were considered most important in relation to violence risk.	Panellists were given the results from Round 1 and were asked to rank domains in order of importance.	Mean ranks were calculated for each domain, and then ranked from smallest to largest for the next round.	Kendall's *W* was calculated to quantify agreement among panellists.
3	**1.** To obtain a final list of the cognitive domains ranked in order of importance. **2.** To present the panellists with preliminary meta-analytic findings and their corresponding effect sizes. **3.** To see if panellists would change their rankings based on results from Round 2, and the meta-analysis.	**Panellists were asked to:** **1.** Rank top 10 domains from Round 2 in order of importance. **2.** See results from Round 2 and meta-analytic findings. **3.** Rank the same list of domains as Round 2 plus domains from the meta-analysis.	Domains were ranked by means to obtain the final list.	Kendall's *W* was calculated to quantify agreement among panellists.

Note. Cognitive domains to be ranked or rated were sent in a random order in all rounds.

5.3 Results

5.3.1 Response Rates

In round one, a total of *n*= 694 researchers and clinicians were invited to participate (*n* = 427 researchers, *n*= 42 chartered forensic and neuropsychologists from the BPS website, and *n*= 267 practicing forensic and neuropsychologists from the APA website). Round one saw *n*= 56 responses, resulting in an 8% response rate. As only panellists who wished to participate in the following rounds left their email address, a full breakdown of the exact proportion of participants from each group (e.g., research, BPS, or APA) could not be calculated. After round one, 80% (*n*= 45) of the panellists who responded to the questionnaire, indicated that they would like to take part in round two. In round two, *n*= 29 responses were received, a 64% response rate, and for round three, the questionnaire was sent to *n*= 24 panellists, and *n*= 19 responses were received, a 79% response rate. The time required for each of the Delphi questionnaires in this study was approximately 15 to 20 minutes per round.

5.3.2 Panellist Profiles

In round one, panellists responded from 20 different countries, where 28% were from the United States and 18% were from the United Kingdom (Table 5.2). Both countries made up the majority of panellists for successive rounds, and whilst the countries being represented from round one decreased due to attrition, the sample for all rounds remained diverse and representative of an international sample. Over half of the panellists in round one (66%), had expertise in both research and clinical practice (Table 5.3), and this group remained in the majority for all rounds. Likewise, professional background was diverse, and mostly evenly split across disciplines (Table 5.3).

Table 5.2 Number of Panel Members from Each Country for Each Round

Country	Round 1 *n* (%)	Round 2 *n* (%)	Round 3 *n* (%)
United Kingdom	10 (18)	6 (21)	4 (21)
United States	16 (29)	6 (21)	4 (21)
Canada	3 (5)	2 (7)	1 (5)
Netherlands	3 (5)	1 (3)	1 (5)
Portugal	3 (5)	2 (7)	2 (11)
Norway	1 (2)	1 (3)	1 (5)
Finland	1 (2)	1 (3)	1 (5)
Germany	1 (2)	0 (0)	0 (0)
New Zealand	2 (3)	1 (3)	0 (0)
Sweden	1 (2)	1 (3)	1 (5)
Romania	1 (2)	1 (3)	1 (5)
Italy	2 (3)	1 (3)	1 (5)
Croatia	1 (2)	1 (3)	0 (0)
France	1 (2)	1 (3)	1 (5)
Ecuador	1 (2)	0 (0)	0 (0)
Australia	4 (7)	2 (7)	1 (5)
Spain	2 (3)	1 (3)	0 (0)
South Korea	1 (2)	0 (0)	0 (0)
Brazil	1 (2)	1 (3)	0 (0)
Egypt	1 (2)	0 (0)	0 (0)
Total panellists in each round	56	29	19

Note. This table presents the number of participants from each country for each round, and a row for the total participants in each round. Due to rounding, not all percentages sum to 100.

Table 5.3 Professional Background and Expertise of Panellists for Each Round

Professional Background	Round 1 *n* (%)	Round 2 *n* (%)	Round 3 *n* (%)
Neuropsychology	4 (7)	3 (10)	0 (0)
Forensic Psychology	5 (9)	6 (21)	4 (21)
Clinical Psychology	5 (9)	3 (10)	2 (11)
Psychiatry	9 (16)	2 (7)	2 (11)
Neuropsychology and Forensic	2 (3)	0 (0)	0 (0)
Forensic and Clinical	12 (21)	6 (21)	3 (16)
Neuropsychology, Forensic and Clinical	6 (11)	3(10)	3 (16)
Neuropsychology and Clinical	7 (13)	2 (7)	2 (11)
Other	6 (11)	4 (14)	3 (16)
Researcher Only	13 (23)	6 (21)	6 (32)
Clinician Only	6 (11)	0 (0)	0 (0)
Both Clinician and Researcher	37 (66)	23 (79)	13 (68)

Note. This table presents background information on the participants for each round. Due to rounding, not all percentages sum to 100.

In round two, for panel members who were both clinicians and researchers, the mean years of experience as a clinician was 22.20 (12.03) years, and the mean years of experience in research was 14.50 (10.30) years; while those whose main role was solely research, had a mean of 8.00 (1.66) years' experience. Mean years of experience for panel members in round three were, clinicians 23.69 (11.44) years; researcher (dual role) 21.84 (12.64) years; and researcher only 9.66 (4.50) years. When asked if the panel members' main focus was in forensic risk, or conducting risk assessments, for round two, 65% of the 'research only' group answered yes, and 83.3% in round three; 67% of the dual group (researcher and clinician) answered yes to forensic risk being a main focus of their research, and 62% in round three, and 57% answered yes, that conducting risk assessments was part of their role as a clinician in round two, and 62% in round three.

5.3.3 Delphi Round One

The literature review identified 41 separable cognitive domains relevant to violence risk, which formed the questionnaire for round one. Sixteen domains met the criteria (rated as *essential* or *relevant* by 80% or more of the panel) to be considered as a core cognitive domain for subsequent rounds, and the remaining domains were excluded from further analyses (Table 5.4). Figure 5.1 demonstrates the domains that were rated as *essential* or *relevant* by 100% of the research only group, 100% of the clinician group only, those which were rated as *essential* or *relevant* by 100% of both groups. For the whole sample, impulsivity and poor inhibitory control were both rated as 100% essential or relevant by the panel, followed by risk taking (98.3%), affect recognition (96.4%), and reasoning (94.8%). Most of the domains which met criteria for inclusion in further rounds were executive functions and social cognitive domains.

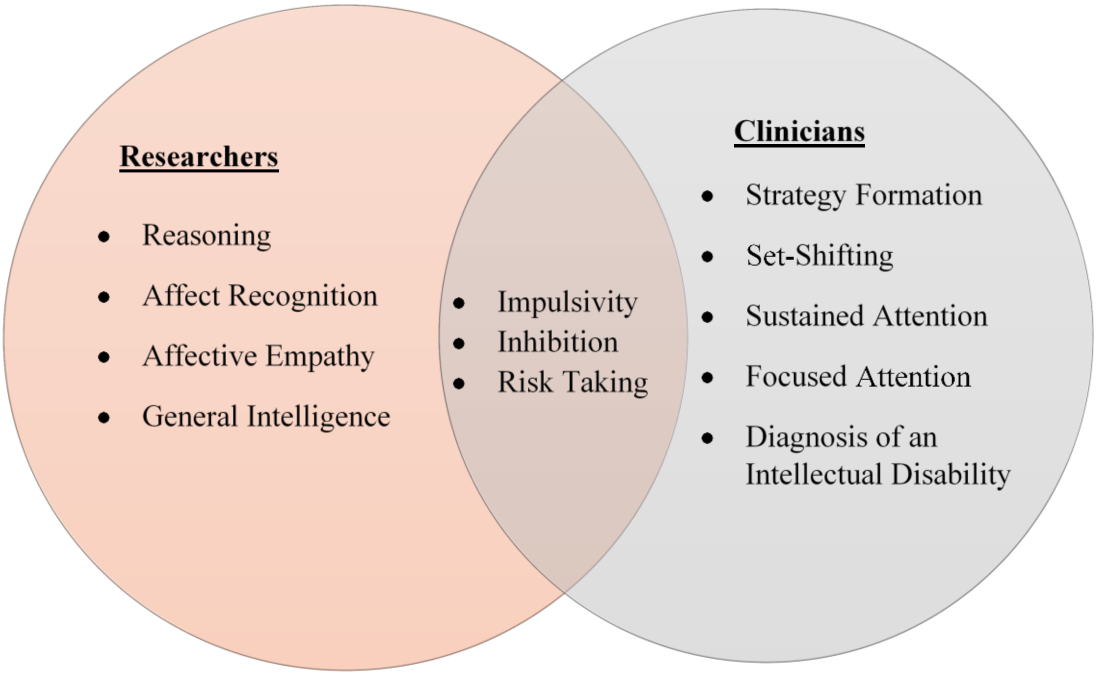
Free text boxes encouraged further insight into the development of a structured risk assessment, for instance panellists suggested that the battery not be longer than two hours, lest its use becomes impractical in corrections/probation/parole settings, likewise, other panel members questioned whether there was enough evidence to support these domains at a group level, rather than an individual one, and that they should not be looked at in isolation of the broader system of negative influences, and risk and protective factors.

Table 5.4 Results of Round One

Domain	% Essential	% Relevant	% Irrelevant	% Essential/ Relevant
Behavioural Impulsivity	94.6	5.4	0	100
Poor Inhibitory Control	85.7	14.3	0	100
Risk Taking	55.4	42.9	1.8	98.3
Affect Recognition	44.6	51.8	3.6	96.4
Reasoning	48.2	46.6	5.4	94.8
Affective Empathy	55.4	39.3	5.4	94.7
Cognitive Empathy	48.2	41.1	10.7	89.3
Planning	41.1	48.2	10.7	89.3
Set-Shifting/Attention Shifting	26.8	60.8	12.5	87.6
Cognitive Insight	28.6	57.1	14.3	85.7
Diagnosis of Intellectual Disability	30.4	53.6	16.1	84
Response Monitoring	35.7	48.2	16.1	83.9
Perseveration	26.8	57.1	16.1	83.9
Clinical Insight	30.4	51.8	17.9	82.2
General Intelligence	26.8	55.4	17.9	82.2
Theory of Mind	37.5	44.6	17.9	82.1
Abstract Thinking Ability	16.1	60.7	23.2	76.8
Focused Attention	23.2	51.8	25	75
Verbal Intelligence	19.6	55.4	25	75
Sustained Attention	26.8	46.4	26.8	73.2
Verbal Comprehension	23.2	50	26.8	73.2
Affective Mentalising	26.8	44.6	28.6	71.4
Mentalising	25	42.9	32.1	67.9
Concept Formation	14.3	53.6	32.1	67.9
Verbal Language Functioning	14.3	51.8	33.9	66.1
Strategy Formation	16.1	62.5	21.4	64.3
Performance Intelligence	8.9	53.6	37.5	62.5
Associative Learning	8.9	53.6	37.5	62.5
Response Reversal	9.4	49.1	41.5	58.5
Verbal Working Memory	17.9	39.3	42.9	57.2
Episodic Memory	16.1	41.1	42.9	57.2
Neurocognitive Insight	10.7	46.4	42.9	57.1
Verbal Learning	16.1	39.3	44.6	55.4
Verbal Memory	16.1	39.3	44.6	55.4
Visual Working Memory	14.3	35.7	50	50
Slow Reaction Time	7.1	41.1	51.8	48.2
Visual Memory	14.3	30.4	55.4	44.7
Visual Perception	8.9	30.4	60.7	39.3
Visuospatial Perception & Organisation	5.4	33.9	60.7	39.3
Visuospatial/Construction Abilities	7.1	26.8	66.1	33.9
Motor Dexterity	3.6	17.9	78.6	21.5

Note. n=56; light shading highlights the domains which were kept for successive rounds; dark shading highlights variables that were excluded.

Figure 5.1 Domains that 100% of Researchers and Clinicians Rated as Essential or Relevant in Round One



5.3.4 Delphi Round Two

In round two, the panel was instructed to rank the list of 16 domains identified in round one, in order of importance. The top five domains included poor inhibitory control, behavioural impulsivity, risk taking, affective empathy and response monitoring. To quantify the consensus of the panel, Kendall's *W* was calculated and resulted in an effect size of W= .22, p <.0001, although significant, indicating ratings were better than chance, it demonstrates a weak agreement. Like round one, the top 10 domains consisted of executive functions and social cognitive domains, and there was almost no differentiation made between cognitive and clinical insight (Table 5.5).

Table 5.5 Results of Round Two

Rank	Domain	Mean Rank (SD)	Median Rank	Lower Quartile	Upper Quartile
1	Poor Inhibitory Control	3.21 (3.59)	2.00	1	4
2	Behavioural Impulsivity	3.48 (3.60)	2.00	1	5
3	Risk Taking	5.28 (3.59)	5.00	3	7
4	Affective Empathy	5.34 (3.93)	4.00	2	9
5	Response Monitoring	6.21 (3.53)	6.00	4	9
6	Cognitive Insight	6.48 (3.74)	6.00	3	8
7	Clinical Insight	6.48 (4.45)	5.00	2	8
8	Reasoning	6.52 (3.83)	6.00	4	9
9	Cognitive Empathy	6.93 (3.97)	7.00	4	9
10	Affect Recognition	6.93 (4.76)	6.00	2	11
11	Theory of Mind	7.03 (4.77)	6.00	2	11
12	Planning	7.31 (4.64)	8.00	3	12
13	Perseveration	7.38 (3.83)	7.00	4	10
14	Set-Shifting	7.97 (4.57)	8.00	4	12
15	Intellectual Disability	9.10 (4.69)	9.00	6	13
16	General Intelligence	9.17 (4.19)	9.00	6	13

Note. This table presents the ranked results from round two. SD= standard deviation.

5.3.5 Delphi Round Three

Round three consisted of two phases, in phase one, the panel members ranked the top 10 domains from round two, and in phase two, they were given both the results from round two and those of a preliminary meta-analysis on this subject before being asked to rank the top 10 domains from round two, again, plus two domains from the meta-analytic results[5], one of which was previously absent (i.e., processing speed, though it was covered under performance intelligence), and the other (i.e., attention) was brought back from round one. While the results of phase one provided a final consensus for the Delphi study, those of phase two sought to identify the optimum list when all information was made available, thus phase one results were not influenced by bias in what domains have been researched (i.e., meta-analytic findings), but gave an indication of those which were considered relevant but may not have been the

[5] The meta-analysis was based on the original literature review conducted for this study, which had broader inclusion criteria than the one in Chapter 4, and only studies between 2010-2016 were included.

subject of study. For round three, phase one, the consensus for the 10 top domains as represented by Kendall's *W* increased slightly from the previous round, $W = .32$, $p < .0001$, and while significant, consensus remained weak, however, standard deviations for each domain decreased from round two, indicating less variance among ratings. With the results of this round, our research question was considered to be answered, and the domains in Table 6 were considered to be our core list of the most essential cognitive domains, in order of importance, for consideration in assessing violence risk.

Table 5.6 Results of Round Three, Phase One

Rank	Domain	Mean Rank (SD)	Median Rank	Lower Quartile	Upper Quartile
1	Poor inhibitory Control	2.00 (1.34)	2.00	1.00	2.00
2	Behavioural Impulsivity	2.53 (2.21)	2.00	1.00	2.50
3	Risk Taking	4.00 (2.79)	3.00	2.00	5.50
4	Affective Empathy	4.42 (2.62)	4.00	2.00	6.50
5	Response Monitoring	4.63 (2.58)	4.00	3.00	6.50
6	Affect Recognition	5.32 (3.15)	5.00	3.00	8.00
7	Cognitive Insight	5.63 (2.58)	5.00	4.50	7.00
8	Reasoning	5.63 (2.62)	6.00	3.50	7.50
9	Cognitive Empathy	5.75 (2.40)	7.00	3.50	7.50
10	Clinical Insight	6.32 (3.23)	8.00	3.50	9.50

Note. This table presents the ranked results for round three, phase one. SD= standard deviation.

The results for round three, phase two are displayed in Table 7. Of the 18 domains presented to the panel, only one of the top 10 core domains, cognitive insight, moved down to rank 12, and general intelligence moved up to rank 10, indicating that the panel remained firm in their views, with the order of domains changing only slightly. Standard deviations for each domain increased from phase one of this round, however, this is likely due to the longer list of domains that the panel was asked to rank. Consensus for this phase, for all 18 items resulted in $W = .40$, $p = < 0.0001$, indicating a moderate/weak consensus. The agreement for the top 10 domains also indicated a moderate consensus ($W = .45$, $p < .0001$).

Table 5.7 Results of Round Three, Phase Two

Rank	Domain	Mean Rank (SD)	Median Rank	Lower Quartile	Upper Quartile
1	Behavioural Impulsivity	2.37 (2.30)	2.00	1.00	2.50
2	Poor Inhibitory Control	2.63 (2.68)	2.00	1.00	2.50
3	Risk Taking	4.26 (3.29)	3.00	2.50	5.00
4	Affective Empathy	5.68 (3.64)	4.00	3.50	9.00
5	Response Monitoring	7.68 (4.96)	7.00	4.00	10.00
6	Cognitive Empathy	7.74 (4.31)	6.00	5.00	11.00
7	Reasoning	8.47 (4.32)	8.00	5.50	12.00
8	Affect Recognition	8.63 (4.89)	9.00	5.00	12.00
9	Clinical Insight	8.68 (5.42)	7.00	4.50	14.00
10	General Intelligence	9.05 (4.33)	9.00	5.00	13.00
11	Planning	9.11 (4.29)	9.00	4.50	12.00
12	Cognitive Insight	9.42 (5.08)	9.00	5.50	13.50
13	Processing Speed	9.79 (4.31)	9.00	6.00	13.00
14	Theory of Mind	10.26 (4.84)	10.00	6.50	14.00
15	Set-Shifting	10.63 (4.44)	12.00	7.00	14.50
16	Attention	11.79 (3.64)	12.00	9.50	14.50
17	Intellectual Disorder	11.79 (5.52)	15.00	6.00	16.50
18	Perseveration	11.84 (3.99)	13.00	9.50	15.00

Note. This table presents the ranked results for round three, phase two. SD= standard deviation.

5.4 Discussion

Emerging evidence from neuropsychological studies of forensic populations has promoted the hypothesis that cognitive measures may have the ability to add incremental validity to existing violence risk assessments. However, due to the limited quality and quantity of research examining the association between cognitive impairments and violence, and the lack of consistency in findings across the literature, the current study aimed to achieve an expert consensus on the cognitive domains thought to be most relevant to assessing the risk of violence. Thus, the Delphi method was implemented to obtain a core list of cognitive domains, and a consensus was reached from an international panel of experts. To the researcher’s knowledge, this is the first Delphi study to examine this subject.

5.4.1 The Importance of Gaining Expert Opinion

In addition to the limited quality and quantity of research in this field, studies have suggested that research takes an average of 17 years to translate into clinical practice (e.g., Morris et al., 2011), and likewise, there was an assumed lag between what is observed in clinical practice and what is examined in research. These notions made the opinions of experts in the field of great value to the aims of this thesis, and to facilitate the development of a neuropsychological battery to pilot for the development of a structured risk measure. The combination of empirical findings and expert opinion was seen as a way to increase the robustness of the risk measure, as well as the contribution to the research literature. Moreover, the inclusion of opinions from practicing clinicians may help to fill in the gaps in the literature, as clinicians often observe patterns in behaviours in their own patients, of which researchers may not be fully aware. For example, after round one, there was 100% agreement by both groups that impulsivity, poor inhibitory control, and risk taking were important, however, domains which were agreed on by 100% of clinicians were dominated by attention domains, whereas the researchers agreed more on social cognition domains (Figure 5.1). This discrepancy may be explained by a lag between research and clinical practice, as social cognition is a more recent concept for which fewer clinically available measures exist, and researchers, who may not work with these populations daily, may be less aware of their poor attention span. Furthermore, as cited in previous chapters, study samples are often biased in this field, and are chosen based on specific characteristics, whereas clinicians will gain knowledge based on all the patients they care for, including those who do not have the capacity to consent for research. Finally, the inclusion of both practicing clinicians and researchers gave some insight into the different cognitive abilities which are viewed as essential by the two groups.

5.4.2 Top 10 Domains and Their Relationship to Violence Risk

The ranked list of cognitive domains serves as a foundation for the risk factors perceived as important by experts, and how they value each one. It is not surprising that executive functions, which many of the top 10 domains in phase one of round three can be categorized as, were perceived by experts as being important for the assessment of violence risk. Executive functions play a key role in an individual’s ability to behave pro-socially, and aid in the organization and deployment of cognitive and emotional resources necessary for goal achievement (Barrasso-Catanzaro & Eslinger, 2016). Inhibitory control, perceived as the most important domain in the current study, is an executive function and can be defined as the ability to stop a mental process or action with or without trying to (MacLeod, 2007). Inhibitory control blocks behaviours such as impulsivity and unnecessary risk taking, thus, when inhibitory control is deficient, these behaviours are more likely to manifest, and may contribute to violence. Underlying inhibitory control, are response monitoring and reasoning, both of which are necessary for the proper functioning of inhibitory control, as they allow individuals to observe, evaluate and change their behaviours in the presence of mistakes (Geier, 2013), and to use logical reasoning (Houdé & Borst, 2015). Behavioural and cognitive products of deficient inhibitory control, such as impulsivity, risk taking, and inattention are commonly seen in justice involved individuals, and are considered risk factors for antisocial and violent behaviour. Further, executive functions have been thoroughly examined in relation to violence, and measures of them have demonstrated the ability to significantly differentiate antisocial populations from non-antisocial controls with a mean effect size of d= 0.62 (Ogilvie et al., 2011), with antisocial populations consistently demonstrating poorer performances.

Empathy and affect recognition, considered social cognitive factors, were also perceived as important for the assessment of violence risk by the expert panel. The reason for this may be because intact social cognition allows individuals to recognize emotions from facial expressions, voices, and body movements (Kok et al., 2014). Further, intact emotion recognition is essential for appropriate social interactions, as they allow individuals to predict the behaviour of others and enables them to behaviourally react or respond appropriately, as well as managing their own behaviour appropriately. Impaired social cognition is often seen in individuals with mental illnesses such as schizophrenia, autism, and in individuals with TBI. Impaired social cognition may also engender negative reactions in others, or to perceive positive or neutral emotions negatively, leading to aggression or violence. Few studies have prospectively examined social cognition in relation to violence, however, those which have, have found positive results. For example, in a cohort study of mentally ill offenders, O'Reilly et al. (2015) found that social cognition had a direct effect on violence independent of a neurocognition composite variable, violence proneness, and symptom severity.

Insight, already an established risk factor for violence in mentally ill offenders, was also perceived as important for assessing risk in the current study. Insight has been studied in relation to violence risk, particularly in individuals with schizophrenia, where it has been established that individuals with schizophrenia have limited insight and may have difficulties in recognising their need for treatment, as well as their own impairments (Ekinci & Ekinci, 2013). Lack of insight is thought to contribute to violence risk in different ways, for instance, individuals with delusions or hallucinations may doubt the reality of their symptoms, consequently avoiding or refusing treatment, which may lead to worsening or relapse of symptoms. Additionally, lack of insight into psychotic symptoms, may increase an individual's conviction in delusions, which may increase violent propensities (Ekinci & Ekinci, 2013), though violence is still rare amongst this population. Although clinical and cognitive insight were defined in the questionnaire and are differentiated in the literature and on most assessments, few respondents appeared to consider their influence to differ until the final round. Nevertheless, cognitive, and clinical insight are included as risk factors in the HCR-20v3 (Douglas et al., 2013). Lack of insight as a construct, has been examined in prospective studies of violence risk, with correlations ranging from 0.18 to 0.37 (approximately 3-6% shared variance) (see Chapter 4), and although small, these effect sizes are consistent with cognitive domains in the wider forensic literature (e.g., Reinharth et al., 2014).

Taken together, the three rounds of the Delphi study were successful and informed the aim, that is to obtain a core list of cognitive domains deemed, by experts in the field, to be important for consideration in risk assessments, irrespective of whether they have yet been examined in the literature. Further, in round three, phase two, panellists were shown results from the previous round, in addition to meta-analytic findings, for intelligence, attention, processing speed, insight, and impulsivity, and then were given the opportunity to change their rankings if they chose, however, only intelligence moved up into the top 10, and attention and processing speed remained quite low, indicating that current research on this topic may be underutilised. The underutilisation of research amongst some professionals is not uncommon. For example, in an international survey on risk assessments, Singh et al. (2014) found that despite strong supporting evidence for better reliability and validity of structured over unstructured risk assessments, clinicians in certain countries continue to assess risk using only unstructured clinical judgement. Further highlighting the lag in translating research into practice.

5.4.3 Limitations

The way in which *experts* were defined and the complexity of the subject presented limitations. Delphi panel members were chosen for two reasons, a) they were active researchers in the field as evidenced by publishing a relevant paper between 2010 and 2016, and b) they were clinicians who specialised in both forensic and neuropsychology. To increase the generalisability of the results, the inclusion of an international sample was sought. The choice to contact corresponding authors may not always have succeeded in identifying those most expert on the topic, however participants were asked to only participate if they had the specified expertise, and as clinician experience (in a dual role) ranged from approximately 22 to 24 years in rounds two and three, researchers experience (in a dual role) ranged from approximately 14 to 22 years, and those in a research role only had experience ranging from approximately eight to 10 years, suggests that this method was successful. Moreover, the literature suggests that panel members responding to the questionnaire should have a good working knowledge in the appropriate area (Hanson & Ramani, 1988), but in order to gain consensus or answer the research question, a high degree of expertise is not necessary (Armstrong, 2001; Welty, 1972). Further, practicing clinicians were invited to take part in this study as a way to gain insight into cognitive domains observed to be relevant to practice, and that may be less researched in the literature. This may have added to the limitations as all the domains were identified through a literature review, and some of the terms may only be differentiated in the research literature, rather than clinical settings. This was highlighted by the panel, some of whom queried the research derived definitions that were provided, indicating they were too technical. Furthermore, all individuals who reported being clinicians only, dropped out of the study after round one, and although this may have had to do with time constraints, it may also have been due to the complexity of the terms, albeit they were all accompanied by definitions. Additionally, although significant, a strong level of consensus was not reached when analysed using Kendall's *W*, however, it did increase as the rounds continued, standard deviations gradually decreased, and the top items did not significantly change throughout. This is likely due to the panel's heterogeneity, including both researchers and clinician-researchers which conversely brought the advantage of a more comprehensive appraisal. Moreover, a minority indicated that in practice their rankings would alter according to offender type making the choice of definitive ranking difficult. Also, although the response rate for round one may be perceived as poor, the response rates for subsequent rounds grew to reach close to 80% for the final round. Poor response rates in round one may be attributed to the time that the questionnaire was sent out, where a majority of the sample was from the United Kingdom and the United States, and August is a popular time for summer holidays, whereas subsequent rounds took place during term time. Furthermore, the time commitment required to participate in a Delphi study, 15 to 20 minutes per round in this case, can be particularly onerous for clinical professions. Finally, as most of the sample was from the UK and the US, the findings may be less generalisable, and likewise, as these findings are based on expert *opinion* the findings may also be biased.

5.4.4 Implications

The development of an expert consensus on the cognitive domains relevant to offending has utility in clinical practice. First, it provides clinicians with the guidance for integrating the measurement of cognitive impairments into their own risk assessments, and second, it provides researchers with a list of cognitive domains that experts in the field consider relevant for future research in the field of violence risk. These findings can be valuable for informing assessments, treatments, and formulations of risk. Future research should move toward identifying a) the predictive utility of the domains, b) the assessments which most accurately and robustly measure these domains, and c) the comparative predictive validity of the domains in differing offender populations (e.g. adolescent, forensic mental health, general psychiatric populations, incarcerated adults). The implementation of cognitive testing used alongside existing risk assessments, particularly in individuals with known risk factors, such as TBI, may have the ability to increase the specificity and sensitivity of existing risk assessments resulting in increased predictive power, more targeted treatment programmes, and the management of violent and antisocial behaviours.

5.4.5 Recommendations for Future Delphi Studies and Research

The Delphi method was successful in achieving the goal of this study and will serve as a useful technique for future studies in this field where a consensus is required. Recommendations for enhancing future Delphi studies in this field include, adding a round before the first round of questioning to identify a sample of experts who are interested in the study and who agree to remain in the study until it is complete to reduce attrition. Researchers should choose a survey software program which automatically analyses simple statistics to increase efficiency and decreases time between rounds in an effort to cut down on attrition and depending on how much input the research question requires; researchers may consider allowing the panel to answer open-ended questions in order to get a better idea of their opinions and ideas. Finally, in a field of research where there is significant heterogeneity, due primarily to methodological differences and the number of neuropsychological measures used, exploratory research methods like the Delphi method should be encouraged and are recommended. The method serves as robust way to gain insight on research topics and may assist in refining research questions or methods which may contribute to the increased quality of future research in this field.

5.5 Conclusion

The current study implemented the Delphi method to achieve an expert consensus on the cognitive domains posited to be the most relevant to violence risk. The results are in line with literature in this area, where impairments, particularly in executive functions and those which are closely related, such as impulsivity, in addition to insight, are hypothesised to be the most relevant to violence risk. These findings also have practical applications, in that they present a starting place for future research, as well as for practicing clinicians. As the research findings in this field are heterogeneous due to the variations in methodologies between studies, these findings may aid in the elimination of less relevant cognitive domains, to not only cut back on the number of domains that are being examined, but to also encourage replicability of existing studies on these risk factors so the field can reach a definitive answer on whether behaviourally measured cognitive domains add value to existing risk assessments. Finally, it can be concluded that the assessment of cognitive functioning is indeed viewed as relevant to violence risk by experts, thus, their further examination in future, quality studies is encouraged.

6 Assessing the Cognitive Contributors to Violence: Part I

Chapter Preface

This chapter sets out the approach and design of a feasibility and pilot study to examine the neuropsychological contributors to violence risk. A prospective, observational approach was planned with a retrospective component. The methodological considerations were based on the criticisms of this type of research outlined in Chapter 3, and the neuropsychological battery and measured functions were developed based on the findings from the systematic review and meta-analysis (Chapter 3) and the Delphi Study (Chapter 4). The primary aim of this study was to investigate the feasibility of using a neuropsychological battery to aid in the identification of cognitive impairments relevant to violence risk, and the assessment of violence risk in an inpatient and community setting, and within that, to pilot a neuropsychological battery of measures thought to be relevant to risk factors for violence. This study is presented in four parts, this chapter (Part I) discusses methodology of the study, and the following chapters discuss the feasibility results (Part II), as well as the results for the primary and secondary hypotheses (Part III). The final part (Part IV; Chapter 9) presents the discussion and conclusions for this study.

6.1 Methodological Considerations

6.1.1 Rationale for Pilot and Feasibility Design

Pilot and feasibility studies are useful for developing a preliminary understanding of the research questions at hand. Conducting pilot studies can answer questions about feasibility, recruitment potential, and patient acceptability, as well as, sourcing a justification for larger, future studies, and can provide best practice guidelines for logistical issues for a larger study. Further, conducting a pilot before the main study can enhance the likelihood of success of the main study (Thabane et al., 2010). Due to the current state of the literature and the inconsistencies seen across studies, a small- scale pilot and feasibility study was necessary in order to identify the most appropriate measures and procedures, and to estimate effect sizes. The study aimed to identify neuropsychological measures which are sensitive and specific to violence and are tolerable to a population of individuals who are known to have a range of cognitive impairments. Further, from the beginning it was acknowledged that recruitment would be a challenge, especially in the presence of a two-hour long assessment period, thus, testing the feasibility of this was of utmost importance. Ultimately, as stated in Chapter 3, a possible explanation for the methodological weaknesses repeatedly observed in this type of research is that researchers may not be paying close attention to the basic, but essential elements of scientific research, therefore, this study was designed with close attention to detail, to develop a strong and rigorous foundation for a future, larger study.

6.1.2 Study Design

When examining cognitive abilities and violence risk, many studies have utilised cross-sectional, retrospective, and prospective study designs. As it was the aim of this thesis to explain the variance in future behaviour, this study followed an observational prospective design. However, due to the potential for a low base rate of violent incidences during the follow-up period, secondary and proxy outcomes were also investigated, including a retrospective component. Retrospective data were collected for the 6 months prior to consent, and a self-report questionnaire, used as a proxy for violence, was administered to all participants at baseline. Additional secondary outcome measures included aggression, loss of institutional privileges, antisocial behaviour, and breaches of conditions in the community sample.

6.1.3 Sample

The study included an all-male sample. The rationale for this came from the notion that Scotland's high secure hospital treats males only, and women make up only a small proportion of patients in the other Forensic Network sites, thus, it was posited that introducing women may contribute to additional variance in the data. Likewise, only men with a violent charge or conviction were chosen to participate, as this pilot study is a steppingstone to develop a violence risk assessment tool, which are historically used only on offenders who are already known to behave violently.

Further considerations surrounding the sample for the current study queried which violent groups to include. In keeping with the overall aim of the study, the decision was first made to include forensic psychiatric inpatients with a diagnosis of a psychotic illness. The rationale for including this group of individuals was guided by the literature which demonstrated that forensic psychiatric patients are a diverse population of individuals, who often have high prevalence rates of aetiological risk factors for cognitive impairment, such as TBI and substance use, and many of them have a history of violent behaviour. Further, individuals who are diagnosed with psychotic illnesses, especially, schizophrenia, have evidenced distinct neurocognitive profiles (e.g., Goldberg & Gold, 1995). Following this decision, it was acknowledged that including an inpatient only sample, and as a result, examining only inpatient violence, may limit generalisability of the findings. Thus, to balance this out, the decision was made to include a sample of violent offenders who were service users of Criminal Justice Social Work, and who lived in the community.

6.1.4 Neuropsychological Risk Factors

The literature cited in Chapter 3 suggested that risk factors for violence should have a theoretical basis, otherwise, the research may lose direction, and without knowledge of the mechanisms which cause the behaviour, case formulations and treatment may not be successful. The entire thesis is based on converging findings from biological and neuropsychological literature supporting the structural and functional differences seen in violent and forensic populations relative to normal populations. As the focus of neuropsychology is the brain-behaviour relationship, neuropsychological risk factors can be conceptualised as behaviours or characteristics caused by underlying brain mechanisms, thus, for example, damage to the brain does not cause violence, but it can contribute to the loss of important cognitive functions, which in some situations, may result in an individual being unable to desist from offending behaviours. The aim of deriving theoretically grounded risk factors is that theory gives an explanation on *how* these risk factors cause or contribute to violence, rather than relying only on trends seen in research, for instance. Chapter 1 of this thesis outlined the underlying brain mechanisms of violence, as well as the neuropsychological characteristics seen in violent populations, Chapter 2 summarized the common aetiological factors for cognitive impairments, Chapter 4 evidenced a list of empirically derived neuropsychological factors associated with violence, and Chapter 5 provided a list of factors viewed by experts as being the most important for inclusion in risk assessments.

6.1.5 Measures

Considerations for measures to include in the current study were based on the length of time they took for administration, psychometric properties, and clinical utility. It was important that short or abbreviated measures were included, and those which allowed for breaks in between without affecting the results, and likewise, measures with strong psychometric properties were important to ensure the validity of the findings. Measures with clinical utility, as opposed to those used in research only, provide normative means which can be used to guide the interpretation of the mean scores found in the current sample, for example, comparison to norms allows for judgements to be made on whether the sample has means which are average, below average, or impaired. Moreover, measures with clinical utility may be more easily accessed or available in clinical settings, thus, it was expected that the implementation of these measures in the current study would more easily translate into clinical practice.

6.1.6 Outcomes

As evidenced in Chapter 3, there is a lack of consensus on how violent outcomes are operationalised and the lack of specificity in definitions are a lasting limitation in violence research. To avoid replicating these, the following a-priori decisions were made in relation to the outcomes: 1) the definition of violence would be specific to *contact* or *physical* violence toward others; 2) an objective tool to code severity of violent incidents would be implemented; 3) a tool to code the type of violence that occurred (reactive vs instrumental) would be included; 4) if there were sufficient data, it was planned to use continuous outcomes for the main analyses to avoid splitting the groups dichotomously (e.g., violent/non-violent). Secondary outcomes (e.g., loss of privileges, breaches of orders, anger, antisocial behaviour) and a proxy measure of violence administered at baseline, were also included to mitigate the risk of having a base rate of observations that was too low to analyse, and to ensure each participant had at least one completed outcome measure in the case of attrition. Although research suggests that individuals with more severe cognitive impairments are more likely to engage in reactive violence, research also suggests that individuals who engage in instrumental violence are more likely to have a higher IQ and higher cognitive functioning (Hoffer et al., 2018). Therefore, a tool to code for reactive and instrumental violence was included for two reasons: 1) To increase specificity in the findings leading to the ability to report the level of functioning which predicts a specific type of violence, and 2) To be able to identify cognitive similarities and differences between those who engage in reactive vs instrumental violence in the current sample. However, as careful consideration was given to the participant groups included in the thesis, for example, those who were more likely to have risk factors for cognitive impairments, such as psychosis, substance misuse, and traumatic brain injuries, it was expected that the samples were more likely to engage in reactive rather than instrumental violence. Though, based on the literature (e.g., Meloy, 2006; Levi et al., 2010), it was hypothesised that individuals in the reactive violence group would demonstrate impaired inhibitory control, attention, executive functioning, insight, and affect recognition, whereas the instrumental group may have less pronounced impairments, aside from high risk taking, and low cognitive empathy.

6.1.7 Risk Assessments

As the overarching aim of the thesis was to develop a supplemental SPJ tool that would encompass cognitive predictors of violence for use alongside of existing risk assessments to improve predictive validity, clinical judgements, and risk formulations, it was a priority to ensure that the tool would easily translate into clinical practice. In addition to this, it was necessary to be mindful of the time and resources that the current studies required, first and foremost for the participants, but also for the services which were assisting in recruitment. With acknowledgement of these two factors, and the notion that the tool was meant to be used alongside of *any* risk assessment, the decision was made to examine the risk assessment tools that were the standard in each service (e.g., the HCR-20v3 in the inpatient sample, and the LSI-R:SV in the community sample), and thus had already been completed. While the HCR-20v3 is a violence risk assessment used to predict future behaviour, it is also meant to structure clinical judgement and inform risk management, all of which the researcher hypothesised the assessment of cognitive functions could further assist in. Although the use of an actuarial tool such as the VRAG may have been viewed as more appropriate for solely predicting violence (a main aim of the thesis), it was not the standard used in the secure hospitals in NHS Scotland, thus it would have required completion in addition to the neuropsychological measures, requiring more time and resources. Moreover, measuring cognitive functioning for the prediction of violence only, in which static risk factors are often only assessed, would dismiss the cognitive measures' potential utility and benefits for clinicians when making clinical judgements, treatment plans, and risk management plans, as cognitive functions are dynamic in nature, and therefore amenable to treatment. In sum, aside from the HCR-20v3 being the standard in NHS Scotland forensic services, the assessment of cognitive functions alongside the assessment of other relevant risk factors would allow clinicians to draw on the additional information about cognitive functioning to increase the accuracy of their clinical judgement in addition to the tool's predictive accuracy. Similar to the circumstances with the HCR-20v3, the risk of the community sample was primarily assessed using the LSI-R:SV, which is an actuarial tool that includes both static and dynamic risk factors and informs the management of offenders. With acknowledgement that this tool was designed to assess general offending, as opposed to violence, additional non-violent outcomes were assessed in addition to violent outcomes in order to determine the predictive accuracy for all outcomes, violent and general (e.g., antisocial behaviour, anger, breaches).

6.2 Plan of investigation

The present study took place in two settings, inpatient secure hospitals (high, medium, low) and within Community Justice Social Work (CJSW). The same procedures were followed in all inpatient settings, and only small moderations were made for the social work setting which are outlined below. Data collection consisted of three main phases; phase one consisted of neuropsychological testing of participants, phase two consisted of case-note review to examine risk and violence-related outcomes, and in phase 3, all follow-up data were collected.

6.2.1 Pre-collected Data Approval Requirements

Following approval from the State Hospital (TSH) Research Committee in June 2017, full ethical approval from the West of Scotland Research Ethics Committee 03 (WoS REC 3) was granted on 1st September 2017. Final management approval was granted on 15th September 2017. Approvals for each site were granted subsequently. Ethical approvals can be found in Appendix G, Appendix H, Appendix I, Appendix J, and Appendix K.

6.2.2 Identification and Recruitment of Participants

The inclusion and exclusion criteria for both groups of participants are presented in Table 6.1 and are followed by a rationale for the criteria.

Table 6. 1 Inclusion and Exclusion Criteria

Group	Inclusion	Exclusion
Inpatients	• Male • 18-60 years old • Completed admission case review • Not being considered for transfer • Can give informed consent • English as their first language • Diagnosis of a psychotic illness*	• Patients who have intellectual disabilities • Non-violent/non-contact sex offences or other offences without the presence of a violent offence • Disabilities which may impede their ability to engage in the assessment process (e.g., significant hearing, sight, and motor impairments)
Community	• Male • 18-60 years old • Can give informed consent • English as their first language • Be under licence and therefore in the care of criminal justice social work for at least 6 months after testing • Has been convicted of a violent offence.	• Participants who have intellectual disabilities • No conviction of a violent offence • Disabilities which may impede their ability to engage in the assessment process

Note. *All diagnoses were made by responsible medical officers/treatment teams.

6.2.3 Rationale for Inclusion/Exclusion Criteria

The age range for included participants was limited to 18-60 years old as the normative data for The Awareness of Social Inference Test (TASIT) only includes individuals up to 60 years of age, and all were 18 or over as Scotland currently has no forensic inpatient services for children or adolescents. Completed admission case reviews were required, as this indicated that the participant had a completed risk assessment. It was also decided to exclude patients on the transfer list, and community offenders who were almost done with their time as a service-user to decrease the number lost to follow-up. English language was a requirement, as many cognitive measures are normed in populations whose primary language is English. Finally, participants with intellectual and physical disabilities were excluded as research shows that individuals with intellectual disabilities may have different or additional risk factors for violence (e.g., Hounsome et al., 2018), and certain physical disabilities would make the successful completion of some tasks difficult.

6.2.4 Recruitment

NHS sites. Due to the protection of patient confidentiality, the researcher was not allowed to view patient files to examine inclusion criteria, and therefore needed to rely on psychology teams at each site to gather information. For each site, the researcher first attended a psychology team meeting or met with a lead psychologist to discuss the project. At TSH, a psychologist from each ward was contacted to ask if they would agree to assist the researcher in the identification and recruitment of patients for the study, and in all other sites, one psychologist was appointed to the researcher as the main contact. After contacts were confirmed, letters explaining the study (Appendix M) were emailed to responsible medical officers (RMOs) to identify patients who did not have the capacity to consent to research. RMOs were instructed to email the psychology contact with names of patients who did not have capacity and therefore would not be asked to take part in the study. Once psychology contacts had the names of patients, the researcher set up meetings to anonymously discuss patients who fit the inclusion criteria, and the number of patients and reasons why others did not meet criteria were recorded. Once a list of patients was identified, psychologists approached patients to invite them to take part in the research, and following agreement, the researcher approached patients and went over the study in detail using the information sheet, also ensuring they had an opportunity to ask questions. If the patient agreed to take part, an appointment was made for a period of no less than seven days, allowing the patient to re-read the information sheet and to further consider taking part. After the seven-day period, the researcher approached the patient again and asked if they would still like to take part, if the patient agreed, consent was taken, and assessments commenced immediately.

For all NHS sites, participants were asked to consent to the following (Appendix O):

- Participate in the study (undertaking neuropsychological assessments).

- Researcher access to their medical files/records in paper and digital format.

Social Work Sites. Participants from CJSW were primarily recruited from two areas of social work, the community intervention team, and the unpaid work team. The community intervention team works with individuals and families involved in CJSW, and the unpaid work team organizes and supervises offenders who are sentenced to complete hours of community service.

First, the community intervention team invited the researcher to a team meeting to introduce the study and discuss details with the social workers. Like the NHS sites, the researcher was not allowed to approach service users before their social worker discussed the study with them, and they needed to be introduced by a common party (social worker). When the social work team identified a service user who fit the criteria, they contacted the researcher and set up an initial meeting to discuss the study with the service user.

The researcher also met with the unpaid work teams to introduce and discuss the study. The same ethical procedures were followed as outlined above. However, the unpaid work team inducts new service users weekly, and the first part of their order is to attend an unpaid work induction, which the researcher also attended to explain the study to potential participants. Additionally, in the unpaid work team, it is common practise that all service users are offered an opportunity to participate in 'Other Activities' in addition to their unpaid work to learn new skills and continue to work off their unpaid work hours. As part of this initiative, the sector manager of the unpaid work team offered to list this research project as an option for 'Other Activities' as it was seen as a way for service users to volunteer their time, learn about themselves through the results of the measures, and as an incentive to work off their unpaid work hours quicker. This also resulted in an increase in recruitment. Once this was added, the researcher attended unpaid work inductions, and gave a short presentation to possible candidates. After the induction, service users who were interested in hearing more about the study, and who met the inclusion criteria, met with the researcher, were given an information sheet and shared their contact information. Candidates were then followed-up to see if they would like to make an appointment to participate in the study.

Shortly after commencing community recruitment, the decision was made to introduce a monetary incentive of £20 to reimburse participants for their time and to encourage participation. This was accompanied by appropriate ethical approvals (Appendix L). The addition of the monetary incentive significantly lifted recruitment numbers.

For CJSW sites, participants were asked to consent to the following:

- Participate in the study (undertaking neuropsychological assessments).
- Researcher access to their medical/social work files/records in paper and digital format.
- Researcher requests a list of offences during follow-up periods.

6.2.5 Research Settings

High Secure. The State Hospital (TSH) is the only high secure psychiatric facility for Scotland and Northern Ireland and is one of four high secure hospitals in the UK. The hospital has 140 beds for male patients which require maximum secure care, and 12 beds for patients with a learning disability. The aim of the hospital is to care for patients, and to protect the public, staff, and patients from harm. It provides assessment, treatment, and care in 'conditions of special security' for individuals with mental illness who cannot be cared for in any other setting due to dangerous, violent, or criminal tendencies, which is defined under the NHS Act 2006. Patients are admitted to TSH under the Mental Health (Care and Treatment) (Scotland) Act 2003/2015 and related legislations.

Medium Secure. In 1999, it was proposed that medium secure care centres be developed across Scotland to bridge the gap between high secure and low secure care units. This was in conjunction with a new Mental Health Act which required NHS boards to ensure that patients were not being held at a higher security level longer than was required. There are three medium secure care units in Scotland, namely, the Orchard Clinic in Edinburgh, Rohallion Clinic in Perth, and the Rowanbank Clinic in Glasgow. The Rowanbank Clinic was the only medium secure care clinic which granted access to the researcher. The Rowanbank Clinic serves both men and women and has eight wards, five of which are male only, and a total of 74 beds.

Low Secure. Low secure units are for patients who have complex issues and cannot be safely cared for in acute inpatient wards (The NHS Confederation, 2012). Like high and medium secure care, patients are detained under the Mental Health Act and require rehabilitation. The two low secure units that patients were recruited from were, Radernie Ward in Fife, Scotland and Leverndale in Glasgow, Scotland. Radernie Ward is a small unit with 10 male beds, and Leverndale has four forensic wards (one of which is female), and 49 beds (5 female, 8 learning disabilities; NHS Quality Improvements Scotland, 2010).

All inpatient assessments were administered in a quiet room attached to the hospital wards.

Criminal Justice Social Work (CJSW). In the 1960's, Scotland became distinctive in their arrangements for the community supervision of individuals convicted of offending behaviours (Daniel & Scott, 2018). Instead of having a separate probation service, like England, the 1968 Social Work (Scotland) Act brought probation within local authority social work services for the first time. CJSW services act on behalf of the Scottish Courts to supervise offenders age 16 and older who have been made subject to a community disposal, provide supervision for offenders released from prison, and provide reports to courts to aid in decisions around sentencing (Association of Directors of Social Work, 2018). Local authorities are responsible for providing CJSW services, including assessment reports, supervision and monitoring and provision of unpaid work, in addition to assessing the risk of offenders in the community and custody. Since the employment of the Management of Offenders (Scotland) Act 2005, there are eight Community Justice Authorities, which ensure support between prisons and community services to aid in the reduction of offending and rehabilitation of offenders (Association of Directors of Social Work, 2018). While CJSW is made up of numerous different services and groups, service users for the current study were only recruited if they were in the care of the Community Intervention Team or if they were sentenced to complete an Unpaid Work Order. All assessments were administered in a quiet room within the social work offices.

6.3 Procedures

All neuropsychological assessments, case-note data collections, data management and analysis, and dissemination of results were completed by the researcher (PhD student).

The neuropsychological battery took approximately two hours to complete and was most often administered to the inpatients in two 1-hour sessions to avoid testing fatigue. Breaks were offered to participants throughout to ensure that they stayed well engaged throughout the assessment process. Where patients were unable to focus for a lengthy period, additional sessions were offered.

The community testing procedure was broken down differently due to concerns from the social work sector managers that 2 hours of testing was too taxing for the population. Therefore, the procedure was broken into 2 phases:

- **Phase 1**: 30-40 minutes of testing.
- **Phase 2**: After phase 1 testing was completed, the participant was given the option to opt into the second phase of the study which consisted of further cognitive testing lasting approximately 80-90 minutes.

Following the completion of assessments, data were collected through case-note and file review. Data were collected in a comprehensive manner using a case-note review checklist (Appendix P) and was used for the purpose of describing the sample. Historical and risk-related data, and a record of potential contributors to the aetiology of cognitive impairments were recorded from participant files where they were available (Table 6.2).

Table 6.2 Data Collected from Participant Files

Historical	Risk-Related	Potential Contributors [a]
• Age, ethnicity, education level/attainment • Date of current admission/ became service user • Source of current admission/services • Legislation for current detention (inpatients only) • Current conviction status (if applicable) • Primary and secondary diagnosis (if applicable) • Year of first diagnosis with a psychotic illness (inpatients only)	• Reason(s) for current admission/services • Offence(s) or alleged offence(s) leading to current admission/ services • Violent incidences during current admission/ services (if applicable) • Total number of previous convictions • Most serious previous offence • Previous types of offences	• Birth trauma • Abnormal infant development • Childhood history of physical or sexual abuse or neglect • Diagnosis/history of drug or alcohol dependence/misuse • Abnormal infant development • Neurological injuries or neurologically relevant diagnoses

Note. [a] This information was only available for a subset of community offenders.

Additional data collected included:

6.3.1 Violence and Offending Proneness

The retrieval of routinely administered assessments (where available) included, the Historical, Clinical, Risk Management-20 Version 3 (HCR-20v3), (Douglas et al., 2014) as a measure of violence proneness, and the Level of Service Inventory-Revised: Screening Version (LSI-R:SV) (Andrews & Bonta, 1995) as a measure of offending proneness. The HCR-20v3 is composed of 20 items, which are broken into three scales. The Historical scale comprises 10 static risk factors associated with violence that are not expected to change over time, and the Clinical and Risk Management scales include five dynamic risk factors each, which are expected to change over time. Items are rated 0 if the risk factor is not present, 1 if it is partially present, or 2 if it is definitely present, items are summed to obtain a total score for each scale, and a then a total score for the entire measure (Douglas et al., 2014). The LSI-R:SV is composed of eight items, namely, two or more prior convictions, arrested younger than 16 years of age, currently unemployed, criminal friends, alcohol, or substance use problems that interfere with school or work, whether a psychological assessment has been indicated, non-rewarding parental relationships, and attitudes or orientations which are supportive of crime. The items are scored as 0 if the risk factor is absent, and 1 if it is present, and then all items are summed to obtain a composite risk score. Risk/needs scores are interpreted as, low 0-2, medium 3-5, high 6-8 (Andrews & Bonta, 1995).

6.3.2 Neuropsychological Battery

The development of the battery followed the National Institute of Mental Health's MATRICS initiative's development of the MATRICS Consensus Cognitive Battery (MCCB), which sought to identify not only domains identified by the research literature, but also those that have yet to be the subject of sufficient research, through expert consensus (Nuechterlein et al., 2008). Using the findings from the Delphi study and meta-analysis, a list of neuropsychological functions relevant to violence risk was developed for further investigation. The functions of interest included, reasoning, inattention, response inhibition, response monitoring, risk taking, affect recognition, cognitive empathy, and lack of insight (inpatients only). A description of each measure, as well as their reliability and validity in the context of prior research are detailed below. All measures were administered according to the standardization guidelines set out in the measure manuals.

Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II; Wechsler D., 2011)**.** The WASI-II is an assessment of intelligence comprised of four subtests and takes an average of 30 minutes to administer. Subtests include block design, vocabulary, matrix reasoning, and similarities. It provides index scores that estimate Verbal Comprehension (VCI) and Perceptual Reasoning (PRI), and two composite scores that estimate intellectual ability- Full Scale IQ (FSIQ)-four subtests and FSIQ-two subtests. Reasoning was identified in the meta-analysis and Delphi study as an important function to investigate and was operationalised by the Perceptual Reasoning Index. Full-scale IQ and verbal comprehension were included as possible covariates.

Psychometric Properties. The WASI-II demonstrates high reliability coefficients for both subtests and composite scores, where the average reliability coefficient for subtests range from .90 to .92, and from .92 to .96 for VCI and PRI, and from .95 to .97 for FSIQ-4. These coefficients are not only consistent with those of the WAIS-IV, but also suggest very minimal measurement error. The FSIQ-4 score is also highly correlated with the WAIS-IV FSIQ-4 score, accounting for 85% of the variance. In addition to exceptional reliability, the WASI-II has also evidenced validity based on the scales' internal structure. The intercorrelations for the WASI-II were found to be similar to the WAIS-IV and WISC-IV, indicating support for the *g* factor, and providing evidence of both convergent and discriminant validity (D. T. Campbell & Fiske, 1959). Additionally, the WASI-II has been shown to be valid in special groups, two of which were of interest for the current study, namely those diagnosed with attention deficit hyperactivity disorder (ADHD) and a traumatic brain injury.

Conners' Continuous Performance Test -Version 3 (CPT-3; Conners, 2015)**.** The CPT-3 is a computerised, task-oriented measure of inattentiveness, impulsivity, sustained attention, and vigilance. It takes 14 minutes to administer and is valid for administration to individuals aged 8 years and older. During this task, letters are presented one at a time on the screen and participants are instructed to press the space bar every time a letter flashes on the screen except when the letter "X" appears, when the letter "X" appears, individuals are instructed to wait until the next letter appears. The amount of time between the appearance of the letter, known as inter-stimulus intervals (ISIs) are 1, 2, and 4 seconds with a display time of 250 milliseconds (Conners, 1995). There are 6 blocks with 3 sub-blocks which consist of 20 trials each. Key areas which are measured in the CPT-3 are, administration validity, response style, clinical likelihood, dimensions of attention, detectability, error rates, and reaction time statistics. All participants first took a practice test to ensure they understood the instructions, followed immediately by the full task. The CPT-3 is automatically scored using computerised scoring software and generates an assessment report with gender specific norms; higher scores reflect higher impairment.

Scoring for individual variables:

- <u>Omissions</u>: when the target is presented, and the participant does not respond. High omission error rates indicate that the participant was not responding to the target stimuli and are a good indicator of inattentiveness.
    - For the current study, inattention was operationalised as omissions.
- <u>Commissions</u>: the participant responds when the target is not present. High commission error rates indicate impulsiveness.
- <u>Perseverations</u>: are responses made in less than 100 milliseconds following the presentation of a stimulus and can present as a repeated response. In this task, perseverations are an indicator of either impulsiveness or a very liberal response style.
- <u>D-prime</u> (*d'*): measures how well the participant discriminates non-targets from targets.

Psychometric Properties. The CPT-3 demonstrated good split-half reliability for all scores in the adult normative sample, ranging from .73-.99, and .79-.99 for the clinical sample, and demonstrated the ability to produce significantly different scores between clinical and normative samples, as well as standard error scores, all of which confirm psychometric soundness.

Stroop Colour and Word Test (SC&WT; Golden, 1978). The SCWT is a widely used measure used to examine executive functions such as concentration effectiveness, response inhibition and mental flexibility. There are three parts to the task, each of which only lasts 45 seconds. For the first part, participants are given the Stroop record form, which has the words, "red", "green" and "blue" printed in black ink, in 5 columns of 20 words, totalling 100 words on the entire sheet. Participants are instructed to read the words aloud going down the columns (as opposed to across or jumping around), as quickly as they can until they are told to stop. If the participant reaches the last word on the sheet (the 100th word), and they were not yet told to stop, they are instructed to start from the beginning until the 45 seconds is up. The second part of the task is made up of non-word text (e.g., "xxxx") printed in red, green, and blue ink. For this portion of the task, participants are instructed to read the colours aloud as quickly as they can until they are told to stop. In the final part of the task, participants are given a sheet with the words "red", "green", and "blue" printed in an ink colour that does not match the word (e.g., the word "red" is printed in the colour blue). Participants are instructed to name the colour of the ink that the words are printed in, ignoring the word that is printed. Again, they are instructed to read aloud as quickly as they can until they are told to stop.

The Stroop produces three basic scores and an interference score:

- Raw Word score: the number of words read correctly on the word page.
- Raw Colour score: the number of colours correctly read on the colour page.
- Raw Colour-Word score: the number of items correctly read on the colour-word page.
- Interference score: Colour-Word score minus the predicted Colour-Word score.
    - Response Inhibition was operationalised as the interference score.

For the three basic scores, high scores reflect better performance (e.g., T-scores above 40), and for the interference score, lower scores (e.g., a T-score of 40 or less) often indicates that there is a problem.

Psychometric Properties. The Stroop has been shown to have good test-retest reliability for the three basic scores, .86, .82, and .73, and for the Raw Interference score, which is in the .70 range and is equivalent to the Raw Colour-Word score. This task has also been validated in several clinical samples including brain injury, dementia, Turner Syndrome, schizophrenia, and depression.

Modified Wisconsin Card Sorting Test (MWCST; Nelson, 1976). The MWCST is used to measure individuals' ability to infer abstract rules, to shift and maintain set, and utilize feedback. It is a variation of the original Wisconsin Card Sorting Test (WCST; Heaton, 1981), which is based on a card-sorting task developed in by Grant and Berg (Grant & Berg, 1948). The original WCST consists of four key cards and 128 response cards. Participants are handed a response card one at a time and asked to match it to one of the key cards, they are only told if they are correct or incorrect but are never told the correct sorting category. The MWCST eliminated 80 of the 128 response cards that share more than one characteristic with a key card, resulting in 48 response cards (Nelson, 1976).

There are several advantages to utilising the modified version of the WCST, especially for research. First, the MWCST takes significantly less time to administer (approximately 7-10 minutes), which was a main concern in the developmental stages of the current study. Second, in the original WCST, some response cards share more than one characteristic with a key card, resulting in some responses being coded as both correct and as a perseverative error. Because the MWCST removed response cards which share more than one characteristic with a key card, it eliminated the risk of recording both a correct response and an error simultaneously and resulted in unambiguous scoring on the MWCST. As an additional result of excluding cards which share characteristics, the administration and scoring are more easily understood, and the task is less frustrating for individuals with cognitive impairments or older populations (Nelson, 1976). Lastly, the scoring of this version only takes approximately 2-3 minutes to complete.

Four main scores are derived from the MWCST:

- <u>Number of Categories Correct</u>: the number of completed trials where 6 cards were correctly sorted.
- <u>Number of Perseverative Errors</u>: the number of cards which were sorted either according to the previous rule after the participant was told that it changed or sorting cards according to the same rule even after receiving feedback that it was incorrect. These are identified during administration and are circled and marked with a "P".
    - Response monitoring was operationalised by number of perseverative errors.

- Number of Total Errors: is the number of errors, including perseverative errors, made by the participant. Errors are circled during administration.
- Percent of Perseverative Errors: the number of perseverative errors divided by the number of total errors and multiplied by 100.

Raw scores are converted to a scaled score and then into a T-score based on age and education, T-scores are also converted to percentiles and are given a qualitative label ranging from extremely superior to extremely abnormal. In addition to the main scores derived, there is also standard score for executive functioning. All T-scores have a mean of 50 and SD of 10, with higher scores indicating better performance, and standard scores have a mean of 100 and SD of 15.

Psychometric Properties. The MWCST demonstrated good reliability over a 2-year test-retest period (Number of Categories Correct = .65, Number of Perseverative Errors= .64), and the Executive Function score, using a subsample of *n*=103, had a test-retest coefficient of .50 over 5.5 years (Lineweaver et al., 1999). The measure was also shown to be positively correlated with several measures of cognitive abilities, demonstrating good convergent and discriminant construct validity as a measure of executive functioning. The MWCST has also been shown to be valid as a clinically sensitive instrument in patients with executive dysfunction and is validated in clinical samples including those diagnosed with schizophrenia (*n*= 110), bipolar disorder (*n*=127), dementia (*n*=176) and other mixed clinical samples (*n*=65).

Iowa Gambling Task Version 2 (IGT-2; Bechara et al., 1994). The IGT-2 is used to assess real word and risky decision making mediated by the prefrontal cortex. It has been used reliably in aggressive, schizophrenic, obsessive-compulsive, and ADHD populations, among several others. It is a computerised task which replicates real-life decision making, and closely resembles real world situations (Bechara et al., 1994). It takes approximately 10-15 minutes to complete and finishes after the participant has clicked on 100 cards. Participants are instructed to try to win as much money as they can or to try to lose the least amount that they can. They are put into a situation designed to parallel real decision-making, which may lead to risky choices/decision making. Participants are presented with four decks of cards labelled A, B, C, D, where decks A and B are disadvantageous, as the immediate win is large, but it is followed by a high penalty, whereas decks C and D are advantageous, as immediate wins are smaller, but losses are smaller too. The measure is scored using the computerized software, and T-scores were provided for evaluating normative scores. T-scores of 39 or less were considered impaired, 40-44 fall in the below average range, and scores of 45 and above indicate no impairments.

Primary Scores:

- Total Net Score (Net Total): is calculated by taking the difference between the total number of cards selected from decks C and D (advantageous decks) and the number of cards selected from decks A and B (disadvantageous decks), e.g., Net Total= [(Deck C + Deck D) – (Deck A + Deck B)]. A positive net score demonstrates that decision-making was advantageous, and a negative score indicates that decision-making was disadvantageous/risky. Scores are interpreted using T-scores.
    - Risk taking was operationalised by Net Total scores
- Block Net Score: is calculated in the same way as Net Total, except it is calculated for each block. To find this score, the 100 card selections are divided into five blocks of 20 cards each, and a net score is calculated for each block. Like Net Total, negative scores indicate disadvantageous/risky decision-making, and positive scores indicate advantageous decision-making. Scores are converted to T-scores for interpretation.

Psychometric Properties. Studies examining the validity of the IGT have compared its performance to measures of executive functioning, such as the WCST, the National Adult Reading Test (NART) and the WASI. Results in typically developing adults have shown associations between specific executive functions, such as categorization, set-shifting, and cognitive flexibility, but not with general intelligence (Brand et al., 2007). Studies have also compared this measure to several other executive functioning measures, namely, Trail-Making Test B (TMT-B; Reitan, 1971); Stroop tasks (Stroop, 1935), CPT-II (Conners, 2015), and the Porteus Maze (Porteus, 1965). When compared to risk-taking tasks, significant correlations were found in a low impulsivity subgroup using the Balloon Analogue Task (BART; Lejuez et al., 2002), and on the N-back task (Kirchner, 1958) between the 2-back measure and the IGT Total Net Score (Bechara et al., 1994). Lastly, when compared to self-report measures, there were notable associations between the IGT performance and executive functioning abilities and impulsivity. Taken together, the evidence presented suggests that the IGT has adequate construct validity. The IGT has been found to have a test-retest reliability of $r= .60$ in one session (Lejuez et al., 2003).

The Awareness of Social Inference Test (TASIT; McDonald et al., 2003). The TASIT provides a systematic examination of social perception. It was developed to assess one's ability to read and integrate social cues. It is comprised of three main parts; Part One: Emotion

Evaluation Test, Part Two: Test of Social Inference (Minimal), and Part Three: Test of Social Inference (Enriched), however, only parts one and two of form A were administered for this study. Part One of the measure tests the ability of the participant to recognise the emotions of other people and Part Two examines the ability of the participant to determine the intention, attitude, and meaning of the speaker; both require the participant to watch short videos.

Administration and Scoring

Part 1: For this part of the measure, participants are given a practice card that has seven emotions on it in a random order (e.g., surprised, anxious, happy, revolted, angry, neutral, sad). When the participant fully understands the instructions, they are given one of five response cards, all of which have the same seven emotions on them in a random order; for each video, the response card gets alternated (e.g., video one goes with response card one, video two goes with response card two and so on). If the video has two people in it, the participant is told which person to focus on. This continues until all 28 items have been completed.

Part 2: For Part Two, participants are told that they will be watching short videos lasting 15 to 60 seconds long, and that after each scene, they will be asked to answer four simple questions. At this time, they are shown a response card that explains what each question will consist of:

1. The first question will ask what the participant thinks someone is *doing* to the other person.
2. The second question will ask what the participant thinks someone is *trying to say* to the other person.
3. Third, the participant will be asked what they think the person is *thinking*.
4. Last, they will be asked to what they think the person is *feeling.*

After each video was complete, it was paused and the participant was asked the four questions, of which they can answer*, yes, no,* or *don't know.* During this part, the characters either exhibit sincere, simple sarcastic or paradoxical sarcastic exchanges, thus the participant needs to pay close attention to the intention of the speaker, otherwise the exchange will be misinterpreted. There are 15 items in this section of the measure.

Both parts of the measure were scored by adding up the correct answers for each section as well as a composite score for each part.

Primary Scores:

<u>Part 1:</u> The first part of the measure gives a sum of correct responses for each of the seven emotions (e.g., happy, surprised, neutral, sad, angry, anxious, revolted) as well as a sum score for positive emotions, negative emotions, and positive + negative emotions.

- Affect recognition was operationalised by the number of correctly identified positive + negative emotions.

<u>Part 2:</u> The second part of the measure derives scores for *sincere* (when a speaker means what they are saying), *simple sarcasm* (when a speaker means the opposite of what they are saying, and intends the recipient to understand the real meaning), and *paradoxical sarcasm* (when dialogue between two speakers does not make sense unless it is understood that one is being sarcastic), as well as, summed scores for what the participant though the actors were *doing, saying, thinking, feeling.*

- Cognitive empathy was operationalised by the feeling score.

Scores which fell in the lower 5th percentile when compared to norms were deemed to be impaired.

Psychometric Properties. The TASIT has demonstrated good test–retest reliability (r =.74–.88). Convergent validity has been shown through its association with measures of second-order Theory of Mind (r =*.68*) and emotion perception (r = .45) (McDonald et al., 2003).

The Ohio State University Traumatic Brain Injury Identification (OSU-TBI) (OSU-TBI; Corrigan & Bogner, 2007). The OSU-TBI is a standardized procedure for learning about a person's lifetime history of traumatic brain injury. This is administered as a brief structured interview and relies on a self-reported history of brain injury and is widely used in various professional settings. This measure was important for investigating the etiology of cognitive impairments. This measure is a structured interview that takes approximately 3 to 5 minutes to complete. Outcome categories include *Worst*: one or more severe TBI; *First*: TBI with a LOC before the age of 15; *Multiple*: two or more TBIs close together; *Recent*: a mild TBI in the last weeks or a severe TBI in the previous months, and Other Sources: any TBI which is combined with another way that may impair normal brain functioning (Corrigan & Bogner, 2007). Number of reported head injuries, number of head injuries with LOC, and severity of head injuries were used as covariates in the current study.

Psychometric Properties. The authors of the OSU-TBI identification method posit that the validity of this measure is not dependent on the participants' account of lifetime TBIs, but it provides summary scores for indices which relate to the likelihood that exposure to TBI have resulted in consequences. Interrater reliabilities for each of the interview questions was completed and resulted in intraclass correlations ranging from r = 0.84 to r = 0.95 (Corrigan & Bogner, 2007). The measure also demonstrated good test-retest reliability with intraclass correlations of greater than or equal to r = 0.60 (Bogner & Corrigan, 2009). Last, this measure also showed good predictive validity between lifetime history indices and cognitive performance measures, interpersonal functioning, and aggression (Bogner & Corrigan, 2009; Corrigan et al., 2012, 2013; Corrigan & Bogner, 2007).

Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). The PANSS is a structured clinical interview to assess the positive, negative, and general symptoms of schizophrenia. It is structured as a 30-item scale and evaluates the presence, absence, and severity of symptoms related to schizophrenia. The interview takes approximately 20 minutes to complete. This measure was used to control for symptoms of psychosis, symptom severity and as a measure of insight. Given its relevance only for those with psychosis, it was administered only to inpatient participants. The PANSS was a routinely administered measure at TSH to all patients with a diagnosis of a psychotic illness or those who had evidence of psychosis. When a recent PANSS was not available for a participant, the researcher completed the measure at the time of testing. This was the last measure that was administered in the battery of tests because of the importance of having a rapport with the participant before asking personal questions about their mental health. All items are scored on a 7-point scale (1= absent; 7= extreme) following the scoring guidelines where each item has a definition and a basis for rating.

Psychometric Properties. The internal consistency of the PANSS was demonstrated through significant correlations between the Positive and Negative scales and the total score, ranging from r = 0.62 to 0.70 (Kay et al., 1987). Test-retest reliability correlations ranged from 0.60 to 0.80 for the positive, negative, composite, and general scales, and demonstrated criterion-related validity indicating that the positive, negative, and composite scales are not influenced by extraneous variables (Kay et al., 1987).

Primary Scores:

<u>*PANSS total score*</u> was used to control for severity of mental health symptoms.

<u>*Positive symptom score*</u> was used to control for positive symptoms.

<u>*Lack of Judgement and Insight score*</u> was used to measure clinical insight (e.g., insight into one's mental illness).

6.3.3 Outcome Measures

Table 6.3 presents the primary and secondary outcomes accompanied by their operational definitions. Feasibility criteria are presented in the following section.

Table 6. 3 Outcomes and Operational Definitions

Outcome	Operationalisation
Primary Outcomes	
Violent Incidents	Institution-recorded incidents of violence will be recorded and triangulated with patient notes to ensure consistency. *"An intentional act of physical aggression against another individual that is likely to cause physical injury"* (Meloy, 2006, p. 539), *where the individual is the clear aggressor or instigator.* Mean number of incidents over 6 months.
Violent Offences & Charges [a]	Violent offences/ charges operationalised by the same definition as above. Mean number of offences/charges over 6 months.
Secondary Outcomes	
Level of Observation (perceived risk)	Daily mean level of observation over 6-months (183 days) prospectively and retrospectively (inpatients only).
Days on Enhanced Levels (perceived risk)	Mean number of days on enhanced levels prospectively and retrospectively over 6 months (inpatients only).
Length of Admission (perceived risk)	Time in months that a FPI has been in their secure environment.
Breaches of License/Conditions	Mean number of breaches of license/conditions prospectively and retrospectively over 6 months. Breaches were not limited to violence and aggression only (community only).
Antisocial Behaviour	Mean number of charges or convictions of non-violent/non-aggressive offences which occurred during the follow-up period in the community sample, over 6 months.
Anger	Mean number of non-contact anger provokedincidents/offences in both samples (e.g., verbal aggression, damage to property, threats, and intimidating behaviours) over 6 months.
Impulsivity [b]	Total score of BIS-11 (Patton et al., 1995)
Severity of violence & Aggression	Mean severity of all violent and aggressive incidents and offences combined over 6 months. Codes range from 0 (completely non-violent) to 4 (severe violence) (Gunn & Robertson, 1976; Robertson et al., 1987).
Reactive vs Instrumental Violence	Median score of all violent incidents and offences over 6 months. Codes range from 1= Purely reactive to 4= Purely Instrumental (Woodworth & Porter, 2002).

Note. BIS= Barratt Impulsivity Scale. [a] When reporting results, a clear distinction will be made between charges and convictions. [b] As impulsivity is already a well-established risk factor for violence, it will serve as a proxy outcome measure in the current study.

The *Gunn-Robertson Violence Scale* (Gunn & Robertson, 1976; Robertson et al., 1987) was used to objectively code prospective and retrospective violent and aggressive incidents. The rating for present incident or offence utilized the abbreviated 5-point scale to objectively quantify violent incidents. The points are as follows: 0 = Completely non-violent; 1 = Minimal violence (e.g. verbally aggressive, shouting or gesturing, even if this was not obviously directed at others); 2 = Moderate violence (e.g. attack on a person resulting in no serious injury, fighting or brawling, or damage to property when this was the main intent); 3 = Moderately severe violence (e.g. attack which resulted in serious injury, or damage to property which was extensive or which could have resulted in threat to life, 4 = Severe violence (victim died or life and health seriously endangered).

6.4 Statistical Methods

A study protocol outlining the primary and secondary hypotheses for this study was published prior to data analysis (e.g., Janes et al., 2021). All statistical analyses were performed using R: A language and environment for statistical computing, version 3.63 (R Core Team, 2020).

6.4.1 Feasibility Outcomes

Descriptive statistics were reported for feasibility outcomes. The predictive accuracy of cognitive measures was assessed utilising point-biserial correlations, Receiver Operating Characteristic (ROC) analysis, and area under the curve (AUC). For this analysis, any participant who was violent during the follow-up period were coded as 1, with 0 indicating the absence of violence.

Feasibility Criteria

Decisions to progress to a larger study were based on the following criteria:

1. Mean tolerability[6] for a single measure must not fall below 3.5. A score of 3 or below indicates that the battery is unpleasant for participants. Scores ranging from 7-4.0

[6] This method was adapted from the test selection phase of The MATRICS Consensus Cognitive Battery (MCCB) (Nuechterlein, et al., 2008). For the MCCB, tolerability was defined as, "the participant's view of the test," and can be influenced by how long the test takes, degree of difficulty, repetitiveness or anything that they feel makes the experience unpleasant (Nuechterlein et al., 2008) (The tolerability scale used for this study can be found in Appendix Q.

indicate extremely pleasant to neutral. If the mean tolerability rating for a single measure falls below 3.5, the reasons for this will be examined, and its replacement with a new measure for the same construct will be considered.

2. The mean recruitment rate for all sites combined must not fall below 30%. In the case that less than 30% of participants are recruited from all those eligible, inclusion criteria and recruitment procedures will be investigated. To keep the sample that is already recruited, only changes to the inclusion criteria which do not threaten the validity of the study will be considered for review. If this is not seen as a possible solution, the study will be deemed infeasible.

3. If the completion rate for the core measures (WASI-II, CPT-3, Stroop, MWCST, IGT-2, TASIT) does not reach 70%, reasons for non-completion will be reviewed and changes will be made to ensure the battery is feasible for these populations to complete.

4. If variables specified in our hypotheses do not show sufficient sensitivity, specificity, predictive accuracy or add incremental validity to existing risk measures, the measures used for this study will be re-considered using literature and further meta-analyses. Sufficient sensitivity + specificity was defined as 1.5 to 2 (Power et al., 2013), and sufficient accuracy and incremental validity were defined as AUC ≥ 0.70.

6.4.2 Primary and Secondary Outcomes

Power calculation. Considering the pilot and feasibility design of this study, a power analysis was not initially warranted. It was originally planned to recruit *n*= 80 participants (*n*= 20 per level of security and *n*= 20 from social work), as it was posited this number would provide adequate data to inform our feasibility and pilot outcomes. However, the recruitment uptake at medium and low secure units was low (e.g., *n*=6 per level of security), resulting in the successful recruitment of *n*= 32 inpatients, which led to increased recruitment in the community (*n*= 31). The final sample included *n*= 63 violent offenders. An ad-hoc power analysis was performed following this to ensure there would be adequate statistical power to allow for valid interpretations of the findings. The power calculation was conducted using G*Power 3.1 (Faul et al., 2007). To detect a medium effect size with 80% power and alpha set at 0.05 in a sample of *n*=32, a maximum of two predictors can be examined at one time, and to examine two predictors in the community group with the same power, a large effect size can be detected, to investigate primary and secondary hypotheses; consequently, the effects were not cumulative. Hypotheses were examined in each group separately.

Patient characteristics were reported using descriptive statistics. Both offending groups were compared on baseline performance using independent sample t-tests and one-way MANOVAs. MANOVAs were performed on each neuropsychological measure (e.g., one on the whole WASI-II, one on the whole IGT-II, etc.), and Bonferroni correction was performed to avoid multiple testing and to remain conservative. Multiple linear regressions were performed to examine primary and secondary outcomes, with the exception of reactive and instrumental violence, in which ordinal logistic regressions were planned to be used. If there were a sufficient number of recorded incidents or offences, the mean number of violent incidents/offences per month of the follow-up period were used as a continuous dependent variable. In the case of a low base rate of incidents, participants who were violent during follow-up were coded as one, and those who were not were coded as zero, and binary logistic regressions were performed. Multiple linear hierarchical regressions were performed to examine the variance in violent outcomes explained by neuropsychological measures when the risk assessment was controlled for (e.g., incremental validity). Bootstrapping using 1,000 replacement samples was performed on all regressions to obtain the 95% confidence intervals, and a range of effect sizes and estimates in a larger population.

6.5 Ethical Considerations

Participant Distress: In consideration of the length of time required to administer the battery, there was a risk for participants to experience fatigue. To combat this, the battery was administered in at least two 1-hour blocks to all inpatients, and community participants were given the option to complete it in either one or two meetings. Participants were also permitted to withdraw from the study at any time and were reminded that this would not affect their care or legal status in any way.

Disclosure of personal information: Participants completed the assessments voluntarily and were not required to disclose any information that they did not wish to. They were also encouraged not to share details about their offences and were not asked personal information during the duration of the study, aside from what was required to complete the PANSS. All personal information was collected from participants' personal files only after consent was obtained.

Risk of breach of confidentiality: There was a low risk of breach of confidentiality in the duration of this study. All participants were given a pseudonymous personal identifier to prevent names from being linked to personal data. All data was stored securely. Confidentiality would only have been breached if a participant disclosed a risk of harming themselves or others, or if they disclosed information about their legal case which had not been previously shared; all of which was outlined in the participant information sheet.

Researcher safety: The researcher completed the State Hospital's security induction and break away training. All assessments were administered in rooms within hospital wards, or within the social work office; the researcher never met with participants outside of these organizations, or without either a psychologist or social worker being aware of the meeting.

Data collection: The collection of data included neuropsychological assessment results and information from either medical or social work files. Due to the high sensitivity of this information, multiple methods were used to ensure confidentiality, including keeping all signed consent forms separate from data, following all best practice guidelines for data storage and management, and ensuring the database included only pseudonymised data, and was held on the secure University of Edinburgh server.

7 Assessing the Cognitive Contributors to Violence: Part II

Chapter Preface

This chapter presents descriptive data for the study samples, baseline measures, prevalence of cognitive impairment in each sample, and the feasibility results.

7.1 Results

7.1.1 Recruitment

Forensic Inpatients. Forensic inpatients were recruited from high, medium, and low secure hospitals across Scotland. Across all levels of security, *n*= 168 patients were considered for participation, and of those 12.5% (*n*= 21) were excluded for not having capacity to consent, 4.7% (*n*= 8) were 60 years or older, 13% (*n*= 22) were on the transfer list, 1% (*n*= 1) did not speak English as their first language, 1.7% (*n*= 3) did not have an HCR-20 version 3, 8.3% (*n*= 14) did not have a diagnosis of a psychotic illness, 2.3% (*n*= 4) had a physical disability that would impede their ability to take part in assessments, 2.3% (*n*= 4) had not had their admissions case review at the time of recruitment, 5.3% (*n*= 9) had diagnosed intellectual disabilities, and 3.5% (*n*= 6) were considered to be unsafe to participate in research. Forty-five percent (*n*= 76) patients fit the eligibility criteria and were approached, 58% (*n*=44) of those patients declined to participate, and 42% (*n*= 32) were successfully recruited.

Community Sample. In the community sample, a total of *n*= 6 individuals were recruited from the social work intervention team, and *n*= 26 were recruited from the unpaid work teams. Due to the large number of service users on each social worker's caseload, they were unable to record reasons for exclusion for their clients, and similarly, the unpaid work teams were also unable. In total *n*= 50 individuals were approached from both unpaid work and the intervention team, and *n*= 31 were successfully recruited, *n*= 18 refused to take part, and *n*= 1 withdrew shortly after consenting. As it is not known how many service users or individuals in unpaid work were eligible aside from those who showed interest, the approximate consent rate for social work was 62%.

7.1.2 Description of Samples

The mean age of the inpatient sample was 42.03 (SD= 10.68) years ranging from 20 to 60 years. The majority of patients (*n*= 23; 72%) were of White Scottish ethnicity and the remaining *n*= 9 (28%) were White British. The community sample had a mean age of 31.06 (SD= 9.33) years ranging from 19 to 51 years. The majority of the community sample (*n*= 25; 81%) were White Scottish, *n*= 1 was White British (3%), *n*= 2 were British Indian (6.5%), *n*= 1 was British Pakistani (3%), *n*= 1 was Malaysian (3%), and *n*= 1 was Black British (3%). All participants in both groups spoke English as their first language. An independent samples t-test with bootstrapped 95% confidence intervals (1,000 sample replicates) revealed that there was a significant difference in age between the groups (t (61)= 4.33, 95% CI [5.63, 16.26]), and a large effect size (Cohen's *d*= 1.09), with the inpatients being significantly older.

Details on the samples' personal demographics, including highest level of education and adverse childhood experiences can be found in Table 7.1. No significant differences were detected between the two groups on years of education (inpatients: 11.22 years SD= 1.38; community: 11.8 years SD= 2.96; t (61)= -0.91, 95% CI [-1.56, 0.63]; Cohen's *d*= 0.23). As can be seen in Table 7.1, portions of the personal demographic information were missing for some of the community participants and was labelled as *unknown*.

Table 7. 1 Personal Demographics

Demographic	Inpatients		Community	
	n	%	*n*	%
Highest Level of Education				
No qualifications	26	81.3	16	52
Standard Grades	1	3.1	3	10
Formal qualifications	0	0	4	13
College	1	3.1	7	22
Diploma	1	3.1	0	0
Undergraduate degree	2	6.3	0	0
Postgraduate degree	1	3.1	1	3
History of Child Sexual Abuse				
Yes	10	31	2	6.4
No	22	69	19	61
Unknown	0	0	10	32
History of Child Physical Abuse				
Yes	8	25	2	6.5
No	24	75	19	61.3
Unknown	0	0	10	32.2
History of Neglect				
Yes	6	19	3	10
No	26	81	20	64
Unknown	0	0	8	26

Psychiatric Demographics. In the inpatient sample, the most common diagnosis was paranoid schizophrenia (*n*=14; 44%), followed by schizophrenia (*n*=6; 19%), and schizoaffective disorder (*n*= 5; 16%). The primary and secondary diagnoses for the inpatient sample can be found in Table 7.2. The inpatient sample had a mean of approximately 12.4 (SD= 9.09) years of psychosis at the point of consent, with a range of 1 to 39 years. They have spent a mean of 31.4 (SD= 36.32) months in hospital, ranging from approximately 2 to 162 months. Psychiatric history was not always clearly reported in social work files. Of the *n*= 28 patients for whom data were available, *n*= 3 (11%) were diagnosed with ADHD, *n*=4 (14%) had a depressive disorder, *n*= 1 (4%) had an anxiety disorder, *n*= 1 (4%) had a diagnosis of PTSD, and *n*= 1 (4%) had possible alcohol dependence. The remaining *n*= 18 (37%) had no diagnostic information in their file.

Table 7. 2 Psychiatric Demographics for Inpatients

Diagnoses	Inpatients	
	n	%
Primary		
Paranoid Schizophrenia	14	44
Schizoaffective Disorder	5	16
Schizophrenia	6	19
Delusional Disorder	3	9
Bipolar Affective Disorder with Psychosis	1	3
Drug Induced Psychosis	1	3
Alcohol Induced Psychosis	1	3
Psychotic Illness/Disorder	1	3
Secondary/Other		
Antisocial/Dissocial Personality Disorder	7	22
Alcohol Dependence	3	9
Drug Dependence	2	6
Depressive Disorder	1	3
Emotional Unstable Personality Disorder	1	3
Mental and Behaviour Disorder due to multiple drug use	4	12.5
Mental and Behaviour Disorder due to harmful Alcohol use	3	9
Obsessional Personality Disorder	1	3
Obsessive Compulsive Disorder	1	3
Anxiety Disorder	1	3
Polysubstance/drug Misuse	7	22
Alcohol Misuse	1	3
Attention Deficit Hyperactivity Disorder	1	3
Harmful use of multiple substances/legal highs	4	12.5
Harmful use of drugs and alcohol	1	3

Relevant Medical Diagnoses. Approximately 41% (*n*= 13) of the inpatient sample had diabetes, *n*= 1 patient had a diagnosis of epilepsy, and *n*= 1 was diagnosed with Asperger's Syndrome. There were no relevant medical diagnoses reported for the community sample.

Birth Trauma and Abnormal Infant Development. Information on birth trauma and abnormal infant development was available only for the inpatient group. Whilst no birth trauma was reported for any of the inpatients, *n*= 3 had reports of abnormal infant development, *n*= 1 patient had a failure to thrive, *n*= 1 had anxious attachment, and *n*=1 had encephalitis at two months old, however they recovered after several months in hospital with no on-going treatment.

History of Substance Misuse and Alcohol Misuse. Seventeen inpatients (53%) had a secondary or other diagnosis related to drug and/or alcohol use. Among these were *n*= 6 patients who had an extensive history of use.

The community sample was less likely to have a clearly defined substance or alcohol use disorder, however, *n*= 5 (16%) participants had a history of excessive and significant use of alcohol, and *n*= 6 (21%) had a significant history of substance misuse.

History and Severity of Head injuries. Information on history and severity of head injuries is presented in Table 7.3. There were no significant differences between the two groups on any of the measures of head injury; number of head injuries, t(60)= 1.02, bootstrapped 95% CI [-0.42, 1.42]; Cohen's *d*= 0.26; number of head injuries with LOC, t(41.61)= 0.90, bootstrapped 95% CI [-0.19, 0.74], Cohen's *d*= 0.23 (degrees of freedom differ because equal variances were not assumed); and severity of head injury, t(60)= 0.90, bootstrapped 95% CI [-0.68, 0.74], Cohen's *d*= 0.

Table 7. 3 History and Severity of Head Injuries in Each Sample

	Inpatients *n*= 31				Community *n*= 31			
Head Injury	*n*	%	Mean	SD	*n*	%	Mean	SD
One or more head injury	25	81	2.55	2.13	25	81	2.06	1.57
Head injury with LOC	12	39	0.71	1.27	14	45	0.48	0.57
Severity of Head injury	--	--	2.06	1.45	--	--	2.06	1.29
Improbable TBI	5	16	--	--	3	10	--	--
Possible mild TBI	8	26	--	--	8	26	--	--
Definite mild TBI	6	19	--	--	11	35	--	--
Moderate TBI	5	16	--	--	3	10	--	--
Severe TBI	1	3	--	--	0	0	--	--

Note. A head injury questionnaire was missing for *n*= 1 inpatient; *LOC*= loss of consciousness; *TBI*= Traumatic brain injury; *Improbable TBI*= the individual experienced no effects after hitting their head; *Possible mild TBI*= no LOC, but reported being dazed and/or having a gap in their memory; *Definite* *mild TBI*= LOC of $\leq$30 minutes, in addition to a gap in memory and/or feeling dazed; *Moderate TBI*= a LOC of 30 minutes to 24 hours; *Severe TBI*= TBI with a LOC for more than 24 hours.

Forensic History. The inpatient sample had a mean of 10.36 (SD= 9.94) convictions, ranging from 1 to 56 per individual, and the community sample had a mean of 8.81 (9.96) convictions, ranging from 1 to 40 convictions per individual. All participants were convicted or charged of at least one violent offence, with the most common offence in the inpatient sample being assault to injury, and the most common offence in the community sample being assault to injury- domestic (Table 7.4). Notably, *n*= 3 inpatients and *n*=3 community offenders, were first time violent offenders. There were no significant differences between the groups on number of previous convictions (t (61) =0.63, bootstrapped 95% CI [-3.61, 7.52]; Cohen's *d*= 0.16).

Table 7. 4 History of Offences

Offences	Inpatients		Community	
	n	%	*n*	%
Non-Sexual Violence				
Assault to Severe Injury/ Serious Assault	15	47	4	15
Attempted Murder	7	22	1	4
Culpable Homicide	4	12.5	0	0
Murder	9	28	1	4
Armed robbery/Assault and Robbery	11	34	2	8
Hostage Taking/Abduction	1	3	1	4
Assault to injury/ Assault to injury-domestic	20	62.5	25	96
Domestic Assault	0	0	7	27
Assault against a child	0	0	1	4
Serious threats/Abuse behaviour/Intimidation	2	6	10	38
Indecency				
Child Sexual Assault	2	6	0	0
Exposure	1	3	0	0
Dishonesty				
Theft	8	25	6	23
Theft by housebreaking	3	9	1	4
Housebreaking	5	16	1	4
Shoplifting	4	12.5	3	11.5
Other Crimes				
Possessing or carrying an offensive weapon	4	12.5	1	4
Drugs	10	31	7	27
Malicious Damage/Vandalism	5	16	0	0
Racial Assault/Crimes	1	3	5	19
Crimes against public justice	3	9	0	0
Breach of Peace	13	41	5	19
Breach of conditions/Probation	4	12.5	9	35
Culpable and reckless conduct	2	6	1	4
Stalking	1	3	0	0
Willful Fire raising/Arson	1	3	2	8
Domestic Abuse	0	0	9	25
Other	1	3	4	15
Driving				
Drink/drug driving	3	9	2	8
Traffic / Driving Violations	3	9	5	19

Level of Security: Inpatients. The majority of patients consented to participate were in high secure care (*n*= 20), followed by medium secure (*n*= 6), and low secure (*n*= 6). All high secure patients were recruited from the State Hospital in Carstairs, Scotland, the medium secure patients were recruited from the Rowanbank Clinic in Glasgow, Scotland, and the n= 4 patients were recruited from Radernie low secure Ward in Fife, Scotland, and *n*= 2 were recruited from Leverndale Hospital, in Glasgow, Scotland.

Legislation. All inpatients were detained in hospital under the Mental Health (Care and Treatment) Act Scotland. Seven patients were held on compulsion orders, and *n*=1 on an interim compulsion order, *n*= 12 had a compulsion order with a restriction order, *n*= 2 had a compulsion order with a restriction order-Insanity, *n*= 9 patients were under a Transfer for Treatment Direction, and *n*= 1 with a Transfer for Treatment Direction with an Order of Lifelong Restriction.

Twenty-five of the community offenders were on a community payback order (CPO), *n*= 1 was on a CPO with a drug treatment order, *n*=2 with a CPO and restriction of liberty order (ROLO), *n*= 1 with a CPO, drug treatment order, and ROLO, *n*= 1 had a mandatory life sentence, and *n*=1 was on conditional release.

7.2 Preliminary Analyses

7.2.1 Data Screening- Inpatients

Data were screened for each group separately, as the primary and secondary hypotheses investigate each sample separately. Data were first cleaned and checked for accuracy and to ensure they were all correctly defined in R, and then were checked for missing data. The only data missing in the inpatient sample was a head injury questionnaire for one patient, thus, all analyses on head injuries included *n*= 31 patients. Using a subset of the data set which included only the core neuropsychological variables (e.g., reasoning, inattention, response monitoring, response monitoring, affect recognition, cognitive empathy, and lack of insight) and all outcomes, Mahalanobis Distance was calculated to detect multivariate outliers, however no outliers were present. Data were further screened to check the assumptions for multiple regression which were all met. Zero-order correlations were performed to examine the intercorrelations between measures. Only subtests within the same measures evidenced correlations greater than 0.90, which was expected (correlation matrices can be found in Appendix R, Appendix S).

Distributions for all key variables are presented in Table 7.5. When examining for skewness values greater than 2 indicate a problem, and values greater than 7 indicate a problem for kurtosis (West et al., 1996). Descriptive analyses showed that the outcome variables violence prospective and retrospective, aggression prospective and retrospective, and number of days on elevated levels were all skewed, in addition to head injuries with LOC. Variables which had a kurtosis value of greater than 7 were violence prospective and retrospective, and head injuries with LOC. Using the Shapiro-Wilkes Test, all the data were tested for normality, which resulted in several measures being identified as non-normal. The *p*-values for the Shapiro-Wilkes Test are presented in Table 7.5, where values less than 0.05 indicate a significant divergence from normality.

Table 7. 5 Distribution of Inpatient Data

Variable	*n*	Skew	Kurtosis	Normality
Number of head injuries	31	0.90	0.75	.01
Head injuries with LOC	31	2.52	7.22	<.001
OSU-TBI-Severity	31	0.08	-1.12	.03
Full Scale IQ	32	0.30	-0.87	.21
Verbal Comprehension	32	0.65	0.70	.23
Reasoning	32	0.06	-1.12	.21
Inattention	32	0.86	-0.62	<.001
Response Inhibition	32	0.36	-0.23	.67
Response Monitoring	32	0.23	-0.78	.34
Risk Taking	32	0.12	-0.43	.35
Affect Recognition	32	-0.90	-0.09	.01
Cognitive Empathy	32	-0.13	-1.14	.16
Lack of Insight	32	0.66	-1.09	.001
Symptom Severity	32	1.14	0.28	<.001
Positive Symptoms	32	1.32	1.32	<.001
Violence Proneness	32	-0.64	-0.52	.05
Violence Prospective	32	5.00	24.14	<.001
Violence Retrospective	32	4.03	16.87	<.001
Violence Severity Prospective	32	1.66	2.03	<.001
Violence Severity Retrospective	32	0.90	-0.76	<.001
Aggression Prospective	32	2.53	5.14	<.001
Aggression Retrospective	32	2.33	4.12	<.001
Levels Prospective	32	1.98	2.86	<.001
Levels Retrospective	32	1.55	1.45	<.001
Days on Elevated Levels Prospective	32	2.03	3.01	<.001
Days on Elevated Levels Retrospective	32	1.32	0.41	<.001
Length of Admission	32	1.77	3.13	<.001
Self-Reported Impulsivity	32	0.47	-0.02	.65

Note. OSU-TBI= Ohio State University- Traumatic Brain Injury questionnaire; values below 0.05 under Normality indicate a significant divergence from a normal distribution.

7.2.2 Data Screening- Community Sample

Data screening for the community sample was completed using the same procedures as above. Data were missing for risk assessments, namely, *n*= 18 participants had a completed

Level of Service Inventory- Revised: Screening Version (LSI-R:SV), *n*= 7 had a completed Level of Service/ Case Management Inventory (LS/CMI), *n*= 2 had a completed Spousal Assault Risk Assessment (SARA), and the remaining *n*= 4 did not have any risk assessment. As the LSI-R:SV had the most completed data, it was used to examine the primary hypothesis, and only the subset of participants who had a completed measure were included in the analysis. There were *n*= 2 participants who did not have prospective follow-up information available, to remain conservative, they were excluded from prospective analyses. Data screening revealed that violence severity prospective and retrospective were the only two outcomes which were not naturally binary. Further, no outliers were detected. The distribution of the data for the community sample are presented in Table 7.6. For skewness, one outcome variable, violence severity prospective, had a skew value of 3.10, and no other variables were skewed. Violence severity prospective also had a high kurtosis value. Both outcome variables were significantly non-normal in addition to several other measures. Zero-order correlations were performed to examine in the intercorrelations between measures. Only subtests within the same measures evidenced correlations greater than 0.90, which was expected (correlation matrices can be found in Appendix T, Appendix U).

Table 7. 6 Distribution of Community Data

Variable	*n*	Skew	Kurtosis	Normality
Number of head injuries	31	0.45	-0.45	.04
Head injuries with LOC	31	0.58	-0.81	<.001
OSU-TBI-Severity	31	-0.39	-1.09	.002
Full Scale IQ	31	0.04	-1.10	.39
Verbal Comprehension	31	-0.21	-0.95	.39
Perceptual Reasoning	31	-0.19	-0.79	.79
Inattention	31	1.17	-0.22	<.001
Response Inhibition	31	0.97	1.15	.02
Response Monitoring	31	-0.27	-1.50	.01
Risk Taking	31	0.35	-0.76	.21
Affect Recognition	31	-0.19	-0.74	.12
Cognitive Empathy	31	-0.09	-0.94	.04
LSI-R:SV	18	0.30	-1.50	.28
Violence Severity Prospective	29	3.10	10.68	<.001
Violence Severity Retrospective	31	0.25	-1.67	<.001
Self-Reported Impulsivity	31	0.46	1.82	.08

Note. OSU-TBI= Ohio State University- Traumatic Brain Injury questionnaire; LSI-R:SV= Level of Service Inventory- Revised: Screening Version; values <.05 under Normality indicate a significant divergence from a normal distribution.

7.3 Baseline Assessments

7.3.1 Between-Group Differences

Means and standard deviations were calculated for all variables, and then MANOVAs were performed for each measure with Bonferroni correction to avoid inflated type I errors from multiple testing, and Pillai's multivariate statistic was used to account for violations of MANOVA assumptions. The MANOVA for the Stroop C&W test was significant, and post-hoc analyses revealed that there was a statistically significant difference between the groups on response inhibition, however, following Bonferroni correction, it was no longer significant. The MANOVA for the TASIT Part II was also significant, and post-hoc analyses using the Kruskal-Wallis test revealed that all variables were significant, though following Bonferroni correction, only cognitive empathy (Feel; H (1) = 8.80, p = .003) and sincere sarcasm were significant (H (1) = 12.0, p = .0005). Overall, the inpatient group performed more poorly relative to the community sample on the WASI, all CPT subtests except for distractibility (d'), all the Stroop C&W test, except for the *word* subtest, the entire M-WCST, and all the IGT except for Nets 3, 4, and 5, and risk taking (Net Total). On the TASIT, the inpatients incorrectly identified more positive and negative emotions, and recognized fewer examples of simple and paradoxical sarcasm. Means, SD, Cohen's *d* and results of the MANOVAs for each measure are presented in Table 7.7.

Table 7. 7 Descriptive Statistics for Baseline Measures and Results for Between Group Analyses

	Inpatients			Community							
	Mean	SD	95% BCa CI[a]	Mean	SD	95% BCa CI[a]	Cohen's *d*	Pillai's Trace	F	df	*p*
WASI-III											
Full Scale IQ	80.91	14.68	[76.17, 85.93]	82.00	10.85	[78.08, 85.93]	0.08				
Reasoning	83.91	15.76	[78.68, 89.05]	87.06	14.81	[81.57, 92.55]	0.21				
Verbal Comprehension	80.03	14.12	[75.38, 85.14]	80.67	8.92	[77.54, 83.68]	0.05	.06	1.24	3, 59	.30
CPT-II											
d' (Distractibility)	59.56	12.71	[54.89, 64.24]	59.93	12.37	[55.64, 64.26]	0.03				
Perseverations	58.75	17.28	[53.37, 64.91]	55.48	12.38	[51.18, 60.00]	0.22				
Commissions	57.69	10.83	[53.67, 61.47]	59.00	13.79	[54.31, 63.77]	0.11				
Inattention	58.68	14.84	[53.44, 64.30]	57.16	15.95	[52.21, 63.14]	0.09	.03	0.49	4, 58	.74
Stroop C&W Test											
Color	44.03	13.21	[39.39, 48.16]	42.90	8.12	[40.20, 46.02]	0.10				
Word	36.87	11.68	[32.69, 40.66]	40.16	8.06	[37.24, 43.04]	0.32				
Color-Word	43.40	12.57	[39.18, 47.60]	46.74	7.90	[43.86, 49.59]	0.32				
Response Inhibition	46.81	6.12	[44.70, 49.07]	50.42	6.74	[48.10, 52.64]	0.56	.16	2.68	4, 58	.03[b]
M-WCST											
Categories Correct	48.63	11.39	[44.51, 52.18]	53.45	6.72	[50.96, 55.77]	0.51				
% Perseverative Errors	45.69	10.89	[42.00, 49.00]	49.74	13.39	[44.78, 54.73]	0.33				
Response Monitoring	46.87	9.59	[43.71, 49.87]	53.19	9.66	[49.74, 56.74]	0.66				
Total Errors	49.18	10.70	[45.51, 52.60]	53.93	8.31	[51.00, 56.77]	0.49				
Executive Function	95.93	16.82	[90.03, 101.08]	105.67	12.08	[101.66, 109.93]	0.66	.11	1.88	4, 58	.13
IGT-II											
Net 1	51.66	10.75	[48.15, 54.92]	53.09	6.03	[50.78, 55.29]	0.16				
Net 2	45.16	10.09	[41.68, 48.38]	45.80	7.44	[43.60, 48.17]	0.07				
Net 3	45.91	8.36	[42.79, 49.07]	43.67	8.66	[40.51, 46.73]	0.26				
Net 4	45.75	11.75	[41.75, 50.34]	42.74	11.75	[39.44, 46.08]	0.28				
Net 5	46.00	9.87	[42.44, 49.65]	42.39	9.84	[39.00, 45.93]	0.36				

Risk Taking	45.81	9.48	[42.64, 49.13]	44.71	6.91	[42.37, 46.96]	0.13	.08	0.79	6, 56	.58
TASIT Part I											
Happy	3.03	0.78	[2.76, 3.33]	3.29	0.69	[3.03, 3.53]	0.35				
Surprised	3.44	0.76	[3.17, 3.65][c]	3.74	0.51	[3.51, 3.93][c]	0.47				
Neutral	2.34	1.12	[1.94, 2.74]	2.42	0.62	[2.21, 2.62]	0.08				
Sad	3.09	0.77	[2.82, 3.36]	3.39	0.76	[3.10, 3.63]	0.38				
Angry	2.97	1.03	[2.59, 3.28]	3.32	0.65	[3.10, 3.53]	0.41				
Anxious	3.34	1.03	[2.96, 3.69]	3.32	0.79	[3.04, 3.58]	0.02				
Revolted	3.06	1.04	[2.69, 3.40]	3.07	0.83	[2.79, 3.40]	0.04				
Affect Recognition	21.28	4.32	[19.74, 22.70]	22.55	2.36	[21.67, 23.51]	0.36	.14	1.06	8, 54	.40
TASIT Part II											
Do	10.50	2.31	[9.73, 11.31]	11.77	1.58	[11.20, 12.33]	0.64				
Say	10.68	2.69	[9.74, 11.61]	11.90	2.12	[11.06, 12.68]	0.51				
Think	9.84	2.37	[9.03, 10.63]	11.45	2.01	[10.77, 12.13]	0.73				
Cognitive Empathy	10.78	2.45	[9.88, 11.65]	12.61	1.45	[12.10, 13.16]	0.91**				
Sincere	16.19	3.29	[14.90, 17.22]	14.90	3.47	[13.78, 16.06]	0.38				
Simple Sarcasm	11.87	5.36	[10.00, 13.66]	16.48	3.08	[15.24, 17.68]	1.05***				
Paradoxical Sarcasm	13.50	5.24	[11.53, 15.35]	16.35	3.18	[15.14, 17.48]	0.66				
Total Correct (Part II)	41.93	9.10	[38.87, 45.00]	47.58	6.56	[44.94, 50.02]	0.71	.32	3.25	8, 54	.004**
Lack of Insight	2.56	1.63	[2.09, 3.09]	--	--	--	--	--	--	--	--

Note. This table presents means, standard deviations (SD), bootstrapped 95% Bias-corrected and accelerated confidence intervals (95% Bca CI) for the mean, Cohen's *d* (Effect Size), and results of MANOVAs for each measure. Pillai's Trace is the MANOVA test statistics, *F*= F-statistic, *df*= degrees of freedom, *p*= p-value for the MANOVA.

[a] Based on 1,000 bootstrap samples unless otherwise noted.

[b] No longer significant following Bonferroni correction.

[c] Based on 999 bootstrapped samples.

** *p*-value < 0.01

*** *p*-value < 0.001

7.3.2 Cognitive Impairments

Table 7.8 presents the proportion of participants in each sample who were impaired on each measure. To analyse these, all scores were first transformed into z-scores using normative means. For standard scores (e.g., WASI-II scores), a mean of 100 and SD = 15 was used, for T-scores (e.g., the CPT-II, M-WCST, IGT-II) a mean of 50 and SD=10 was used, and for the TASIT, the normative means reported for each item in the administration manual were used. Following Table 7.8 is Figure 7.1, a line graph representing the mean z-scores for each sample on each measure, with error bars representing the standard error of the z-score mean. In general, more inpatients were impaired on the measures, however, although there were only significant differences between groups on measures of cognitive empathy and simple sarcasm, proportions of participants with impairment were large in both samples indicating the severe cognitive impairments seen in violent offenders. Notably, for example, approximately 2.5% of a normal population should fall into the borderline or impaired ranges on a Wechsler scale, however, in this sample of n=63 participants, 44% overall fell into this range.

Table 7. 8 Proportion of Participants with Impairments on Neuropsychological Measures

Measure	Inpatients				Community				Whole Sample
	Low Average (%)[a]	Borderline (%)[b]	Impaired (%)[c]	Borderline + Impaired (%)	Low Average (%)	Borderline (%)	Impaired (%)	Borderline + Impaired (%)	Borderline + Impaired (%)
WASI-III									
Full Scale IQ	8 (25)	5 (16)	10 (31)	15 (47)	9 (29)	9 (29)	4 (13)	13 (42)	28 (44)
Reasoning	6 (19)	6 (19)	7 (22)	13 (41)	10 (32)	4 (13)	4 (13)	8 (26)	21 (33)
Verbal Comprehension	6 (19)	10 (32)	7 (22)	17 (54)	12 (39)	9 (29)	5 (16)	14 (45)	31 (49)
CPT-II									
d' (Distractibility)	2 (6)	10 (31)	5 (16)	15 (47)	7 (23)	1 (3)	9 (29)	10 (32)	25 (40)
Perseverations	4 (13)	0	9 (28)	9 (28)	6 (19)	5 (16)	3 (10)	8 (25)	17 (27)
Commissions	8 (25)	8 (25)	3 (9)	11 (34)	4 (13)	9 (29)	5 (16)	14 (45)	25 (40)
Inattention	7 (22)	2 (6)	5 (16)	7 (22)	3 (10)	1 (3)	6 (19)	7 (23)	14 (22)
Stroop C&W Test									
Color	6 (19)	3 (9)	4 (13)	7 (22)	12 (39)	4 (13)	1 (3)	5 (16)	12 (19)
Word	8 (25)	6 (19)	7 (22)	13 (41)	15	3 (10)	3 (10)	6 (19)	19 (30)
Color-Word	5 (16)	6 (19)	3 (9)	9 (28)	4 (13)	4 (13)	0	4 (13)	13 (21)
Response Inhibition	9 (28)	1(3)	0	1 (3)	3 (10)	0	0	0	1 (2)
M-WCST									
Categories Correct	3 (9)	2 (6)	3 (9)	5 (16)	2 (6)	0	0	0	3 (5)
% Perseverative Errors	6 (19)	4 (13)	2 (6)	6 (19)	3 (10)	3 (10)	3 (10)	6 (19)	12 (19)
Response Monitoring	7 (22)	4 (13)	0	4 (13)	6 (19)	1 (3)	0	1 (3)	5 (8)
Total Errors	5 (16)	2 (6)	1 (3)	3 (9)	3 (10)	0	0	0	3 (5)
Executive Functions	4 (13)	6 (19)	2 (6)	8 (25)	4 (13)	0	0	0	8 (13)
IGT-II									
Net 1	4 (13)	0	1 (3)	1 (3)	1 (3)	1 (3)	0	0	1 (2)
Net 2	4 (13)	3 (9)	2 (6)	5 (16)	6 (19)	3 (10)	0	0	5 (8)

Net 3	8 (25)	1 (3)	1 (3)	2 (6)	6 (19)	2 (6)	3 (10)	5 (16)	7 (11)
Net 4	6 (19)	3 (9)	2 (6)	5 (16)	8 (26)	4 (13)	3 (10)	7 (23)	12 (19)
Net 5	7 (22)	2 (6)	2 (6)	4 (13)	9 (29)	4 (13)	3 (10)	7 (23)	11 (17)
Risk Taking	3 (9)	5 (16)	1 (3)	6 (19)	10 (32)	3 (10)	0	3 (10)	9 (14)
TASIT									
Happy	0	16 (50)	7 (22)	23 (72)	0	14 (45)	4 (13)	18 (58)	41 (65)
Surprised	0	11 (34)	3 (9)	14 (44)	0	6 (19)	1 (3)	7 (23)	21 (33)
Neutral	7 (22)	0	8 (25)	8 (25)	17 (55)	0	1 (3)	1 (3)	9 (14)
Sad	16 (50)	0	6 (19)	6 (19)	9 (29)	0	5 (16)	5 (16)	11 (17)
Angry	11 (34)	0	9 (28)	9 (28)	15 (48)	0	3 (10)	3 (10)	12 (19)
Anxious	4 (13)	0	7 (22)	7 (22)	12 (39)	0	4 (13)	4 (13)	11 (17)
Revolted	10 (31)	0	8 (25)	8 (25)	13 (42)	0	7 (23)	7 (23)	15 (24)
Affect Recognition	7 (22)	4 (13)	12 (38)	16 (50)	7 (23)	7 (23)	6 (19)	13 (42)	29 (46)
TASIT Part II									
Do	0	3 (9)	22 (69)	25 (78)	4 (13)	10 (32)	12 (39)	22 (71)	47 (75)
Say	4 (13)	2 (6)	16 (50)	18 (56)	7 (23)	5 (16)	6 (19)	11 (25)	29 (46)
Think	0	2 (6)	21 (66)	23 (72)	0	4 (13)	10 (32)	14 (45)	37 (59)
Cognitive Empathy	0	5 (16)	17 (53)	22 (69)	0	10 (32)	6 (19)	16 (52)	38 (60)
Sincere	5 (16)	1 (3)	2 (6)	3 (9)	3 (10)	4 (13)	5 (16)	9 (29)	12 (19)
Simple Sarcasm	5 (16)	1 (3)	20 (63)	21 (66)	7 (23)	3 (10)	6 (19)	9 (29)	30 (48)
Paradoxical Sarcasm	0	2 (6)	20 (63)	22 (69)	0	3 (10)	12 (39)	15 (48)	37 (59)
Total Correct (Part II)	3 (9)	3 (9)	19 (59)	22 (69)	0	8 (26)	11 (35)	19 (61)	41 (65)

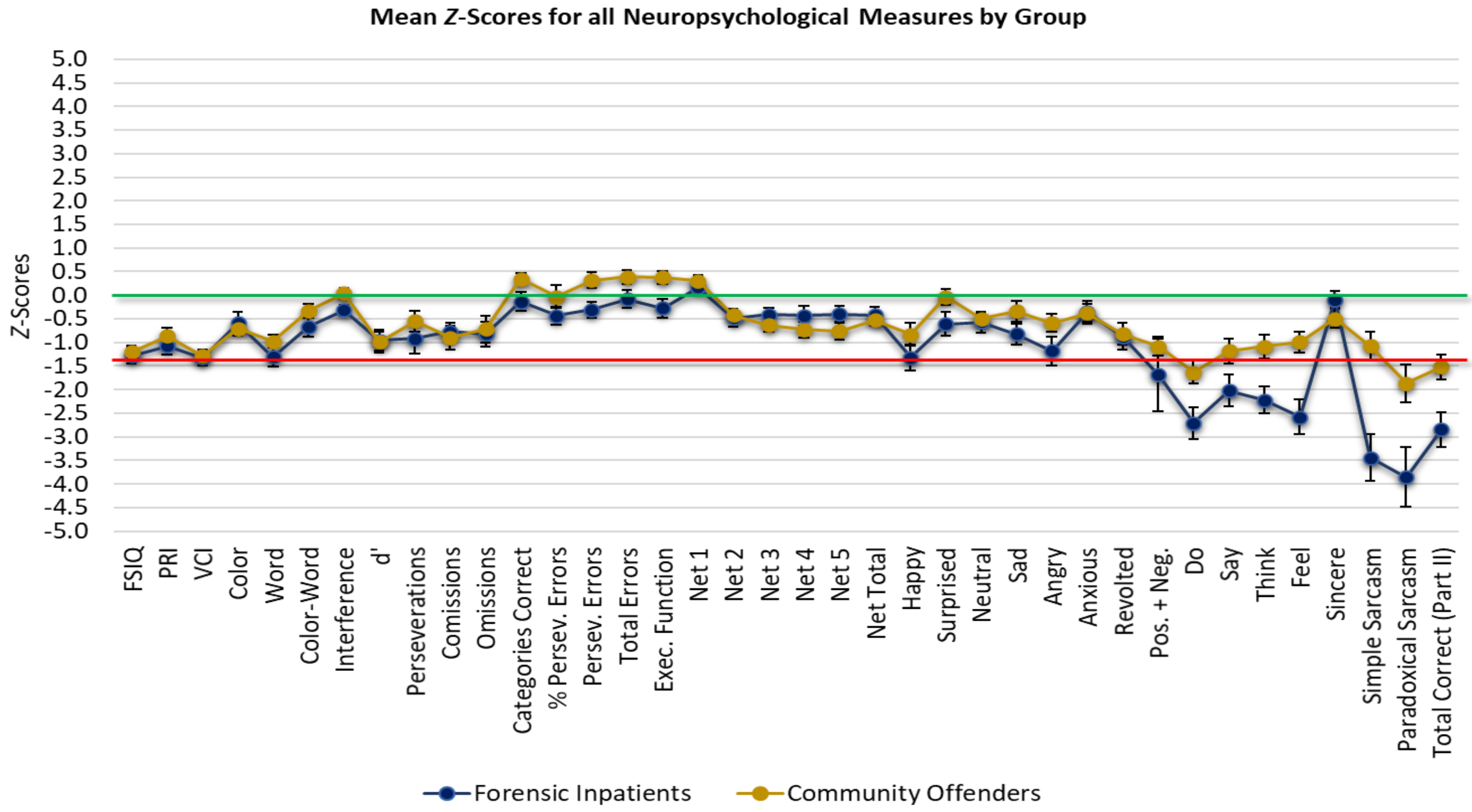
Note. This table presents the proportion of participants in each group, and overall, who had scores in the low average, borderline, impaired, and borderline + impaired ranges on each measure.

[a] Low Average: *z*-scores < -0.67

[b] Borderline: *z*-scores < -1.33

[c] Impaired: *z*-scores < -2.00

Figure 7. 1 Mean *Z*-scores by Group



Note. Line markers represent the mean z-score for each group; error bars represent the standard error for the mean z-score for each group. The red line marks z-score= -1.33; scores below this are in the borderline and impaired range; the green line represents the normative mean (e.g., *z*-score= 0.00). *FSIQ*= Full scale IQ; *PRI*= perceptual reasoning index; *VCI*= verbal comprehension index; *d'*= distractibility; *Persev.*= Perseverative; *Pos.+Neg.*= Positive emotions + negative emotions.

7.3.3 Violence and Offending Proneness

Means, standard deviations, and ranges are presented in Table 7.9 for the HCR-20v3 categories for the inpatient sample and indicates that this sample had more historical risk factors overall. A breakdown of the data for the sub-sample of community offenders who had a completed LSI-R:SV can be found in Table 7.10. The items that were most common in the low-risk group (e.g., 0-2 endorsed risk/need items) were having two or more prior convictions and non-rewarding parental relationships, similarly, two or more prior convictions was almost a common item in the medium risk group (e.g., 3-5 endorsed risk/need items), as well as alcohol or drug problems associated with school or work. There were only *n*= 3 individuals in the high-risk group (e.g., 6-8 endorsed risk/need items), however, 100% of them had five out of the eight risk/need factors for future offending.

Table 7. 9 Descriptive Statistics for Inpatients on the HCR-20v3

HCR-20 Scores	Mean	SD	Range
Historical	16.19	1.06	9-20
Clinical	6.78	3.02	1-11
Risk	5.87	2.36	1-10
Total	28.84	6.72	13-38

Note. Table represents the means, standard deviations (SD), and range for scores on the Historical, Clinical, Risk Management-20, version 3 (HCR-20v3).

Table 7. 10 Proportions of Endorsed Risk Factors by Level of Risk and Need on the LSI-R:SV

	Level of Risk or Need					
	Low (*n*= 7)		Medium (*n*= 8)		High (*n*= 3)	
LSI-R:SV Item	*n*	%	*n*	%	*n*	%
Two or more prior convictions	3	43	7	87.5	3	100
Arrested younger than 16	0	0	2	25	3	100
Currently Unemployed	0	0	2	25	3	100
Some criminal friends	0	0	4	50	3	100
Alcohol or drug problems (school or work)	1	14	7	87.5	3	100
Psychological Assessment Indicated	1	14	2	25	1	33
Non-rewarding parental relationships	3	43	3	37.5	1	33
Attitudes or orientations supportive of crime	2	29	6	75	2	67

Note. *N*= 18 community offenders; LSI-R:SV= Level of Service Inventory-Revised: Screening Version.

7.3.4 Follow-up Data

Inpatient Outcomes. During the prospective follow-up period, *n*= 2 patients displayed violent behaviour (physical contact), one with *n*=7 recorded incidents, and the other with *n*=1 incident, and there were *n*=9 patients who displayed aggression (no contact), with incidents ranging from 1-18 per patient. Retrospectively, *n*= 5 patients displayed violent behaviour with incidents ranging from 1-7, and *n*= 10 patients had incidents of aggression ranging from 1-7 recorded incidents (Table 7.11). Thus, the prospective base rate was 6% for violence, 28% for aggression, and the retrospective base rate was 16% for violence, and 31% for aggression. All violent incidents were coded using the Woodworth and Porter (2002) scale to determine whether they were reactive or instrumental, however, all incidents during the follow-up periods were reactive, thus, the analysis of reactive and instrumental violence was not possible.

Table 7. 11 Descriptive Statistics for Primary and Secondary Outcomes for Inpatients

Primary and Secondary Outcomes	Mean	SD	Median	Range
Frequency of Violence Prospective	0.04	0.21	0.00	0-1.16
Frequency of Violence Retrospective	0.07	0.22	0.00	0-1.16
Severity of Violence Prospective	0.35	0.64	0.00	0-2.50
Severity of Violence Retrospective	0.51	0.73	0.00	0-2.00
Frequency of Anger Prospective	0.31	0.77	0.00	0-3.00
Frequency of Anger Retrospective	0.15	0.33	0.00	0-1.16
Levels of Observation Prospective	1.14	0.26	1.00	1.00-2.00
Levels of Observation Retrospective	1.29	0.46	1.00	1.00-2.70
Days on Enhanced Levels Prospective	4.03	7.86	0.00	0.00-30.50
Days on Enhanced Levels Retrospective	5.90	9.00	0.00	0.00-30.50
Length of Admission[s]	33.06	35.91	18.69	2.21-162.08
Self-Reported Impulsivity	66.31	8.98	66.00	49.0-89.0

Note. This table presents the means, standard deviations (SD), median, and ranges for outcomes in the inpatient sample.

[a]Length of admission is in months.

Community Sample. Prospectively, *n*= 2 participants were violent, one time each, *n*= 2 were aggressive, *n*= 5 had incidents of antisocial behaviour (only *n*= 2 had more than one offence), and *n*= 1 had a breach for missing a court-ordered curfew. Retrospectively, *n*= 13 were violent, *n*= 2 were aggressive, *n*= 8 had incidents of antisocial behaviour, and *n*= 6 had breaches of their court ordered sentences (only *n*= 1 had more than one breach). The prospective base rate was 7% for violence, 7% for aggression, 17% for antisocial behaviour, and 3% for breaches. Retrospectively the base rate was 42% for violence, 6% for aggression, 25% for antisocial behaviour, and 19% for breaches. All violent incidents were coded using the Woodworth and Porter (2002) scale to determine whether they were reactive or instrumental, however, all incidents during the follow-up periods were reactive, thus, the analysis of reactive and instrumental violence was not possible. Table 7.12 presents data for the continuous outcomes.

Table 7. 12 Descriptive Statistics for Primary and Secondary Outcomes

Primary and Secondary Outcomes	Mean	SD	Median	Range
Violence Severity Prospective	0.22	0.80	0.00	0-4
Violence Severity Retrospective	0.93	1.03	0.00	0-3
Self-Reported Impulsivity	70.97	8.80	72.00	52-99

7.3.5 Risk Factors for Cognitive Impairment and Violence

Data were examined to calculate the proportion of participants who had a risk factor for cognitive impairment (e.g., one or more head injury, a head injury with LOC, and history of alcohol/substance misuse in conjunction with one or more head injury). Overall, n= 50 (81%) of the entire sample reported having one or more head injury, n= 26 (42%) reported having a head injury with LOC, and n= 18 (29%) had both a history of alcohol and/or substance misuse and one or more head injury. During prospective follow-up, overall, n= 4 participants were physically violent (n= 2 in each group), and all four participants reported having one or more head injury, n= 1 inpatient and n = 1 community participant reported a head injury with LOC, and n= 1 inpatient reported one or more head injury, and a history of substance and alcohol misuse. Retrospective violence was also analysed, as it had a higher base rate in both samples (n= 5 inpatients and n= 13 community participants were violent). Of the n= 18 violent participants, n= 13 (72%) reported one or more head injury, n= 9 (50%) reported a head injury with LOC, and n= 2 (11%) reported both one or more head injury and a history of substance and/or alcohol misuse. Despite postulations that participants with risk factors for cognitive impairments would behave more violently, this data suggests that although participants with head injuries accounted for a large proportion of the incidents prospectively and retrospectively, a large proportion of individuals with these risk factors were not violent.

7.4 Feasibility Results

There were seven feasibility questions of interest for this study (listed below). Answers to questions 1 and 2 were previously presented in Table 7.1. The battery took approximately 1 hour and 40 minutes to complete in the inpatient sample, and 1 hour and 36 minutes in the community sample. Sixty-six percent of inpatients (*n*= 21) completed the battery in two sessions, and 34% (*n*= 11) completed it in three sessions. Nine percent (*n*= 3) of the community sample completed testing in two sessions, and 91% (*n*= 29) completed it in one session. The rate of completion was 100% in both samples. Data on the tolerability of measures is presented in Table 7.13, and the answer to feasibility question 7 is presented in Table 7.14.

7.4.1 Feasibility Research Questions

1. What percentage of potential participants were considered by their teams to lack the capacity to consent to participate?
2. What percentage of eligible participants were successfully recruited?
3. How many minutes, and over how many sessions does the test battery take?
4. What is the rate of completion of each measure and the battery as a whole?
5. What are the reasons for non-completion of each measure or the battery as a whole?
6. How tolerable did the participants find each measure?
7. Are the proposed measures of cognitive impairment sufficiently sensitive and specific to be associated with violence (inpatient and community) prospectively, and add incremental validity?

Table 7. 13 Tolerability of Measures by Group

	Inpatients		Community	
Measures	Mean (SD)	Median	Mean (SD)	Median
Wechsler Abbreviated Scale for Intelligence, version 2	5.2 (1.28)	5	5.19 (0.84)	5
Connors' Continuous Performance Test, version 3	4.35 (1.14)	4	4.19 (1.25)	4
Stroop Colour and Word Test	4.95 (1.09)	5	4.93 (0.99)	5
Modified Wisconsin Card Sorting Test	5.55 (0.99)	5.5	5.35 (0.87)	5
Iowa Gambling Task, Version 2	4.95 (1.19)	5	4.48 (1.29)	5
The Awareness of Social Inference Test	4.85 (1.34)	4	4.74 (1.21)	5
Full Battery	4.97 (1.18)	5	4.81 (1.16)	5

Note. Tolerability scores were as follows: 7= Extremely Pleasant; 6= Very Pleasant; 5= Pleasant; 4= Neutral; 3= Unpleasant; 2= Very Unpleasant; 1= Extremely Unpleasant, Tolerability measure for reference (Appendix Q).

7.4.2 Sensitivity, Specificity, and Accuracy

Receiver Operating Characteristics (ROC) and area under the ROC curve (AUC) were performed to examine the predictive performance of the neuropsychological measures. First, ROC curve analysis was performed to obtain AUC values which represent the predictive accuracy of the measures for the outcome. As the occurrence of violence is rare, observational research often results in a low base rate of violent incidents, and as this method is independent of the number of incidents (McDermott et al., 2008), it was deemed to be appropriate for the current study. This statistic can identify the most accurate cut-off point on a measure and provides an index of the sensitivity and specificity for this particular cut off point (Singh, 2013). Moreover, the AUC is frequently used in risk assessment literature, which will make the current results more comparable across studies and easier to interpret. In this thesis, the AUC was defined as the probability that a randomly selected participant who demonstrated violent behaviour during a follow-up period, had a higher score on a risk factor (higher score here indicates higher impairment) than a randomly selected non-violent offender. Further, as there has been variability across the violence literature in the benchmarks used to interpret the magnitudes of AUC values (Singh, 2013), AUC values ≥ 0.70 were considered sufficient for interpretation. It should be noted that AUC scores are less accurate in samples less than *n*= 200 (Hanczar et al., 2010; Singh, 2013), thus bootstrapping was performed to obtain the range of effects in a larger sample. Point-biserial correlations were reported to denote the direction of the association with the outcome.

To interpret sensitivity and specificity, research literature suggests that for a measure to be useful, the sensitivity + specificity should be at least 1.5 where 1 indicates that the measure is not useful, and 2 indicates that the measure is perfect (Power et al., 2013), however, a low prevalence of the outcome may bias these findings. Thus, sufficient predictive accuracy and sensitivity and specificity were operationalised as AUC values of ≥ 0.70 and sensitivity + specificity of 1.5 to 2, and to avoid bias, prospective and retrospective violence were both analysed in both samples, as both samples had a low base rate of violence prospectively. Due to low power, all findings should be interpreted cautiously (Singh, 2013). ROC analyses were performed in R Studio using the *ROCit* (Khan & Brandenburger, 2014) and *pROC* packages (Robin et al., 2011).

Inpatients. All results for this analysis are presented in Table 7.14. For prospective violence, reasoning, inattention, response inhibition, and affect recognition all had AUC values ≥ 0.70. Sensitivity, or the proportion of correctly identified patients who were violent during the follow-up period was 100% for reasoning, inattention, response inhibition, risk taking, affect recognition and cognitive empathy, however, it should be noted that there were only *n*= 2 patients (6%) who displayed physical violence during the follow-up, which may have inflated the sensitivity, however, this was not the case for lack of insight. All measures showed good sensitivity, though not all showed good specificity, however, as sensitivity predicts patients who were actually violent, sensitivity may be more important. Measures that demonstrated high specificity, or the proportion of patients who were non-violent during the follow-up and were correctly identified as so, were reasoning, response inhibition, affect recognition, and lack of insight. The three measures which demonstrated the best sensitivity, specificity, and accuracy were reasoning, response inhibition, and affect recognition. The cut-off score represents the cut-off score for the measure where the sensitivity and specificity are highest for predicting the outcome. Retrospective violence was also examined as there was a higher prevalence (*n*= 5; 16%). The same measures were analysed and demonstrated slightly lower accuracy, and specificity and sensitivity, however, all measures had sufficient sensitivity + sensitivity except for inattention, response inhibition, response monitoring, and affect recognition. Reasoning continued to be the measure with the highest sensitivity, specificity, and accuracy. For both outcomes, the HCR-20 AUC scores, sensitivity, and specificity were presented to examine the neuropsychological measures in comparison to a validated risk assessment. In both outcomes, the HCR-20 demonstrated excellent scores in all three performance measures.

Community. Like the inpatient sample, the prevalence of violent incidents prospectively was *n*= 2 (6%) which may have introduced bias into the findings, however, only inattention showed sufficient accuracy, sensitivity, and specificity. To investigate whether these findings would be different with a higher base rate or prevalence rate, retrospective violence was analysed as the prevalence was higher (*n*= 13; 42%). Although the base rate for retrospective violence was more than 5 times that of the prospective violence, no AUC scores reached a sufficient AUC value (e.g., AUC= .70), and no measures met the sufficient threshold for sensitivity and specificity, thus, these measures may not be useful for predicting violence in this sample (Table 7.15). Moreover, the bootstrapped confidence intervals for both outcomes were wide, indicating uncertainty in the true effect sizes, further indicating that these measures may not be useful for predicting violence in this population. The LSI-R:SV performance scores demonstrated sufficient sensitivity and specificity for prospective violence; however, the AUC value was under .70, and the 95% confidence intervals suggested that the AUC would not be higher than .65 in a larger population. Retrospectively, the LSI-R:SV had sufficient AUC value, but not sensitivity and specificity, although it was approaching a sufficient value (sensitivity + specificity= 1.45).

Based on these findings, it appears that most of these measures are sufficiently sensitive, however, they may be less specific, and less accurate, thus, in the case of a future study, measures shown to have higher performance indicators for violent outcomes may need to be identified. As violence was the primary interest of the current study, AUC scores for secondary outcomes were not explored.

Table 7.14 Performance Indicators for Measures- Inpatient Sample

	Violence Prospective						Violence Retrospective					
Measures	r^{pb}	AUC	95% CI	Cut-off	Sensitivity	Specificity	r^{pb}	AUC	95% CI	Cut-off	Sensitivity	Specificity
Reasoning	-.26	.83	[.70, .97]	70	1.00	.83	-.38	0.79	[.62, .92]	81	1.00	.70
Inattention	.25	.76	[.47, 1.0]	59	1.00	.56	.20	0.68	[.41, .90]	59	.80	.59
Response Inhibition	-.27	.86	[.73, .97]	42	1.00	.80	-.17	0.66	[.41, .88]	50	1.00	.41
Response Monitoring	-.16	.68	[.37, .96]	50	1.00	.47	-.24	0.70	[.45, .91]	51	1.00	.44
Risk Taking	.05	.63	[.38, .87]	45	1.00	.47	-.18	0.67	[.46, .85]	46	1.00	.56
Affect Recognition	.26	.88	[.73, .97]	25	1.00	.77	-.33	0.70	[.39, .96]	18	.60	.88
Cognitive Empathy	-.08	.62	[.40, .82]	12	1.00	.50	-.28	0.74	[.53, .91]	12	1.00	.56
Lack of Insight	.07	.52	[.10, .93]	5	.52	.80	.44	0.78	[.48, .98]	5	.80	.89
HCR-20	.24	.82	[.57, 1.0]	33	1.00	.63	.45	0.92	[.79, 1.0]	33	.56	.91

Note. Sensitivity and specificity are influenced by base rate and should be interpreted cautiously; r^{pb}= Point biserial correlation; AUC= Area Under the Curve, 95% CI= 95% bootstrapped confidence intervals with 1000 stratified samples.

Table 7.15 Performance Indicators for Measures- Community Sample

	Violence Prospective						Violence Retrospective					
Measures	r^{pb}	AUC	95% CI	Cut-off	Sensitivity	Specificity	r^{pb}	AUC	95% CI	Cut-off	Sensitivity	Specificity
Reasoning	.03	.53	[.07, .94]	106	.50	.85	.22	.62	[.42, .81]	100	.38	.88
Inattention	-.20	.81	[.67, .92]	48	1.00	.78	-.14	.60	[.40, .82]	56	.84	.50
Response Inhibition	.01	.59	[.22, .94]	55	.50	.81	.05	.51	[.27, .70]	42	.15	1.00
Response Monitoring	-.04	.56	[.24, .89]	62	1.00	.33	-.22	.61	[.38, .81]	59	.77	.56
Risk Taking	.12	.68	[.46, .87]	46	1.00	.48	-.25	.65	[.43, .83]	41	.46	.83
Affect Recognition	-.13	.58	[.16, .96]	21	.50	.85	.11	.52	[.31, .72]	27	.15	1.00
Cognitive Empathy	.06	.56	[.26, .85]	12	1.00	.18	.28	.67	[.48, .87]	13	.69	.67
LSI-R:SV	.07	.62	[.12, .65]	4	1.00	.56	-.13	.68	[.41, .91]	4	.75	.70

Note. Sensitivity and specificity are influenced by base rate and should be interpreted cautiously; r^{pb}= Point biserial correlation; AUC= Area Under the Curve; 95% CI= 95% bootstrapped confidence intervals with 1000 stratified samples, LSI-R:SV prospective had *n*=1 incident, and *n*= 17 participants.

7.4.3 Incremental Validity

Incremental validity was examined in measures that achieved an AUC score of ≥.70. To complete this, the independent variance (e.g., standardised residuals) of the neuropsychological measure beyond that explained by the risk assessment was obtained by performing a series of simple linear regression analyses where the risk assessment score was the independent variable, and the neuropsychological measure was the dependent variable (Buffington-Vollum et al., 2002; Edens et al., 2006). ROC analyses were then performed using the standardised residuals, and the AUC value was used to assess the predictive validity of neuropsychological measures for violent outcomes independent of the influence of the HCR-20 or LSI-R:SV (i.e., incremental validity), if the AUC remained ≥.70, it was deemed to have the ability to add incremental variance to the risk assessment. As both samples had a low base rate of prospective violence, to avoid potential bias due to low power, retrospective violence was also explored. The results for the inpatient sample demonstrated that after controlling for the HCR-20, all measures maintained sufficient accuracy for prospective violence, though, the high AUC values accompanied by perfect sensitivity and high specificity implies that these numbers are likely to be skewed based on the low base rate, and therefore, low power. For retrospective violence, only lack of insight maintained a sufficient AUC value, and sufficient sensitivity and specificity. In the community sample, only inattention was examined for prospective violence and resulted in sufficient performance indicators, AUC= 0.81, 95% CI [.62, 1.0]; sensitivity= 1.0; and specificity= 0.81, this finding is also likely to be skewed due to low power. Inpatient results are presented in Table 7.16.

Table 7.16 Performance Indicators for Incremental Validity for Inpatients

	Inpatients			
Measures	AUC	95% CI	Sensitivity	Specificity
Prospective Violence				
Reasoning	.73	[.57, .90]	1.00	.70
Inattention	.72	[.40, 1.0]	1.00	.50
Response Inhibition	.84	[.70, .97]	1.00	.80
Affect Recognition	.93	[.80, 1.0]	1.00	.87
Retrospective Violence				
Reasoning	.59	[.36, .81]	1.00	.33
Inattention	.59	[.28, .87]	.80	.52
Response Monitoring	.60	[.34, .84]	1.00	.33
Affect Recognition	.66	[.34, .96]	.60	.89
Cognitive Empathy	.61	[.39, .81]	1.00	.33
Lack of Insight	.70	[.35, .96]	.80	.85

Note. 95% CI= 95% bootstrapped confidence intervals with 1000 stratified samples.

8 Assessing the Cognitive Contributors to Violence: Part III

Chapter Preface

This chapter presents the findings for all primary, secondary, and exploratory analyses for each sample, and completes the results for this study. The discussion and conclusion for the current study are presented in the following chapter (Chapter 9).

8.1 Primary and Secondary Findings

8.1.1 Power Calculation

A power calculation was performed using G*Power 3.1 (Faul et al., 2007), and to detect a medium effect size with 80% power and alpha set at 0.05 in a sample of *n*=32, a maximum of two predictors can be examined at one time, and to examine two predictors in the community group with the same power, a large effect size can be detected. Due to the small sample size, and the assumption that all data represented meaningful information about the sample, the transformation of data and the removal of outliers were not performed. To account for the small sample, and the possibility of violated assumptions, pre-planned non-parametric bootstrapping was performed for all analyses using *n*= 1,000 replacement samples. Non-parametric bootstrapping does not make any assumptions about the distribution of the data making it a viable alternative to transforming data, especially in smaller samples, and gives a robust estimation of confidence intervals of the wider population, based on the current sample (e.g., Mooney & Duval, 1994). All bootstrapping was conducted to find bias-corrected and accelerated (BCa) confidence intervals, and regressions were bootstrapped using the *case* method, rather than the *residual* method, as the residual method is only valid if the original model meets all the assumptions of regression first. Additionally, due to the pilot and feasibility nature of the study, the researcher was looking for indicators that a larger, future study would be warranted, and variables which explain the data, rather than predict. Significance of findings was assumed when the 95% confidence intervals did not cross zero, and when they did not cross 1 for odds ratios. R^2 was interpreted as small (0.02), medium (0.13), large, (0.26) (Cohen, 1988), and odds ratios were interpreted as small (1.68; 0.60), medium (3.47; 0.29), large (6.71; 0.15) (Chen et al., 2010).

Covariates, including, symptom severity (inpatients only), number of head injuries self-reported, number of head injuries with a loss of consciousness, severity of TBI, and history of substance and alcohol misuse were separately controlled for, as there was not enough power to investigate them cumulatively. Hypotheses were made prior to data analysis and were published in a protocol (e.g., Janes et al., 2021). Inattention and response inhibition were chosen as the two main variables to be used to investigate our hypotheses, as inattention was significantly associated with violence in the meta-analysis, and response inhibition was consistently the top ranked risk factor from the Delphi study. Lack of insight and impulsivity, which were also significant in the meta-analysis were not chosen for primary hypotheses as both are already risk factors on existing risk measures.

8.1.2 Primary Hypothesis 1. **Performance on measures of inattention and response inhibition will explain the variance in prospective violent incidents/offences, over, and above traditional risk assessments.**

Operational Definitions:

- **Inattention**: This was operationalised as the number of omissions on the CPT, where higher scores are indicative of higher impairments. It was expected that higher scores would be associated with increased violent incidents and increased violence severity.
- **Response Inhibition:** This was operationalised as the Interference score on the Stroop, where lower scores are indicative as higher impairments. It was expected that lower scores would be associated with increased violent incidents and increased violence severity.
- **HCR-20:** The HCR-20 is operationalised as the HCR-20 total score collected from patient files at baseline for the inpatient sample.
- **LSI-R:SV:** The LSI-R:SV is operationalised as the number of risk/needs factors on the measure and was collected at baseline for a subset of community participants (*n*=18), however, *n*= 2 community participants were missing prospective outcome data, and *n*=1 of those two participants was in the subset of participants who had a completed LSI-R:SV, and there was only one data point, thus, this analysis could not be performed on the community sample.
- **Prospective Violent Incidents:** Violent incidents were the mean number of physically violent incidents that were recorded during the prospective 6-month follow-up period.
- **Prospective Violence Severity:** Violence severity was operationalised as the mean severity, coded using the Violence Rating Scale (Robertson et al., 1987) of all the aggressive and physically violent incidents that were recorded over the prospective follow-up period.

Analysis. To investigate this hypothesis, it was planned that the mean number of physically violent incidents (prospectively) would be used as the outcome, however, there was a low base rate where only *n*= 2 (6%) community participants exhibited this behaviour during the follow-up period, resulting in naturally binary data. Due to the low base rate, *prospective violence severity* was used to investigate this hypothesis in the community sample, and both outcomes were analysed in the inpatient sample. First zero-order correlations were examined to detect collinearity among variables, next, hierarchical regressions were performed to control for the risk assessment, using two separate linear regression models, which were then compared using an ANOVA. First, a model with the risk assessment only was built to find the amount of individual variance it explained in the outcome, and subsequently a second model was built to include the risk assessment, response inhibition, and inattention to obtain the proportion of variance that response inhibition and inattention explain in the outcome when controlling for the risk assessment. The rationale for this analysis was that if response inhibition and inattention remained significant after controlling for the risk assessment, it would indicate that they can add incremental validity to the risk assessment and may be useful for increasing their predictive performance. To find the proportion of variance explained after controlling for the risk assessment (delta R^2), the R^2 value (proportion of explained variance) from model 2 was subtracted from the R^2 value from model 1 (the model examining the risk assessment alone). An ANOVA was then performed to obtain the F-change statistic, and to indicate whether the additional explained variance was significant.

Results: Inpatient Sample. No variables indicated problems with collinearity (Appendix R). Table 8.1 presents the results of the regressions for the inpatient sample, however, models for both outcomes suggest that the addition of measures of inattention and response inhibition to the HCR-20 did not significantly increase the variance explained and only explained an additional 11% and 2%, respectively, both small effect sizes. The bootstrapped confidence intervals did not indicate that these findings would be significant in a larger population, as all crossed the null line. For the mean frequency of violence, the model for the HCR-20 only was not significant (F (1, 30)= 0.62, p= .44), and neither was the full model (F (3,28)= 1.47, p= .24), and the F-change statistic (F (2, 28)= 0.15, p= .15), further confirmed the non-significant findings. Similar findings emerged for the violence severity outcome. For the HCR-20 alone, the F-test was F (1, 30)= 1.86, p= .18, and for model 2, F (3, 28)= .77, p= 0.52. The adjusted R-squared value decreased from adj. R^2 = .03 in model 1 to adj. R^2 = -.02 in model two, indicating that the addition of the cognitive measures decreased the overall fit of the model. The ANOVA indicated that the F-change statistic was not significant (F (2, 28)= 0.26, p= .77). Durban-Watson was 2.16 and 2.25 respectively, indicating no problems with autocorrelation, tolerance and VIF were both in range for both outcomes, and there were no multivariate outliers. However, the standardized residuals were not normal for both outcomes, and the mean frequency outcome was highly skewed, thus interpretation of the bootstrapped confidence intervals are likely to be more valid than the model estimates. Based on these finding, Hypothesis 1 was not supported by the data in the inpatient sample.

Table 8. 1 Results for Inpatients with Mean Frequency and Mean Severity of Prospective Violence

Variable	*B*	95% BCa CI	*β*	95% BCa CI	sr^2	Fit	Difference
Mean Frequency of Violence							
(Intercept)	-0.08	[-0.43, 0.00]					
HCR-20	0.00	[0.00, 0.02]	0.14	[0.05, 0.26]	.02		
						R^2 = .02 95% CI [.00, .05]	
(Intercept)	-0.07	[-1.37, 0.09]					
HCR-20	0.00	[-0.00, 0.01]	0.04	[-0.17, 0.29]	.00		
Inattention	0.00	[-0.00, 0.02]	0.29	[-0.15, 0.65]	.08		
Response Inhibition	-0.00	[-0.03, 0.00]	-0.11	[-0.27, 0.14]	.01		
						R^2 = .13 BCa CI [.00, .20]	ΔR^2 = .11
Mean Severity of Violence							
(Intercept)	-0.31	[-1.39, 0.63]					
HCR-20	0.02	[-0.01, 0.06]	0.24	[-0.18, 0.50]	.06		
						R^2 = .059 95% CI [.00, .27]	
(Intercept)	0.47	[-1.20, 3.07]					
HCR-20	0.02	[-0.01, 0.06]	0.24	[-0.17, 0.49]	.06		
Inattention	-0.00	[-0.02, 0.01]	-0.10	[-0.35, 0.29]	.01		
Response Inhibition	-0.01	[-0.06, 0.02]	-0.11	[-0.35, 0.22]	.01		
						R^2 = .076 BCa CI [.00, .20]	ΔR^2 = .017

Note. B= Unstandardised Beta, 95% BCa CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, 95% BCa CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

Results: Community Sample. Zero-order correlations showed no signs of collinearity (Appendix T). Results of the regression (Table 8.2) showed that the addition of measures of inattention and response inhibition added approximately 5% of variance to the LSI-R:SV, a small effect size. There were no indications of autocorrelation (Durbin-Watson= 2.3), multicollinearity, or multivariate outliers, however, the residuals were not normal even after bootstrapping, thus findings are likely to be biased. Bootstrapped confidence intervals for the model estimates and R^2 did not indicate that these results would reach significance in a larger population. Neither model was significant (LSI-R:SV only model (F (1, 15)= 0.08, p= .78; model 2: (F (3, 13)= 0.28, p= .84), nor was the F-change statistic (F (2, 13) = 0.37, p= .69).

Table 8. 2 Results for Community Sample with Prospective Violence Severity

Variable	B	BCa 95% CI	β	BCa 95% CI	sr^2	Fit	Difference
(Intercept)	0.04	[-0.19, 0.64]					
LSI-R:SV	0.01	[0.00, 0.25]	0.07	[-0.17, 0.24]	.00		
						R^2 = .01 95% CI [.00, .06]	
(Intercept)	-0.15	[-3.41, 0.35]					
LSI-R:SV	0.02	[-0.01, 0.19]	0.08	[-0.28, 0.34]	.01		
Inattention	-0.00	[-0.02, 0.00]	-0.18	[-0.34, 0.11]	.03		
Response Inhibition	0.01	[0.00, 0.08]	0.15	[-0.14, 0.41]	.02		
						R^2 = .06 95% CI [.02, .11]	ΔR^2 = .05

Note. B= Unstandardised Beta, 95% BCa CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, 95% BCa CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

The following set of hypotheses investigated secondary outcomes, namely, *prospective perceived risk* which included *length of admission* in secure settings at the time of consent, *mean level of observation*, and mean number of *days on elevated levels of observation* over the 6-month follow-up period in inpatients only, followed by *aggression*, *impulsivity*, and retrospective outcomes in both samples.

8.1.3 Secondary Hypothesis 1. **Performance on measures of inattention and response inhibition will explain the variance in perceived risk- <u>Inpatient sample only</u>.**

Operational Definitions: Secondary Hypothesis 1

- **Length of Admission:** This variable is operationalised as the length of admission in months that the patient has been in a secure setting at the time of consent.
- **Prospective Mean Level of Observation:** Level of observation was operationalised by the mean level of observation that the inpatients were on over the 6-month follow-up periods. Levels of observation are indicative of risk, where, as levels of observation increase, often patient privileges are revoked until they are on a decreased level.

- **Prospective Mean Days on Enhanced Levels:** Mean days on enhanced levels is operationalised by the mean number of days that a patient was on enhanced levels of observation as a measure of perceived risk.

Analysis and Results: Secondary Hypothesis 1. Zero-order correlations were first examined, and then multiple linear regressions were performed where independent variables were entered simultaneously. Correlations indicated no problems, and the only significant correlation found was between mean level of observations and mean days on elevated levels, which was expected (r= 1.00). The results of the regressions (Table 8.3) did not support the hypothesis for any of the perceived risk outcomes, and all effect sizes were small. The F-tests for the regressions were as follows,

1. Length of admission: (F (2, 29) = 0.41, p= .67)
2. Level of observation: (F (2, 29) = 0.72, p= .49)
3. Days on enhanced levels: (F (2, 29) = 0.61, p= .55).

All models met model assumptions except for normality, however, bootstrapping improved this.

Table 8.3 Results for Prospective Perceived Risk in Inpatients

Variable	B	BCa 95% CI	β	BCa 95% CI	sr^2	Fit
Length of Admission						
(Intercept)	71.74	[-34.46, 199.47]				
Inattention	-0.41	[-1.11, 0.23]	-0.16	[-0.34, 0.42]	.03	
Response Inhibition	-0.31	[-2.76, 1.81]	-0.05	[-0.36, 0.39]	.00	
						R^2 = .027 BCa CI [.00, .09]
Mean Level of Observation						
(Intercept)	1.68**	[1.07, 2.78]				
Inattention	-0.00	[-0.01, 0.00]	-0.16	[-0.41, 0.04]	.02	
Response Inhibition	-0.01	[-0.02, 0.00]	-0.18	[-0.42, 0.76]	.03	
						R^2 = .048 BCa CI [.00, .16]
Mean Days on Levels						
(Intercept)	19.38	[1.02, 53.32]				
Inattention	-0.08	[-0.29, 0.03]	-0.14	[-0.38, 0.09]	.02	
Response Inhibition	-0.23	[-0.76, 0.05]	-0.17	[-0.04, 0.07]	.03	
						R^2 = .041 BCa CI [.00, .23]

Note. *B*= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.
* indicates p < .05.

** indicates p < .01.

8.1.4 Secondary Hypothesis 2. **Performance on measures of inattention and response inhibition will explain the variance in, or will be associated with, prospective anger provoked incidents.**

Operational Definitions:

- **Anger.** Anger was operationalised as the mean number non-contact incidents, namely, verbal aggression, damage to property, and intimidating behaviours, over the 6-month follow-up periods. In the inpatient sample, there were sufficient data to use anger as a continuous variable, however, in the community sample, the data were naturally binary, and there were only two incidents, thus anger could not be analysed in the community sample.

Analysis. A multiple linear regression was performed to examine the hypothesis in the inpatient sample, where both independent variables were entered simultaneously. Normality was violated, though it was corrected following bootstrapping.

Results. Inattention and response inhibition explained 3% of the variance in anger (F (2, 29) = 0.49, p= .61) in the inpatient sample, and the 95% confidence intervals did not indicate that in a larger population these results would differ (Table 8.4). The findings do not indicate that that these variables are important indicators of anger in this sample.

Table 8.4 Results for Prospective Anger in Inpatients

Variable	B	BCa 95% CI	β	BCa 95% CI	sr^2	Fit
(Intercept)	0.49	[-1.19, 3.31]				
Inattention	0.01	[-0.01, 0.04]	0.13	[-0.20, 0.54]	.02	
Response Inhibition	-0.01	[-0.07, 0.01]	-0.10	[-0.35, 0.18]	.01	
						R^2 = .033 BCa CI [.00, .16]

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

8.1.5 Secondary Hypothesis 3: **Performance on measures of inattention and response inhibition will explain the variance in impulsivity.**

Operational Definitions:

- **Impulsivity.** Impulsivity is operationalised by the total score on the BIS-11 questionnaire. The BIS-11 is a 30 item self-report measure of impulsivity and was used as a proxy for violence in the current study. As impulsivity is already a well-established risk factor for violence, it was posited that measures which explain the variance in impulsivity, may also explain the variance in violence. The BIS-11 was administered to all participants in both samples at baseline.

Analysis and Results: Multiple linear regressions were performed for both samples, with both independent variables entered simultaneously. In the inpatient sample, measures of inattention and response inhibition explained 7% of the variance in self-reported impulsivity (F (2, 29) = 1.72, p= .32), a small effect size. The model for the community offenders explained 15% of the variance, a medium effect size, and inattention emerged as a potentially important explanatory variable for self-reported impulsivity (F (2, 28) = 2.56, p= .09). Further, inattention was in the expected direction indicating that as inattention increased, impulsivity also increased (Table 8.5). All model assumptions were met for both samples. The hypothesis was partially supported in the community sample.

Table 8. 5 Results for Self-Reported Impulsivity in Both Sample

Variable	B	BCa 95% CI	β	BCa 95% CI	sr^2	Fit
Forensic Inpatients						
(Intercept)	49.41**	[13.75, 84.17]				
Inattention	-0.02	[-0.27, 0.11]	-0.03	[-0.29, 0.21]	.00	
Response Inhibition	0.39	[-0.18, 1.69]	0.26	[-0.24, 0.55]	.07	
						R^2 = .075 95% CI [.00, .30]
Community Offenders						
(Intercept)	71.73***	[39.94, 100.60]				
Inattention	0.18*	[0.04, 0.49]	0.33	[0.00, 0.62]	.11	
Response Inhibition	-0.22	[-0.78, 0.36]	-0.17	[-0.59, 0.27]	.03	
						R^2 = .15 BCa CI [.10, .41]

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = BCa 95% bootstrapped confidence intervals, sr^2= semi-partial correlation squared.
* indicates p < .05.

** indicates p < .01.

8.1.6 Secondary Hypothesis 4: **Performance on measures of inattention and response inhibition will explain the variance in retrospective violent incidents, over and above the HCR-20 (Inpatients).**

Operational Definitions: Secondary Hypothesis 4

- **Retrospective Violent Incidents:** Violent incidents were the mean number of physically violent incidents that were recorded during the retrospective 6-month follow-up period (6- month period before consent).
- **Retrospective Violence Severity:** Violence severity was operationalised as the mean severity, coded using the Violence Rating Scale (Robertson et al., 1987) of all the aggressive and physically violent incidents that were recorded over the retrospective follow-up period.

Analysis: Secondary Hypothesis 4. Hierarchical regressions were performed using two separate linear regression models, which were then compared using an ANOVA. The procedures were mirrored from the same analysis with prospective outcomes in the primary hypothesis.

Results: Secondary Hypothesis 4- Inpatient sample only. There were n= 5 inpatients who displayed physical violence during the retrospective follow-up period, ranging from 1-7 incidents per patient. The results in Table 8.6 show that the HCR-20 explained 8% of the variance in the mean frequency of physically violent incidents retrospectively (F (1, 20) = 2.55, p= .12), however the addition of inattention and response inhibition explained only an additional 0.6% (F (3, 28) = 0.86, p= .47), and the F-change statistic was not significant (F (2, 28) = 0.08, p= .92). When violence severity (retrospective) was analysed, the HCR-20 explained a significant proportion of the variance with a large effect size (F (1, 30) = 14.88, p <.001), however, the addition of inattention and response inhibition explained only 2% of variance when the HCR-20 was controlled for (F (3, 28) = 5.03, p= .01). The F-change statistic was not significant (F (2, 28) = 0.40, p= .67). Thus, Secondary Hypothesis 4 was not supported by these findings. Further, adjusted R^2 decreased when the cognitive measures were added, indicating that their addition decreased the model fit.

Table 8. 6 Results for Retrospective Outcomes in Inpatients

Variable	*B*	BCa 95% CI	*β*	BCa 95% CI	sr^2	Fit	Difference
Mean Frequency of Violence							
(Intercept)	-0.19	[-0.45, -0.06]					
HCR-20	0.01	[0.00, 0.02]	0.28	[0.13, 0.49]	.08		
						R^2 = .078 95% CI [.02, .23]	
(Intercept)	-0.32	[-1.59, 0.03]					
HCR-20	0.01	[0.00, 0.02]	0.26	[0.09, 0.49]	.07		
Inattention	0.00	[-0.00, 0.01]	0.07	[-0.28, 0.30]	.00		
Response Inhibition	0.00	[-0.01, 0.02]	0.04	[-0.22, 0.29]	.00		
						R^2 = 0.084 BCa CI [.02, .17]	ΔR^2 = .006
Mean Severity of Violence							
(Intercept)	-1.28*	[-2.24, -0.62]					
HCR-20	0.06**	[0.03, 0.09]	0.58	[0.36, 0.73]	.33		
						R^2 = .332** 95% CI [.10, .52]	
(Intercept)	-0.34	[-2.55, 1.71]					
HCR-20	0.06**	[0.03, 0.09]	0.54	[0.32, 0.75]	.29		
Inattention	-0.00	[-0.02, 0.01]	-0.04	[-0.33, 0.23]	.00		
Response Inhibition	-0.02	[-0.05, 0.02]	-0.14	[-0.38, 0.13]	.02		
						R^2 = .350** BCa CI [.09, .52]	ΔR^2 = .018

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = BCa 95% bootstrapped confidence intervals, sr^2= semi-partial correlation squared.
* indicates $p < .05$.
** indicates $p < .01$.

8.1.7 Secondary Hypothesis 5. **Performance on measures of inattention and response inhibition will demonstrate incremental validity over and above the LSI-R:SV, for retrospective violent incidents and violence severity.**

Operational Definitions: Secondary Hypothesis 5

- **Retrospective Violent Incidents:** Violent incidents were naturally binary in the community sample; thus, violent incidents were coded as 1, and non-violent incidents were coded as 0.

- **Retrospective Violence Severity:** The operational definition for violence severity in Secondary Hypothesis 4 was maintained for this analysis.

Analysis: Secondary Hypothesis 5. To analyse the binary outcome, a hierarchical logistic regression was performed. A procedure similar to that used for the linear hierarchical regression was used, where a model containing only the risk assessment was entered first, and then all variables were entered together. An ANOVA was performed to determine whether the addition of the variables significantly improved the model fit. To examine the continuous outcome, a linear hierarchical regression was performed using the same procedures as were employed for Secondary Hypothesis 4.

Results: Secondary Hypothesis 5. There was n=8 retrospective violent incidents. The results for the logistic regression examining secondary hypothesis 5 are presented in Table 8.7. The model examining only the LSI-R:SV was not significant (χ^2 (1) = 2.23, p= .14), and neither was the model with all the variables (χ^2 (3) = 3.14, p= .37). Further, the addition of the cognitive measures to the model did not significantly improve it (χ^2 (2) = 0.90, p= .63). All assumptions were met for the logistic regression. All odds ratios were small in magnitude and pseudo R^2 were also small suggesting that the model does not explain the data well.

The results for the linear regression examining violence severity can be found in Table 8.8. Neither the first model (F (1, 16) = 1.24, p= .28) or the second model with all predictors (F (3, 14) = 0.62, p= .61), were significant. Similarly, the addition of the cognitive measures did not improve the model (F (2, 15)= 0.36, p= .70), and explained only an additional 5% of the variance in the outcome. Effect sizes were small, and the findings do not support the hypothesis.

Table 8. 7 Results for Retrospective Violent Incidents in the Community Sample

Variable	*Estimate*	BCa 95% CI	*OR*	BCa 95% CI	Fit
(Intercept)	1.26	[-1.24, 5.00]			
LSI-R:SV	-0.43	[-1.32, 0.37]	0.65	[0.26, 1.45]	
					Pseudo R^2 = .09
(Intercept)	-3.06	[-42.92, 41.33]			
LSI-R:SV	-0.43	[-13.58, 0.61]	0.65	[0.00, 1.89]	
Inattention	0.01	[-0.67, 0.37]	1.02	[0.51, 1.45]	
Response Inhibition	0.07	[-4.13, 0.36]	1.07	[0.02, 1.44]	
					Pseudo R^2 = .13

Note. BCa 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for the model estimate (1,000 replicates); *OR*= Odds Ratio, *BCa* 95% CI= BCa bootstrapped confidence intervals for the odds ratio, *Pseudo-*R^2= McFadden's Pseudo-R^2.

Table 8.8 Results for Retrospective Violence Severity in the Community Sample (*n*=18).

Variable	*B*	BCa 95% CI	*β*	BCa 95% CI	sr^2	Fit	Difference
(Intercept)	1.57*	[0.36, 2.43]					
LSI-R:SV	-0.16	[-0.34, 0.22]	-0.27	[-0.67, 0.33]	.07		
						R^2 = .07 95% CI [.00, .38]	
(Intercept)	-0.34	[-4.38, 4.91]					
LSI-R:SV	-0.15	[-0.37, 0.24]	-0.19	[-0.69, 0.25]	.07		
Inattention	0.00	[-0.06, 0.03]	0.01	[-0.42, 0.54]	.00		
Response Inhibition	0.03	[-0.08, 0.07]	0.01	[-0.25, 0.49]	.04		
						R^2 = .12 BCa CI [.00, .29]	ΔR^2 = .05

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); *β*= Standardised Beta, *β* 95% CI= Bca bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% Bca bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

8.1.8 Secondary Hypothesis 6. **Performance on measures of inattention and response inhibition will explain the variance in retrospective perceived risk- <u>Inpatient sample only.</u>**

Operational Definitions: Secondary Hypothesis 6

- **Retrospective Mean Level of Observation:** Level of observation was operationalised by the mean level of observation that the inpatients were on over the retrospective 6-month follow-up period.
- **Retrospective Mean Days on Enhanced Levels:** Mean days on enhanced levels is operationalised by the mean number of days that a patient was on enhanced levels of observation retrospectively as a measure of perceived risk.

Analysis and Results: Secondary Hypothesis 6. Multiple linear regressions were performed for both outcomes, and both independent variables were entered simultaneously. The results are presented in Table 8.9. The assumption of normality was violated, but was corrected following bootstrapping, all other assumptions were met. Inattention and response inhibition explained 8% of the variance in mean level of observations (*F* (2, 29) = 1.17, *p*= .33), and 11% in mean number of days on elevated levels (*F* (2, 29) = 1.80, *p*= .18), both small effect sizes. The confidence intervals for response inhibition in the second model indicated that in a larger population, it may be an important explanatory variable for the number of days a person was on elevated levels of observation, however, they are wide suggesting uncertainty in the true estimates.

Table 8.9 Results for Retrospective Perceived Risk for Inpatients

Variable	B	BCa 95% CI	β	BCa 95% CI	sr^2	Fit
Mean Level of Observation						
(Intercept)	1.89**	[0.59, 2.89]				
Inattention	0.00	[-0.01, 0.02]	0.10	[-0.41, 0.04]	.01	
Response Inhibition	-0.02	[-0.04, 0.00]	-0.22	[-0.42, 0.76]	.05	
						R^2 = .075 BCa CI [.00, .28]
Mean Days on Levels						
(Intercept)	21.94	[-3.27, 45.59]				
Inattention	0.07	[-0.12, 0.40]	0.10	[-0.22, 0.46]	.01	
Response Inhibition	-0.42*	[-0.88, -0.05]	-0.28	[-0.53, -0.01]	.08	
						R^2 = .11 BCa CI [.01, .36]

Note. *B*= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

8.1.9 Secondary Hypothesis 7. **Performance on measures of inattention and response inhibition will explain the variance in retrospective anger provoked incidents.**

Operational Definitions:

- **Retrospective Anger:** Aggression was operationalised as the mean number non-contact incidents, namely, verbal aggression, damage to property, and intimidating behaviours, over the retrospective 6-month follow-up period. In the inpatient sample, there were sufficient data to use aggression as a continuous variable, however, in the community sample, the data were naturally binary, and only had two incidents, thus, this analysis could not be performed for the community sample.

Analysis. A multiple linear regression was used to analyse aggression retrospectively for the inpatient sample. Bootstrapping was employed to account for violated assumptions of normality. A logistic regression was used to examine the association between inattention and response inhibition and aggression, retrospectively in the community sample. All model assumptions were met.

Results. Inattention and response inhibition explained 6% of the variance in the inpatient sample (Table 8.10) a small effect size, (F (2, 29) = 0.89, p= .42). The hypothesis was not supported, and the spread of the confidence intervals suggests that these findings would not differ in a larger sample.

Table 8.10 Results Mean Frequency of Retrospective Anger in Inpatients

Variable	*B*	BCa 95% CI	β	BCa 95% CI	sr^2	Fit
(Intercept)	-0.01	[-0.96, 0.67]				
Inattention	0.00	[-0.0, 0.02]	0.22	[-0.13, 0.58]	.05	
Response Inhibition	-0.00	[-0.02, 0.01]	-0.05	[-0.32, 0.19]	.00	
						R^2 = .058 BCa CI [.00, .34]

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

8.2 Exploratory Analyses

Operational Definitions: Exploratory Analyses for the Inpatient Sample

- **Affect Recognition:** Affect recognition was operationalised as the total correctly identified positive and negative emotions on the TASIT. The TASIT is a measure of social cognition where participants watch a series of vignettes of individuals expressing certain emotions. Immediately following the vignette, the participant is asked to identify the emotion that the actor was displaying. All emotions are listed on a series of cue-cards, each one presenting the same emotions in a different order.
- **Positive Affect Recognition:** Positive affect recognition is operationalised as the total number of correctly identified positive emotions on the TASIT.
- **Negative Affect Recognition:** Negative affect recognition is operationalised as the total number of correctly identified Negative emotions on the TASIT.
- **Lack of Insight:** Lack of insight is operationalised as the total score from the lack of insight and judgement scale on the PANSS. The PANSS is a clinical rating scale that follows a structured interview that is routinely administered on a 6-monthly basis at the State Hospital. N= 4 inpatients had a recently completed PANSS, which were recorded from their electronic medical records, however, the researcher administered the remaining n= 28 at the time of baseline testing.

- **Prospective Violence Severity:** Prospective Violence severity was used as the dependent variable, as the main aims of the study are to explain prospective violent outcomes.

Analyses. Several exploratory analyses were undertaken to find a combination of neuropsychological measures which may explain the variance in prospective outcomes. Due to the number of independent variables, and to avoid multiple testing, a best subsets regression was used to identify up to two independent variables which explain the outcomes best. The best subsets regression is a type of stepwise analysis and is a regression model selection approach that involves testing all possible combinations of the independent variables, and then selects the best model according to statistical criteria (Kassambara, 2017). To perform this analysis, a subset of the data was created, and the function *regsubsets()* in the *leaps* package in R Studio was used (Lumley, 2020). A model was built using the function with prospective violence severity as the dependent variable, and no independent variables, and a maximum of two variables to be included in the model was specified. Zero-order correlations gave no indications of multicollinearity. Next, multiple linear regressions were performed and both independent variables were entered simultaneously. A hierarchical regression was then performed to examine whether the two variables explain the variance in violence severity over and above the HCR-20 and demonstrate incremental validity. Bootstrapping corrected the violated assumptions of normality in both models (Appendix W).

Results. The best subsets regression identified lack of insight and affect recognition as important variables and resulted in a significant regression model (F (2, 29) = 4.05, p= .03) where both variables individually explained a significant proportion of the variance, and 22% of the variance overall, a medium effect size (Table 8.11). Adjusted R^2= .16 suggesting that the model would explain slightly less variance in a population, as opposed to the current sample. The results of the hierarchical regression are presented in Table 8.12 and show that the model explained the variance in violence severity over and above the HCR-20 (F (3, 28) = 3.02, p= .04), though lack of insight was no longer significant in the presence of the HCR-20. Once the two variables were added to the regression with the HCR-20, they explained an additional 18% of variance in the outcome, a medium effect size (F (2, 29) = 3.45, p= .04). Adjusted R^2 for the HCR-20 only model was .03, and was .16 for the full model, indicating that in a population, the addition of affect recognition and insight to the HCR-20 would explain an additional 13% of the variance in violence severity.

Table 8. 11 Exploratory Regression for Prospective Violence Severity in Inpatients

Variable	*B*	*B* 95% CI	*β*	*β* 95% CI	sr^2	Fit
(Intercept)	-1.27*	[-3.55, -0.03]				
Affect Recognition	0.06*	[0.01, 0.14]	0.35	[0.00, 0.59]	.18	
Lack of Insight	0.17*	[0.03, 0.39]	0.41	[0.02, 0.70]	.10	R^2 = .22* 95% CI [.01, .50]

Note. B= Unstandardised Beta, *B* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); *β*= Standardised Beta, *β* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; 95% CI for R^2 = BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.
* indicates $p < .05$.

** indicates $p < .01$.

Table 8. 12 Exploratory Regression for Prospective Violence Severity in Inpatients

Variable	*B*	BCa 95% CI	*β*	BCa 95% CI	sr^2	Fit	Difference
(Intercept)	-0.31	[-1.39, 0.63]					
HCR-20	0.02	[-0.01, 0.07]	0.24	[-0.16, 0.51]	.06		
						R^2 = .06 BCa CI [.00, .23]	
(Intercept)	-0.34	[-4.02, 0.06]					
HCR-20	0.02	[-0.02, 0.05]	0.16	[-0.20, 0.47]	.03		
Affect Recognition	0.06*	[0.00, 0.13]	0.36	[0.01, 0.60]	.14		
Lack of Insight	0.15	[0.01, 0.37]	0.33	[-0.01, 0.58]	.13		
						R^2 = .24* BCa CI [.01, .51]	ΔR^2 = .18*

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1,000 replicates); *β*= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared
* indicates $p < .05$.

** indicates $p < .01$.

8.2.1 Additional Exploratory Analyses for the Inpatient Sample

Several covariates were used to test the robustness of lack of insight and affect recognition.

Operational Definitions: Additional Analyses

- **Positive Symptoms:** Positive symptoms was operationalised by the PANSS positive symptoms scale. This was used in place of the whole PANSS measure (symptom severity), as lack of insight is counted in the total score for the PANSS, and they have a correlation of *r* (30)= 0.75, *p* <0.001, which would likely artificially inflate the results of the regression. Moreover, as all the inpatients had a diagnosis of a psychotic illness, controlling for positive symptoms was appropriate.

- **Head Injuries with LOC:** Number of self-reported head injuries with a loss of consciousness. This was measured using the OSU-TBI-ID semi-structured interview. All the patients in medium and low secure units were given the interview at the time of baseline testing, and the majority of the interviews for patients in the high secure setting were recorded from their electronic file.
- **Number of Head Injuries:** Number of self-reported head injuries, recorded by the OSU-TBI-ID measure.
- **Severity of Head Injuries:** The worst self-reported head injury for each participant was rated on a 5-point Likert scale ranging from, no history of TBI or very minor head injury to severe TBI.
- **Alcohol use:** Patients with a history of alcohol use, harmful use, and/or a diagnosis of alcohol dependence, were coded as 1, and those with no history were coded as 0.
- **Substance use:** Patients with a history of substance misuse, harmful use, and/or a diagnosis of drug dependence, were coded as 1, and those with no history were coded as 0.

Analyses and Results. Zero-order correlations revealed that positive PANSS and insight were significantly correlated r (30)= 0.64, p <.001, however, correlations below .80-.90 pose little threat to regression models (Field et al., 2012), further VIF was 5.5 and 5.7, respectively indicating no problems with multicollinearity. All other correlations were under 0.5. A series of multiple linear regressions were performed to examine the covariates one at a time. All regression results are presented in Table 8.13.

1. First, the positive PANSS score was added to the model with lack of insight and affect recognition to control for positive symptoms, and the model remained significant, explaining 25% of the variance in violence severity (Adj. R^2= .17), a medium effect size (F (3, 28) = 3.17, p= .04). Lack of insight was no longer significant in the presences of positive symptoms. The slight increase in Adj. R^2 from the model with just affect recognition and lack of insight suggests that positive PANSS slightly improved the model fit.
2. Head injuries with LOC was then added to the model, however, the model was no longer significant, though both affect recognition and lack of insight were significant following bootstrapping (F (3, 27) = 2.64, p= .07). Adj. R^2 decreased from 0.16 from the initial model to 0.14 in this model, suggesting that the addition of head injuries with LOC did not improve the model.

3. Number of head injuries was examined subsequently, and affect recognition and lack of insight remained significant, as well as the model (F (3, 27) = 3.10, p= .04). Head injuries individually explained only 4% of the variance in the outcome in this model, compared to affect recognition which explained 14% and lack of insight explaining 19%, both medium in magnitude. The model explained 26% of the variance in the outcome. Adj. R^2 increased to 0.17 indicating that the addition of number of head injuries improved the model slightly.
4. Following this, severity of head injuries was added to the model with affect recognition and lack of insight, and the model was not significant (F (3, 27) = 2.64, p= .07), however, both affect recognition and lack of insight were. Adj. R^2 decreased to 0.14 suggesting that severity of head injuries did not improve the model.
5. History of alcohol use was examined next, resulting in a significant model (F (3, 28) = 3.23, p= .04), however, only affect recognition and lack of insight were significant. Adj. R^2 increased to 0.18 with the addition of alcohol use, improving the model.
6. History of substance use was explored last and resulted in a significant model (F (3, 28) = 4.44, p= .01), and all variables were significant. Together the three variables explained 32% of the variance in violence severity. The addition of substance use increased Adj. R^2 to .25 suggesting it improved the model, and increased the explained variance by 10%. This suggests that violence severity is at least partially explained by a history of substance use.

These findings indicate that in the presence of all three head injury variables and alcohol use, affect recognition and lack of insight remained significant explanatory variables for violence severity, and as the beta values did not markedly increase or decrease, these variables do not significantly confound the relationship between affect recognition and lack of insight and violence severity. The addition of positive symptoms appeared to wash out the significance of lack of insight, decreasing the individual explained variance, and the addition of substance misuse increased the variance explained by the model as well as lack of insight, suggesting that it may be an important explanatory variable for severity of violence. Affect recognition did not seem to be significantly influenced by the addition of any control variables. Aside from substance use, only the addition of positive PANSS, number of head injuries, and history of alcohol use improved the model fit despite not being significant.

Table 8.13 Exploratory Regression for Prospective Violence Severity in Inpatients

	Variable	*B*	BCa 95% CI	*β*	BCa 95% CI	sr^2	Fit
1.	(Intercept)	-1.45*	[-3.43, -0.11]				
	Affect Recognition	0.06*	[0.01, 0.14]	0.35	[-0.01, 0.60]	.12	
	Lack of Insight	0.11	[-0.07, 0.39]	0.21	[-0.06, 0.53]	.04	
	Positive PANSS	0.03	[-0.03, 0.13]	0.19	[-0.19, 0.51]	.04	
							R^2 = .25* BCa CI [.03, .54]
2.	(Intercept)	-1.47	[-3.05, 0.11]				
	Affect Recognition	0.05*	[0.00, 0.13]	0.33	[0.04, 0.55]	.11	
	Lack of Insight	0.17*	[0.01, 0.37]	0.40	[0.04, 0.69]	.16	
	Head Injuries LOC	-0.06	[-0.21, 0.09]	-0.12	[-0.37, 0.13]	.01	
							R^2 = .23 BCa CI [.03, .55]
3.	(Intercept)	-1.26*	[-3.18, -0.16]				
	Affect Recognition	0.06*	[0.01, 0.14]	0.38	[0.09, 0.60]	.14	
	Lack of Insight	0.19*	[0.03, 0.40]	0.44	[0.05, 0.72]	.19	
	Head Injuries	-0.07	[-0.18, 0.02]	-0.21	[-0.44, 0.10]	.04	
							R^2 = .26* BCa CI [.06, .59]
4.	(Intercept)	-1.19	[-3.02, 0.07]				
	Affect Recognition	0.06*	[0.01, 0.14]	0.36	[0.07, 0.60]	.13	
	Lack of Insight	0.17*	[0.01, 0.36]	0.40	[0.04, 0.69]	.16	
	Head Injuries Severity	-0.05	[-0.21, 0.10]	-0.12	[-0.37, 0.23]	.01	
							R^2 = .23 BCa CI [.02, .51]
5.	(Intercept)	-1.50*	[-3.93, -0.09]				
	Affect Recognition	0.06*	[0.00, 0.15]	0.36	[0.02, 0.62]	.14	
	Lack of Insight	0.21**	[0.05, 0.42]	0.46	[0.16, 0.72]	.21	
	Alcohol Use	0.33	[-0.11, 0.93]	0.20	[-0.04, 0.57]	.04	
							R^2 = .26* BCa CI [.02, .50]
6.	(Intercept)	-1.83*	[-4.01, -0.58]				
	Affect Recognition	0.07*	[0.02, 0.15]	0.43	[0.11, 0.67]	.18	
	Lack of Insight	0.20**	[0.06, 0.41]	0.46	[0.14, 0.73]	.21	
	Substance Use	0.43*	[0.03, 0.87]	0.32	[0.004, 0.64]	.10	
							R^2 = .32* BCa CI [.02, .49]

Note. B= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1000 replicates); *β*= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals. LOC= Loss of consciousness, sr^2= semi-partial correlation squared.

* indicates $p < .05$.

** indicates $p < .01$.

As affect recognition is operationalised as the number of positive and negative emotions correctly identified, positive and negative emotions were each investigated separately with lack of insight. The results of the regression examining the relationship between lack of insight, positive emotions and violence severity are presented in Table 8.14. First a model with only lack of insight and positive emotions were analysed and resulted in a significant model (F (2, 29) = 5.27, p = .01) explaining 27% of the variance in violence severity, a large effect size. Adj. R^2= .21, indicating that if the model were derived from the population rather than this sample, it would explain approximately 6% less variance in violence severity. Covariates were assessed individually resulting in lack of insight and positive affect recognition remaining significant in the presence of head injuries with LOC (F (3, 27) = 3.64, p = .03), number of head injuries (F (3, 27) = 4.28, p = .01), severity of head injuries (F (3, 27) = 3.82, p = .02), history of alcohol use (F (3, 28) = 4.18, p = 0.01), history of substance misuse (F (3, 28) = 5.66, p = .004). When positive symptoms of psychosis were added to the model, only positive affect recognition remained significant (F (3, 28) = 3.74, p = .02). Large effect sizes were found in all models, though adjusted R^2 suggested that only the addition of number of head injuries (adj. R^2= .25), history of alcohol use (adj. R^2= .23), severity of head injuries (adj. R^2= .22), and history of substance use (adj. R^2= .31) improved the model, despite only history of substance use being significant. As the model with history of substance use, positive emotion recognition, and lack of insight explained 37% of the variance, as opposed to 32% with affect recognition, positive emotion recognition may be a better explanatory variable. The model met all assumptions except for homoscedasticity and normality, though these were corrected following bootstrapping. Notably though, the confidence intervals for all effect sizes ranged from small to large, indicating some uncertainty in the true effects.

Table 8.14 Exploratory Regression for Prospective Violence Severity in Inpatients

Variable	B	BCa 95% CI	β	BCa 95% CI	sr^2	Fit
(Intercept)	-1.28*	[-3.29, -0.38]				
Lack of Insight	0.18*	[0.04, 0.37]	0.42	[0.07, 0.66]	.17	
Positive Affect Rec.	0.13*	[0.04, 0.32]	0.41	[0.14, 0.63]	.16	
						R^2 = .27* BCa CI [.03, .54]
(Intercept)	-1.39*	[-3.04, -0.31]				
Lack of Insight	0.13	[-0.03, 0.37]	0.24	[-0.01, 0.60]	.06	
Positive Affect Rec.	0.13*	[0.04, 0.28]	0.39	[-0.06, 0.53]	.16	
Positive PANSS	0.02	[-0.03, 0.11]	0.14	[-0.19, 0.51]	.02	
						R^2 = .28* BCa CI [.03, .56]
(Intercept)	-1.21*	[-2.55, -0.28]				
Lack of Insight	0.18*	[0.03, 0.36]	0.42	[0.09, 0.68]	.18	
Positive Affect Rec.	0.13*	[0.04, 0.28]	0.41	[0.10, 0.61]	.17	
Head Injuries LOC	-0.09	[-0.27, 0.06]	-0.17	[-0.41, 0.07]	.03	
						R^2 = .29* BCa CI [.03, .54]
(Intercept)	-1.30*	[-2.68, -0.38]				
Lack of Insight	0.19**	[0.06, 0.37]	0.46	[0.06, 0.69]	.21	
Positive Affect Rec.	0.16**	[0.05, 0.29]	0.46	[0.19, 0.69]	.21	
Head Injuries	-0.08	[-0.19, 0.01]	-0.25	[-0.49, 0.07]	.06	
						R^2 = .32* BCa CI [.06, .59]
(Intercept)	-1.26*	[-2.68, 0.27]				
Lack of Insight	0.18*	[0.04, 0.35]	0.43	[0.03, 0.65]	.18	
Positive Affect Rec.	0.15**	[0.05, 0.29]	0.45	[0.18, 0.68]	.20	
Head Injuries Severity	-0.09	[-0.24, 0.06]	-0.19	[-0.48, 0.13]	.04	
						R^2 = .30* BCa CI [.03, .57]
(Intercept)	-1.52*	[-3.70, -0.46]				
Lack of Insight	0.22**	[0.08, 0.42]	0.47	[0.07, 0.65]	.22	
Positive Affect Rec.	0.14*	[0.04, 0.32]	0.43	[0.13, 0.64]	.18	
Alcohol Use	0.35	[-0.09, 0.92]	0.21	[-0.22, 0.39]	.04	
						R^2 = .31* BCa CI [.04, .51]
(Intercept)	-1.27**	[-3.59, -0.78]				
Lack of Insight	0.20**	[0.07, 0.41]	0.47	[0.17, 0.69]	.22	
Positive Affect Rec.	0.16**	[0.06, 0.32]	0.49	[0.25, 0.71]	.24	
Substance Use	0.44*	[0.03, 0.83]	0.33	[0.02, 0.63]	.11	
						R^2 = .37** BCa CI [.02, .54]

Note. *B*= Unstandardised Beta, *B* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta; *β*= Standardised Beta, *β* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; 95% CI for R^2 = BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared. LOC= loss of consciousness; Positive Affect Rec.= positive affect recognition.
* indicates p < .05.

** indicates p < .01.

Lack of insight and positive affect recognition were then examined using a hierarchical regression to investigate whether they explain the variance in violence severity over and above the HCR-20. The model with the HCR-20, lack of insight and positive affect recognition was significant, though, lack of insight and positive affect recognition were the only two explanatory variables which were significant (F (3, 28) = 3.57, p = 0.03). An ANOVA was performed to obtain the F-change statistic, which was also significant (F (2, 28) = 2.78, p=.02), and the proportion of variance explained by each model was compared and indicated that the addition of lack of insight and positive affect recognition to the HCR-20 explained an additional 22% of the variance in violence severity (Table 8.15). Adjusted R^2 increased from .03 to .20 suggesting that lack of insight and positive affect recognition improved the model. The model violated assumptions of normality and homoscedasticity; however, this was corrected after bootstrapping was employed. When negative affect recognition was analysed with lack of insight, the model was not significant, and lack of insight was the only significant explanatory variable even following bootstrapping (B= .066 (BCa 95% CI= -.03, .17), β= .26; F (1, 29) = 2.72, p= .08).

Table 8.15 Exploratory Regression for Prospective Violence Severity in Inpatients

Variable	*B*	BCa 95% CI	*β*	BCa 95% CI	sr^2	Fit	Difference
(Intercept)	-0.31	[-1.39, 0.63]					
HCR-20	0.02	[-0.01, 0.07]	0.24	[-0.16, 0.51]	.06		
						R^2 = .06 95% CI [.00, .23]	
(Intercept)	-1.51*	[-3.46, -0.13]					
HCR-20	0.01	[-0.02, 0.04]	0.10	[-0.25, 0.35]	.01		
Lack of Insight	0.16*	[0.03, 0.37]	0.35	[0.04, 0.57]	.12		
Positive Affect Rec.	0.13*	[0.04, 0.27]	0.40	[0.09, 0.60]	.16		
						R^2 = .28* BCa CI [.03, .52]	ΔR^2 = .22*

Note. *B*= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta; β= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared. Positive Affect Rec.= Positive affect recognition.
* indicates p < .05.
** indicates p < .01.

To further explore the finding that *better* positive affect recognition increases violence severity, a post-hoc analysis was performed to examine whether *better* positive affect recognition also increased the odds of an individual behaving aggressively or violently. As only n = 2 inpatients were physically violent during prospective follow-up, and the violence severity score includes aggressive and violent incidents, this score was transformed into a binary variable where 1= violent/aggressive, and 0= non-violent, and resulted in there being nine incidents. A logistic regression was performed using positive affect recognition and lack of insight entered simultaneously, resulting in a significant model (χ^2 (2) = 7.19, p= .03, *pseudo* R^2= .19), however, following bootstrapping, only the intercept was significant. Subsequently, when affect recognition was examined in the same way, the model was not significant (χ^2 (2) = 5.14, p= .08, *pseudo* R^2= .14). These findings indicate that while *better* affect recognition, and *better* positive affect recognition explained severity of violence and anger provoked incidents, they did not increase the odds of a person behaving angrily or violently.

8.2.3 Exploratory Analysis: All Cognitive Variables- Inpatients

Finally, a multiple linear regression was performed to examine the cumulative relationship between all cognitive variables and violence severity. The model was not significant (F (8, 23) = 1.36, p= .26), which was likely due to insufficient power, however, bootstrapped confidence intervals did not indicate that together, these variables would significantly explain the variance in the outcome in a larger population (Table 8.16). Further, adjusted R^2= .08, indicating that if this model were derived from the population, it would explain 24% less variance in violence severity than it did in this sample, indicating that the model lacks generalizability.

Operational Definitions:

- **Response Monitoring:** Represents the total number of perseverative errors on the M-WCST. On this measure, participants were asked to sort cards with different colours and shapes on them according to certain rules, however, the examiner cannot tell the participant the rules, they must use the examiners voice (e.g., correct, or incorrect) to figure out what the rules are for the current round. Perseverative errors were recorded when the participant was told a rule was incorrect, but continued to try the same rule (e.g., a lack of response monitoring).
- **Risk Taking:** This is operationalised by the IGT Net Total score, which represents the proportion of disadvantageous cards chosen in a gambling task. In the IGT, a computer task, participants were presented with four piles of cards, and a sum of fake money. Each card either adds to the money or removes some, and it is the participant's job to choose

cards where they will either win the most money or lose the least money. Choosing cards from the disadvantageous pile results in winning more money, but also losing more money.

- **Reasoning**: Reasoning is operationalised by the perceptual reasoning scale on the WASI.
- **Cognitive Empathy:** Cognitive empathy is operationalised by participants correctly identifying what an actor is *feeling* in vignettes on the TASIT. The score is made up of the total correct.
- The remaining variables were previously operationalised.

Table 8.16 Exploratory Regression for Violence Severity for Inpatients

Variable	*B*	BCa 95% CI	*β*	BCa 95% CI	sr^2	Fit
(Intercept)	0.79	[-5.46, 5.32]				
Inattention	-0.01	[-0.04, 0.04]	-0.12	[-0.55, 0.29]	.01	
Response Inhibition	-0.01	[-0.07, 0.02]	-0.09	[-0.38, 0.13]	.01	
Response Monitoring	-0.01	[-0.04, 0.02]	-0.10	[-0.48, 0.17]	.01	
Risk Taking	0.01	[-0.01, 0.03]	0.09	[-0.13, 0.41]	.01	
Reasoning	-0.02	[-0.05, 0.01]	-0.23	[-0.52, 0.01]	.05	
Affect Recognition	0.05	[-0.04, 0.16]	0.25	[-0.12, 0.67]	.06	
Cognitive Empathy	0.05	[-0.14, 0.15]	0.13	[-0.24, 0.43]	.02	
Lack of Insight	0.12	[-0.09, 0.35]	0.26	[-0.18, 0.49]	.07	
						R^2 = .32 BCa CI [.08, .43]

Note. *B*= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta; *β*= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

Taken together, in the inpatient sample, better affect recognition and poor insight explained a significant proportion of the variance in violence severity even after controlling for number of head injuries with a LOC, number of head injuries, severity of head injuries, positive symptoms of psychosis, and alcohol and substance use. Further, the addition of measures of affect recognition and insight explained the variance in violence severity over and above the HCR-20 and explained an additional 18% of the variance, demonstrating their ability to add incremental validity to the HCR-20. Furthermore, better *positive* affect recognition and lack of insight explained the variance in violence severity over and above the HCR-20, and explained an additional 22%, indicating that better positive affect recognition may be a better explanatory variable than affect recognition in general.

8.2.4 Exploratory Hypothesis 1. The addition of measures of inattention and response inhibition to the LSI-R:SV will add incremental validity to the LSI-R:SV for explaining antisocial behaviour (Community Sample).

As the LSI-R:SV is a general offending risk/needs assessment, it was hypothesised that inattention and response inhibition would add incremental validity to the LSI-R:SV, for explaining a non-violent outcome, such as antisocial behaviour.

Operational Definitions:

- **Antisocial Behaviour:** This was operationalised as the number charges or convictions of non-violent/non-aggressive offences which occurred during the follow-up period in the community sample, over 6 months. However, participants who demonstrated antisocial behaviour only had one incident each, making the outcome naturally binary. Presence of antisocial behaviour was coded as 1, and absence of the behaviour was coded as 0. Behaviour included, misuse of drugs, theft, attempting lockfast with intent, shoplifting, and breach of peace. As there was only a subset of participants who had a completed LSI-R:SV (*n*=18), and *n*= 1 participant in that subset was missing prospective follow-up data, leaving only *n*=3 prospective antisocial incidents. However, retrospective antisocial behaviour had *n*= 5 incidents, thus it was decided that the retrospective outcome would serve as a better dependent variable to examine this hypothesis.

Analysis and Results. A hierarchical logistic regression was performed to examine this hypothesis following the same procedures as previously stated. A likelihood ratio test was then used to examine whether addition of the cognitive variables to the model significantly improved the fit. The first model was approaching significance (i.e., $p < 0.05$) (χ^2 (1) = 3.58, p= .06), however, the second model with inattention and response inhibition was not significant (χ^2 (3) = 4.92, p = 0.18). The likelihood ratio test demonstrated that the addition of the cognitive measures did not significantly improve the model (χ^2 (2) = 1.34, p= .51), though the AUC score for the model is .80, indicating accurate differentiation of antisocial and non-antisocial individuals. The model correctly identified 100% of non-antisocial participants, but only correctly classified 20% antisocial participants. Results are presented in Table 8.17.

Table 8.17 Exploratory Regression for Retrospective Antisocial Behaviour (n=18)

Predictor	*Est.*	BCa 95% CI	*OR*	BCa 95% CI	Fit
(Intercept)	-3.39*	[-7.26, -0.87]			
LSI-R:SV	0.62*	[0.03, 1.68]	1.85	[1.03, 5.34]	
					Pseudo R^2 = .17
(Intercept)	-2.06	[-319.83, 439.25]			
LSI-R:SV	0.71	[-15.96, 6.97]	2.01	[0.00, 106.87]	
Inattention	0.04	[-19.23, 2.53]	1.04	[0.00, 12.66]	
Response Inhibition	-0.08	[-19.41, 17.68]	0.93	[0.00, 4.77e+07]	
					Pseudo R^2 = .25

Note. 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for the model estimate; *OR*= Odds Ratio, *OR* 95% CI= BCa bootstrapped confidence intervals for the odds ratio, *Pseudo-*R^2= McFadden's Pseudo-R^2
* indicates $p < .05$.
** indicates $p < .01$.

8.2.5 Exploratory Hypothesis 2. Measures of inattention and response monitoring will be associated with retrospective breaches (Community).

Operational Definitions:

- **Breaches of License/Conditions:** Number of breaches of license/conditions retrospectively over 6 months. Breaches included missed appointments, not attending unpaid work, and missed court ordered curfews. There was a low base rate of participants with prospective incidents (n = 1), however, retrospectively, there were n = 6 participants with incidents, therefore, the retrospective outcome was used as the dependent variable in this analysis. As it was naturally binary, presence of breaches was coded as 1, and absence of breaches was coded as 0.

Rationale for Hypothesis. Whilst it is acknowledged that the operational definition of breaches does not include violence or anger provoked incidents, it was hypothesised that participants who miss appointments and important, court ordered obligations may have deficient executive functioning. Moreover, these types of failures contribute to a cycle of imprisonment and further offending. In Criminal Justice Social Work, when an individual does not turn up for their unpaid work shift, for example, they get three warnings from an unpaid work officer, and after the third warning, they are reported to the Sheriff. The Sheriff then decides on whether the individual's sentence should be longer or re-visited, or if they need to appear in court. Although antisocial behaviour may contribute to this, the notion of getting three warnings before an action is taken, made the researcher posit whether the individual has deficient executive functions which would contribute to the inability to plan, successfully achieve goals, and to understand the long-term consequences of their actions. Response monitoring is an executive function that allows individuals to observe, evaluate, and change their behaviours when a mistake is made. Response monitoring is closely related to response inhibition, and when impaired, contributes to deficient self-monitoring of actions and behaviours and inhibitory control. Behavioural products of deficient inhibitory control include, impulsivity, risk taking behaviours, and inattention. Inattention may contribute to an individual forgetting an appointment, and impulsivity and risk taking may contribute to the individual acting without considering consequences, or risking getting a lengthier sentence. Moreover, the cognitive and behavioural products of deficient inhibitory control may contribute to an individual's inability to maintain organisation and structure, resulting in missed appointments.

Analysis and Results. A logistic regression was preformed, and all variables were entered simultaneously (Table 8.18). The model was not significant (χ^2 (2) = 1.00, *p*= .61), and did not provide a good fit for the data, as the goodness of fit test was significant (χ^2 (3) = 0.49, *p*= .01). Response inhibition was then entered into the model, and inattention was removed which resulted in a model which fit the data better (χ^2 (3) = 2.42, *p*= .49), however the model and the variables were not significant (χ^2 (2) = 0.97, *p*= .62). The findings did not support this hypothesis.

Table 8.18 Exploratory Regression for Retrospective Breaches- with 1000 Bootstrap Samples (*n*=31)

Predictor	*Est.*	BCa 95% CI	*OR*	BCa 95% CI	Fit
(Intercept)	-3.12	[-13.47, 5.09]			
Inattention	0.03	[-0.03, 0.11]	1.03	[-0.97, 1.11]	
Response Monitoring	0.00	[-0.11, 0.17]	1.00	[0.89, 1.18]	*Pseudo R^2* = .03
(Intercept)	2.23	[-12.73, 15.29]			
Response Inhibition	-0.08	[-0.32, 0.07]	0.93	[0.72, 1.06]	
Response Monitoring	0.00	[-0.12, 0.15]	1.00	[0.88, 1.16]	
					Pseudo R^2 = .03

Note. 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for the model estimate; *OR*= Odds Ratio, *OR* 95% CI= BCa bootstrapped confidence intervals for the odds ratio, *Pseudo-R^2*= McFadden's Pseudo-R^2

8.2.6 Exploratory Analyses: Best Subsets Regressions (Community)

All included variables have been previously operationalised.

Analyses. Additional exploratory analyses were conducted to examine the relationship between the untested variables and outcomes. Like the inpatient sample, best subsets regressions were performed to find the best independent variables to explain the outcomes. Outcomes of interest were examined one at a time in the following order, prospective violence severity, retrospective violence severity, and retrospective antisocial behaviour. After the best subsets regression identified potentially important variables, they were entered simultaneously in a multiple linear regression if the dependent variable was continuous (e.g., violence severity), and a multiple logistic regression if they were binary (antisocial behaviour). For variables which were significant, their robustness was further evaluated by adding covariates to the model one at a time and examining them separately. Following exploration through best subsets regressions, a final multiple linear regression was performed with all cognitive variables with violence severity (retrospective) as the dependent variable, to examine their cumulative relationships.

Results: Best Subsets Regressions

1. Inattention and affect recognition were identified as the best explanatory variables for prospective violence severity, however, the model and independent variables were not significant (F (2, 26) = 1.74, p= .19, R^2= .12).
2. Risk taking and cognitive empathy were identified as important variables for retrospective violence severity, and the model and variables were not significant even after bootstrapping (F (2, 28) = 1.27, p= .30, R^2= .08).
3. Response monitoring and risk taking were identified as the best explanatory variables for retrospective antisocial behaviour. The model was significant (χ^2 (2) = 7.21, p = .03, *Pseudo* R^2= .20), and response monitoring was significant following bootstrapping (*OR*= 1.12, 95% BCa CI [1.02, 1.29]), however, risk taking was not (*OR*= 1.11, 95% BCa CI [0.84, 1.96]). Response monitoring only, was then entered into a logistic regression, and resulted in a significant model (χ^2 (1) = 4.92, p = .03, *Pseudo* R^2= .14), and response monitoring was significant following bootstrapping (*OR*= 1.12, 95% BCa CI [1.02, 1.33]), though in the opposite direction than expected. The model correctly classified 100% of those who were not antisocial during the follow-up period but failed to correctly classify any participants who were antisocial. A hierarchical logistic regression was performed to determine whether the addition of response monitoring to the LSI-R:SV would improve the accuracy of the model, however, the likelihood ratio test indicated this was not the case (χ^2 (1) = 0.52, p = .47). Albeit response monitoring may be an important variable to examine further in antisocial behaviour.
4. For retrospective violent incidents, risk taking, and cognitive empathy were identified as important variables. The results of the logistic regression demonstrated that there was no significant association between these variables and the outcome even after bootstrapping (F (2, 28) = 0.89, p = .42).

5. Finally, the model with all cognitive variables in it was not significant (*F* (7, 21) = 0.69, *p* = .68), and no variables were significant after bootstrapping. The model explained 19% of the variance in the outcome, and semi-partial squared correlations show that inattention, response monitoring, risk taking, and affect recognition each individually explained the most variance (Table 8.19). Adjusted R^2= -.073 indicating that the model does not explain the data well and is not generalizable to a larger population.

Table 8.19 Prospective Violence Severity in the Community Sample (*n*=29)

Variable	*B*	*B* 95% CI	*β*	*β* 95% CI	sr^2	Fit
(Intercept)	2.93	[-0.38, 15.56]				
Reasoning	0.00	[-0.02, 0.07]	0.04	[-0.32, 0.38]	.00	
Inattention	-0.01	[-0.06, 0.00]	-0.20	[-0.42, 0.34]	.04	
Response Inhibition	-0.00	[-0.9, 0.03]	-0.02	[-0.40, 0.37]	.00	
Response Monitoring	-0.02	[-0.09, 0.00]	-0.21	[-0.40, 0.34]	.04	
Risk Taking	0.02	[-0.02, 0.11]	0.13	[-0.57, 0.19]	.02	
Affect Recognition	-0.09	[-0.44, 0.00]	-0.19	[-0.45, 0.37]	.04	
Cognitive Empathy	0.02	[-0.44, 0.25]	-0.03	[-0.39, 0.51]	.00	
						R^2 = .19 95% CI [.11, .22]

Note. *B*= Unstandardised Beta, *B* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1000 replicates); *β*= Standardised Beta, *β* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; 95% CI for R^2 = BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

8.2.7 Exploratory Analysis with the Whole Sample (*n*= 63)

The final analysis combined both samples to investigate explanations for the variance in self-reported impulsivity. An a priori power analysis revealed that to find a medium effect size, six predictors could be examined in one model. The variables for the regression were chosen prior to analyses (Janes et al., 2021), and include inattention, response inhibition, response monitoring, risk taking, reasoning, cognitive empathy, and affect recognition, all of which were entered simultaneously. A dummy variable was also used to control for group (Inpatients= 1; Community= 0). First, zero order correlations (Appendix V) were examined and revealed no problems with collinearity. The regression was performed, but was not significant (*F* (8, 54) = 1.85, *p* = 0.09; R^2= .22), however, following bootstrapping cognitive empathy was significant explanatory variable for impulsivity, albeit in the opposite direction than expected (Table 8.20). All model assumptions were met, though bootstrapping was employed to obtain the confidence intervals for a larger sample. Adjusted R^2= .09, indicating that in a population the model would explain 13% less variance in self-reported impulsivity than this sample did.

Table 8.20 Results for Self-Reported Impulsivity (*n*=63)

Variable	*B*	*BCa* 95% CI	*β*	*BCa* 95% CI	sr^2	Fit
(Intercept)	54.95	[17.15, 91.57]				
Reasoning	-0.03	[-0.19, 0.15]	-0.04	[-0.28, 0.18]	.00	
Inattention	0.14	[-0.02, 0.38]	0.19	[-0.02, 0.47]	.04	
Response Inhibition	0.02	[-0.39, 0.52]	0.01	[-0.27, 0.30]	.00	
Response Monitoring	-0.17	[-0.43, 0.05]	-0.16	[-0.37, 0.05]	.03	
Risk Taking	-0.13	[-0.46, 0.21]	-0.11	[-0.38, 0.19]	.01	
Affect Recognition	0.38	[-0.15, 1.29]	0.12	[-0.04, 0.33]	.01	
Cognitive Empathy	1.24*	[0.06, 2.55]	0.21	[0.03, 0.46]	.04	
Group	-3.04	[-8.88, 2.30]	-0.14	[-3.63, 0.10]	.02	
						R^2 = .22 BCa CI [.08, .29]

Note. *B*= Unstandardised Beta, *BCa* 95% CI= bias corrected and accelerated (BCa) bootstrapped confidence intervals for Unstandardised Beta (1000 replicates); *β*= Standardised Beta, *BCa* 95% CI= BCa bootstrapped confidence intervals for standardised Beta; BCa CI for R^2 = 95% BCa bootstrapped confidence intervals, sr^2= semi-partial correlation squared.

9 Assessing the Cognitive Contributors to Violence: Part IV

Chapter Preface

This brief chapter draws together the findings presented in chapters 7 and 8, and discusses the similarities and differences observed in similar research. First outlining the aims, followed by a discussion of the feasibility results, cognitive profiles of violent offenders, primary and secondary hypotheses, and implications for clinicians and researchers. The chapter concludes with recommendations for clinicians, and suggestions for researchers in this field.

9.1 Study Aims

This study was undertaken to assess the feasibility of conducting a larger, future study, and to pilot a battery of measures of neuropsychological abilities identified through meta-analyses and a Delphi study. A prospective observational design was implemented with two distinct groups of violent offenders to investigate the incremental and predictive validity of neuropsychological measures independent of important covariates and existing risk assessments. Feasibility and pilot studies are often necessary for researchers to determine the optimal processes, resources needed, how the study and data can be best managed, and to ensure it is scientifically valid (van Teijlingen & Hundley, 2001).

Individual studies and reviews examining the link between cognitive impairments and violence in forensic populations are increasing, however, the heterogeneity of effect sizes and methodology, small sample sizes, and poor research quality, have contributed to a pool of somewhat inconsistent findings, as observed in the findings of the review completed for this thesis (Chapter 4). In the review, the heterogeneity between studies made it difficult to compare and synthesise findings, lending to the notion that there is a need for researchers to step back, examine the neuropsychological measures which have sufficient performance indicators for predicting violence, and to identify procedures which will maximise the data quality and the findings. Further, while replication with human participants is often difficult, there is a need for replication in this field, rather than innovation, thus, it was posited that a pilot and feasibility study would be a necessary starting point to establish a solid foundation to build upon.

The study was designed to examine two different groups of violent offenders to gain a better understanding of the neuropsychological profiles of forensic psychiatric inpatients, and offenders living in the community. Robust methodology was implemented after seriously acknowledging the common limitations in this field, and measures were chosen based on validity, reliability, and their utility in clinical practice to ensure that there would be a clear transition from research to practice with the aim of optimising the understanding of neurocognition in violent populations, treatment programmes, and formulations of risk. Feasibility outcomes related to recruitment, participant retention, tolerability of measures, and predictive performance of measures. These outcomes would directly impact a decision to progress to a larger study, and if so, offer insights into what minor amendments to the study design and methodology may be required in a larger study. Primary outcomes related to examining the relationship between neuropsychological measures and violent outcomes. Secondary outcomes related to anger, antisocial behaviour, and proxies for violent behaviour, which were implemented as a fail-safe in the case of a low base rate of prospective recorded incidents, which is common in observational studies of violence. Whilst the feasibility outcomes related directly to the decision of whether to progress to a larger study, the primary and secondary outcomes were expected to facilitate a broad assessment of the potential effects of neurocognition on violent behaviour, as well as give indicators of whether a larger study would be necessary.

9.2 Study Findings

There are a number of findings from this study which have clinical and research implications, both with respect to the aims of this thesis and also with respect to future research in this area. Several of these findings and their implications are reviewed below, and others are discussed in the final chapter of this thesis.

Progression to a Full Study. A study protocol was published before data analyses commenced, which outlined a set of clearly specified criteria which a decision would be based on to progress to a larger study. The pre-specified criteria are presented in Table 9.1. Most of the feasibility criteria were met in both samples, except not all measures demonstrated sufficient predictive accuracy, sensitivity, specificity, or incremental validity. ROC analyses demonstrated that only reasoning (AUC= .83), inattention (AUC= .76), response inhibition (AUC= .86), and affect recognition (AUC= .88) had sufficient predictive performance indicators for prospective violence in the inpatient sample, although the base rate of incidents was low. When the same analysis was performed for retrospective violence, which had a higher base rate of incidents, only reasoning (AUC= .79), cognitive empathy (AUC= .74), and lack of insight (AUC= .78) demonstrated sufficient performance indicators. In the community sample, only inattention (AUC= .81) demonstrated sufficient predictive performance indicators for prospective violence, which also had a low base rate, and no measures demonstrated predictive validity for retrospective violence. AUC values for incremental validity were calculated for both samples using prospective and retrospective violence, in the inpatient sample, reasoning, inattention, response inhibition, and affect recognition were sufficient for prospective violence, though, only lack of insight remained significant independent of the HCR-20 for retrospective violence. In the community sample, inattention remained significant independent of the LSI-R:SV, though all AUC analyses should be interpreted cautiously as they were underpowered. All the measures were tolerable to the samples, and the completion rate was excellent, however, new neuropsychological measures may need to be identified for a future study as many were not sufficiently sensitive and specific to violence, particularly in the community sample. Though, on the contrary, it may be that the measures which demonstrated sufficient performance are measures of the most important risk factors for violence, and their use should be replicated in a larger study. The findings of the feasibility portion of the study demonstrate that the methods and procedures are feasible for a future, larger study with only minor changes.

Table 9. 1 Assessment of feasibility criteria for both samples

Study Outcome	Criterion	Observed	Criterion Met
Tolerability of measures	> 3.5	≥ 4	Yes
Recruitment Rate	> 30%	~ 66%	Yes
Completion Rate of Measures	≥ 70%	100%	Yes
Sensitivity & Specificity of Measures for violence	≥ 1.50	Range 1.18- 1.83	Partially
Predictive Accuracy	≥ .70	Range .52- .88	Partially
Incremental Validity	≥ .70 after controlling for risk assessments	Range .51- .71	Partially

Recruitment Procedures. Recruitment in both samples was by far the greatest challenge of this study. First, the limitations set by NHS ethics made the process of recruitment particularly difficult for both the researcher and clinicians. As a result of this, it was necessary for the researcher to rely heavily on clinicians who already had little time to spare. Understandably this was done in the best interest of the patients, however, it may be useful for research committees within the forensic facilities to liaise with NHS research ethics committees on what may and may not be appropriate for this population in relation to recruitment, as well as ways to maximize the autonomy of the researcher, with only little action necessary from practicing clinicians, within reason. As both samples were included in the study protocol, similar procedures were followed when recruiting the community sample, that is, a large amount of work was expected from the social workers before the researcher was able to meet a service user. Much like professionals in the NHS, social workers have limited time outside of their daily workload. The consequence of this was that they were unable to go through each individual on their case load and record reasons for exclusions or invite every service user who fit the study criteria, to participate. Thus, it is recommended that a future study implement a recruitment plan for the community offenders that would allow for the recording of reasons why they were excluded, and to give every individual who fits the criteria the opportunity to participate. This implementation would need to be discussed with the social work teams to ensure that the proposed plan would also be feasible for them to record this information. Further, due to difficulty with engaging community participants, a monetary component was added to the study to reimburse them for their time. This proved to be successful and is strongly recommended for consideration in future studies with this population.

In addition to the restrictions implemented by ethics, recruiting both samples proved to be challenging in other ways as well. Much of the inpatient sample suffered from paranoia, making it difficult for them to trust the researcher, and likewise, the community sample appeared to be aversive to working with professionals in general. However, the original protocol set out to recruit $n = 20$ participants from high, medium, and low security hospitals for a total of $n = 60$ and $n = 20$ from the community sample for a total sample size of $n = 80$. Following recruitment from the Forensic Network, $n = 32$ patients were successfully recruited, which prompted an amendment to recruit $n = 32$ community offenders, for a total of $n = 63$ participants, indicating that 79% of the recruitment target was reached, which is a promising finding. Though, it is acknowledged that the recruitment rate for all inpatients was only 42% out of those who were eligible, and the true rate of recruitment for the community sample could not be calculated, which are both limitations of this study. It is recommended that researchers planning to conduct studies with these difficult populations have strong clinical skills to optimize engagement, and most importantly, to build a rapport and trust to ensure validity and reliability of research findings.

Neuropsychological Battery. The main concern about the study protocol from all research committees was the length of the testing battery. However, the battery took approximately one hour and 40 minutes to complete, and the tolerability ratings demonstrated that neither sample found the measures intolerable. For the inpatient sample, participants were given a minimum of two sessions to complete the measures, and community participants were given the option to break the assessments into two or more sessions. Frequent breaks were offered to all participants which were often not needed. Moreover, it should be noted that one $n = 1$ participant in the entire study withdrew consent after beginning assessments.

Capacity and Autonomy of the Samples. With acknowledgement that both samples are considered vulnerable research participants, it appears that there is a slight disconnect between what research ethics committees and hospital board research committees believe about the capacity of this population to take part in research and what their actual capacity is. For example, although all participants were assessed to have the capacity to make their own decisions, additional concerns from these committees regarding assessment times, as well as information collected from participant files, gave the impression that these samples would not be able to make the right decision for themselves, even after they were considered to have the capacity to consent. It is acknowledged that there is a thin line between the need to protect a vulnerable sample, and to allow them full autonomy to make decisions. However, when ethical integrity is held to the utmost priority in research, the information sheet and consent form should give the participant enough information to make an informed decision based on what they feel is best for them, as long as they have the capacity to do so.

Primary and Secondary Outcomes. Descriptive analyses provided an important look at the performance of violent offenders on a variety of neuropsychological measures, including, reasoning, response inhibition, inattention, response monitoring, risk taking, affect recognition, cognitive empathy, and lack of insight. However, whilst it was hypothesised that the inpatient sample would perform significantly more poorly relative to the community sample, this was true only for cognitive empathy and simple sarcasm, and following Bonferroni correction, no other measures significantly differentiated the two samples. Though, after transforming measure scores to z-scores, of the core measures, the inpatient sample had higher proportions of impairment on reasoning, response inhibition, response monitoring, risk taking, affect recognition, and cognitive empathy.

Despite significant theoretical support for cognitive impairments being important for the prediction of violence risk, most of the measures did not demonstrate a statistical association with violence, anger provoked incidents, antisocial behaviour, or proxies of violence in these samples. Primary analyses of the inpatient sample showed that affect recognition, specifically, better positive affect recognition, and lack of insight explained a significant proportion of the variance in prospective violence severity independent of number and severity of head injuries, head injuries with LOC, alcohol use, and substance use, though substance use explained an additional proportion of the variance to both models. Further, together, positive affect recognition and lack of insight explained an additional 22% of the variance, over and above the HCR-20, demonstrating significant incremental validity. Better affect recognition (e.g., recognition of both positive and negative emotions) and lack of insight also demonstrated significant incremental validity after controlling for the HCR-20, explaining an additional 18% of the variance in violence severity. A post-hoc exploratory analysis using logistic regression indicated that although both better affect recognition and better positive affect recognition, and lack of insight explain the variance in violence severity, they do not significantly increase the odds of the inpatients behaving violently or aggressively. The only significant finding from the secondary analyses was that poor response inhibition explained a significant proportion of the variance in the mean number of days a patient was on elevated levels (*perceived risk)*. In the community sample, only *better* response monitoring increased the odds of an individual being antisocial (*OR*= 1.12), though, this was no longer significant in the presence of the LSI-R:SV, and inattention was a significant explanatory variable for self-reported impulsivity (R^2= .11). Aside from prospective and retrospective anger, and violence frequency, all of which had too low a base rate to analyse, no variables were significantly associated with the remaining outcomes.

In summary, findings from the inpatient sample indicated that measures of reasoning, inattention, response inhibition, and affect recognition had sufficient predictive performance indicators for prospective violence, and reasoning, cognitive empathy, and lack of insight for retrospective violence. Measures that demonstrated incremental validity for prospective violence were reasoning, inattention, response inhibition, and affect recognition, and for retrospective violence, only lack of insight was sufficient. Better affect recognition and lack of insight explained 22% of the variance in prospective violence severity, and 18% when controlling for the HCR-20; and better positive affect recognition and insight explained 27% of the variance in prospective violence severity, and 22% when controlling for the HCR-20, both independent of most important covariates. Poor response inhibition may be an important explanatory variable for perceived risk. In the community sample, only inattention had sufficient predictive performance and demonstrated incremental validity, though in the opposite direction than expected, for prospective violence. Better response inhibition increased the odds of antisocial behaviour, and inattention individually explained 4% of the variance in self-reported impulsivity. All cognitive variables together explained 32% of the variance in the inpatient sample, and 19% in the community sample for prospective violence severity, though adjusted R^2 indicated that neither model is generalizable to a larger population. The final analysis combined the samples to investigate the association with self-reported impulsivity, though the model was not significant, it explained 22% of the variance, and cognitive empathy emerged as a significant variable, however, it was in the opposite direction than expected.

Neuropsychological Profiles of Violent Offenders. Whilst there were only two measures which significantly differentiated the two samples, the proportion of participants who performed in the borderline or impaired range on the neuropsychological measures is noteworthy for several reasons (Table 7.8). For example, for the neuropsychological measures of interest, 33% of the entire sample of *n*= 63 performed in the borderline or impaired range on reasoning, 45% on inattention, 2% on response inhibition, 14% on risk taking, 46% on affect recognition, and 38% on cognitive empathy. In a normal population that follows a normal bell curve, only 2.5% of the population would be expected to perform in the borderline or impaired range, however, in this sample of violent offenders, up to 46% performed in this range on the measures of interest. These findings alone have substantial clinical implications and should be of interest to clinicians and social workers working with justice involved individuals regardless of whether they demonstrate predictive validity for violence risk. Further, when the proportions of impaired individuals were compared to the group means, it starkly demonstrated that group means can be alarmingly misleading. For example, the observation that many of the group means indicated subclinical impairments, but had high proportions of individual impairment, alluded that the association between cognitive impairment and violence risk may only be relevant on an individual level rather than at the group level. Overall, although the inpatient sample performed more poorly than the community sample, the non-significant differences between groups on most of the measures suggests that violent offenders with and without psychosis share similar neuropsychological profiles, and thus, share many of the same cognitive strengths and weaknesses. From these findings, it is clear that these populations may have trouble with perceptual reasoning, sustaining attention, and recognizing emotions and social cues, all of which should be considered by clinicians and social workers when developing treatment programmes and risk formulations.

Positive Affect Recognition and Lack of Insight May Predict Violence Severity in Inpatients. Although affect recognition was in the opposite direction than hypothesised, *n*= 29 (46%) of the entire sample scored in the borderline to impaired range, however of the *n*= 9 participants who behaved violently or aggressively during prospective follow-up, only *n*= 4 had scores in the impaired range, indicating that in this sample, most individuals who were violent had low average affect recognition abilities. Moreover, when this was further examined to explore whether better affect recognition in this sample increased the odds of being violent or aggressive, the results were not significant suggesting that better affect recognition did not increase the odds of acting violently or aggressively.

Theoretically, poor social cognition is thought to be linked to poor response inhibition and impulsivity, where the misidentification of social cues and poor response inhibition may lead to poor self-regulation and impulsivity and the use of violence. For example, research on forensic psychiatric samples has suggested that intact social cognition, namely, theory of mind (ToM), is a prerequisite for regulation and adjustments, including affect regulation and impulse control (J. Allen et al., 2008; Baumeister & Heatherton, 1996; Weiss et al., 2006), as well as empathy (Blair, 2005). Moreover, previous research has suggested that a better understanding of the emotional states of others may function as an inhibitor of aggression (Dolan & Fullam, 2006), and thus, a lack of understanding may act as a risk factor for violence. Research on individuals who are diagnosed with schizophrenia have confirmed that social cognition in schizophrenia is often compromised (Abu-Akel & Abushua'leh, 2004; Ang & Pridmore, 2009; Bora et al., 2009; Brüne, 2005; Frith, 2004; Harrington et al., 2005; Sprong et al., 2007), and that poor social cognition may contribute to violence (O'Reilly et al., 2015). However, some studies have found results which contradict these findings, for example, in 2004, Abu-Akel and Abushua'leh examined the link between violence and mentalizing abilities in individuals with a diagnosis of schizophrenia and found that affective ToM (the ability to make empathic inferences), *decreased* the likelihood of violent propensities, and cognitive ToM (the comprehension of cognitive mental states), *increased* the likelihood of violent propensities. Further, in a study comparing forensic and non-forensic patients with schizophrenia, Majorek et al, (2009) found that the two groups not only did not differ in performance on a mentalising task, but the forensic patients significantly outperformed the non-forensic group, also finding a link between better mentalising and increased violence. Although this link appears illogical, Abu-Akel and Abushua'leh (2004) attributed this, in addition to deficits in empathy, to an ability which may be used in violent offenders to manipulate and deceive their victims. However, this postulation would also indicate the pre-planning of violent incidents, or instrumental violence, which was not observed in the current sample, though this may partially explain the positive correlation between better affect recognition and more severe violence prospectively. Though, as affect recognition was inversely correlated with retrospective violence, this finding is likely to be an artefact of the small sample size and low power.

The finding of lack of insight as a significant explanatory variable for violence has been demonstrated thoroughly in the literature, and lack of insight is a risk factor for violence on the HCR-20 v3 (Douglas et al., 2014). Two separate meta-analyses which examined risk factors for violence in psychosis (e.g., Witt et al., 2013) and schizophrenia (e.g., Reinharth et al., 2014), both found that lack of insight was associated with violence. However, the current study also demonstrated how the use of the lack of insight and judgement scale from the PANSS can be used as an individual measure to not only predict future violence, but that it has the potential to add incremental validity to the HCR-20. In consideration of these findings, further research is necessary to better understand the link between better social cognition and increased violence severity in populations with schizophrenia and other psychotic illnesses.

Better Response Monitoring May Increase the Odds of Antisociality in Community Offenders. In conjunction with reasoning, response monitoring is necessary for the successful functioning of response inhibition, a crucial executive function. Thus, theoretically, better response monitoring would indicate better response inhibition, less impulsive behaviour, and less criminality. The finding that *better* response monitoring increases the odds of antisocial behaviour is in stark contrast to the hypotheses in this study, as well as the extant literature. With minimal empirical support for this finding, there are additional possible explanations. For instance, this may be an artifact of the study limitations namely a low base rate of incidents, which can contribute to skewed findings. Alternatively, this finding may be a product of the use of MWCST, as when the proportion of participants with impairments on each measure was examined (Table 7.8), it became apparent that this measure picked up on very few impairments in the community sample, relative to the inpatient sample, suggesting that it may have decreased sensitivity to less significant impairments, or as a result of it being modified, it may have been less challenging for the community sample. Finally, response monitoring was no longer significant after controlling for the LSI-R:SV, indicating that in in this study it did not demonstrate the ability to contribute incremental variance, and did not support the central thesis hypothesis.

Null Findings. In the current study, null findings were considered as important, if not more important, than significant findings, considering the current state of the literature. Despite significant theoretical support, many of the measures selected to assess neuropsychological abilities in the current study did not demonstrate predictive validity for primary or secondary outcomes. Whilst it may be tempting to fault the small sample sizes, a priori power analyses were conducted to ensure there would be 80% power when examining the primary and secondary hypotheses, and bootstrapping was employed to obtain a 95% confidence interval to determine if the models would be significant at a population level. Further, effect sizes ranged from R^2= .02 to .37, where the latter was rare, indicating primarily small to moderate effects. Literature on effect sizes suggests that if there are no effects to be found, increasing the study power will not increase the magnitude of effect sizes (Ellis, 2010). This may be an indicator that a future study may not be valuable, namely, for those variables that did not reach significance. Moreover, although there were bootstrapped confidence intervals that did not cross the null line, they remained wide indicating uncertainty in the true estimates and effect sizes, even in a population of n= 1,000.

Another possible explanation for null findings is the low base rate of outcomes. Whilst this is a limitation of the current study, the number of secondary outcomes available for analysis likely offset this. It is acknowledged that retrospective outcomes are not as robust as prospective outcomes, and proxies for violence are not as robust as actual violence, however, as this was a pilot study, secondary outcomes were implemented to allow for analysis to identify indicators that a future, larger study was warranted. Each outcome was robustly examined, and the results did not differ considerably for any of the outcomes regardless of the base rate of incidents.

Null findings are much less likely to be published in peer-reviewed journals contributing to publication bias, which may explain why, despite empirical support for the link between cognitive impairments and violence risk, the majority of the findings of this study, in two separate violent offending samples, were null. Thus, it may be that the *null hypothesis* that measures of neuropsychological abilities do not significantly relate to an individual's risk of violence, is supported by the non-significant findings in this study. Reports of non-significant relationships between cognitive measures and violent and antisocial outcomes have been reported previously (see meta-analyses in Chapter 4), however, the literature argues against this (e.g., Hancock et al., 2010; Morgan & Lilienfeld, 2000; Ogilvie et al., 2011; Reinharth et al., 2014), suggesting that the findings from the current study may relate more strongly to the neuropsychological measures used, which was also postulated in the meta-analysis. However, as seen in the meta-analysis in the current thesis, various measures were used in a variety of samples, and the majority of the findings remained in the subclinical range. Yet, there is the possibility that the relationship between scores on cognitive measures and the outcomes analysed in the current study were moderated by participant characteristics and the environment in which they resided during the study. Although, formal moderation analyses using interactions were not conducted in the current study, the addition of substance misuse to the model with lack of insight and affect recognition increased the strength of their relationship to violence severity, suggesting that substance misuse on top of poor cognitive functioning may increase severity of violence in the inpatient sample, which is in line with research discussed in Chapter 2 of this thesis. Additional participant characteristics such as increased number of or severity of TBI, increased symptoms of psychosis, alcohol use, and adverse childhood experiences are also likely to have a moderated effect on cognitive functioning by increasing cognitive impairments, perhaps strengthening the relationship between cognitive functioning and violence risk, though this needs to be explored in a future study. Alternatively, it could be assumed that treatments such as cognitive remediation, certain cognitive therapies, and psychotropic medications would increase cognitive functioning and weaken the relationship between cognitive measures and violence, which, aside from cognitive remediation which did not take place on the participants during this study, may be an additional explanation for null findings. Further, although the inpatient and community samples did not differ significantly on many of the measures, the inpatient sample demonstrated poorer performance on almost all tasks, further indicating that psychosis and symptoms of mental illness may moderate this association.

9.3 Clinical Implications

The findings from this study have substantial implications for clinicians and researchers. Arguably, the most important finding for clinicians was the proportion of impairments observed in each sample, and overall, on the neuropsychological measures. Neuropsychological assessments of forensic populations should be routinely completed to understand their cognitive strengths and weaknesses. The findings from this study indicate that a large proportion of the sample may be functioning at a similar level as individuals with intellectual disabilities (ID), however, individuals with a diagnosis of ID were excluded from the current study. The DSM-5 defines intellectual disability as a neurodevelopmental disorder that begins in childhood and is characterised by intellectual difficulties, including, "reasoning, problem solving, planning, abstract thinking, judgement, academic learning, and learning from experience", which is confirmed through standardized IQ testing (American Psychiatric Association, 2013, p. 33). A formal diagnosis also includes deficient adaptive functioning which inhibits individuals from conforming to developmental and sociocultural standards, in addition to their responsibility to meet social responsibilities (American Psychiatric Association, 2013). The severity of ID follows a spectrum from mild to profound, though most individuals are classified as having a mild ID, which is characterised by having slower conceptual development and daily living skills, however, these individuals have the ability to learn practical skills allowing them to function with minimal levels of support. Further, the definition of ID can include any individual whose level of intelligence inhibits their ability to function adaptively and pro-socially (Walsh et al., 2020), and in this sample, 44% had FSIQ scores in the borderline and impaired ranges, and 33% had reasoning scores, and 49% had verbal comprehension scores all in the borderline and impaired ranges. Furthermore, a low IQ may indicate that an individual has a low level of social and interpersonal maturity, indicating that they may require different treatment modalities (Walsh et al., 2020). In general, it is not speculated that the samples included in this thesis suffer from profound ID, and equally, many may not meet all required criteria a formal diagnosis, however, a large proportion of this sample has demonstrated performances on neuropsychological measures at a deficient level of cognitive functioning (See Table 7.8). An uneven pattern of cognitive impairments will still avail individuals the capacity to commit crimes but may not allow them the capacity to lead a structured, pro-social lifestyle, or the ability to inhibit risk factors for violence.

Bearing this in mind, it is possible that these individuals may not be receiving the same support that an individual with a formal ID diagnosis receives, which may inhibit their overall ability to benefit from treatments. For example, cognitive-behavioural therapy (CBT), a widely accepted and utilized treatment for aggressive behaviours (Lee & DiGiuseppe, 2018), was once deemed inappropriate for use in individuals with intellectual disabilities, as the verbal skills and accurate reporting of feelings for which cognitive therapies rely on, are often deficient in individuals with ID (e.g., Sturmey, 2004). Moreover, past research has identified the specific skills required to fully benefit from cognitive therapies, such as CBT, namely, the ability to provide self-reports and to use abstract concepts (Biza Stenfert Kroese, 2005), communication skills, intact self-monitoring and memory, and the ability to recognize emotions, as well as the cognitive model of the therapy (Hatton, 2002). More generally, an essential requirement for any talk therapy is sufficient verbal skills and verbal comprehension (e.g., Wright, 2006) a skill in which 49% of this sample were borderline or impaired. This is not to say that all violent offenders will require additional support in therapies such as CBT or similar, as research has found that most individuals with mild ID have the necessary skills to take part in these treatments (Dagnan et al., 2000). However, literature surrounding the use of cognitive therapies suggest that intellectual functioning, specific cognitive and skill abilities and deficits, and levels of motivation and confidence be thoroughly and carefully assessed to formulate the modifications of the intervention that are necessary to ensure they reflect the individual's needs and learning style (Lynch, 2004; Willner, 2005). Whilst it is not known if the participants in this study would meet the full criteria for a diagnosis of ID, it is clear that a large proportion are functioning at a borderline or impaired level, suggesting that proportions of this population require extra support, even on daily and less complex tasks, and to ignore this in the presence of treatment planning and risk management would be imprudent. Thus, assessment of cognitive strengths and weaknesses should always be considered when treating violent offenders, and the findings should be implemented into treatment plans to optimise treatment, and to decrease recidivism.

9.4 Research Implications

These findings may also have implications for researchers. For example, most of the literature in this area of research points to significant results, however, based on the findings from this study and the meta-analysis conducted for this thesis, it appears that the research findings may partially be due to publication bias, and the file drawer problem, that is, negative and null results are less likely to be published. Although a number of studies in this area suggest that cognitive abilities contribute to violence and have the ability to add incremental validity to risk assessments, the effect sizes in this study and the meta-analyses were largely subclinical, and thus, may not be useful in the way of prediction. However, they likely have utility in treating offenders, and may be useful for predicting functional outcomes, which may be more valuable for this population. Based on these findings, it is recommended that research in this area begins to pull away from using neuropsychological measures to predict violence, and instead uses them as they were originally intended, that is, to identify the cognitive strengths and weaknesses of this population to aid in treatment. There are additional research implications relating to the cognitive impairments observed in the sample, which are discussed in the final chapter of this thesis.

9.5 Study Strengths

Although the study findings did not fully support the overall hypothesis that neuropsychological measures are significantly associated with violence, the study itself was successful. The overall hypothesis was acquired from the literature, and the study was carefully designed to ensure robustness of findings including specific operational definitions of primary and secondary outcomes, and clearly operationalised independent variables and covariates. Two different samples of violent offenders were successfully recruited and comprehensively assessed using valid and reliable measures. Further, secondary outcomes were implemented as a fail-safe to avoid the loss of potentially important analyses due to low base rates and they represented both inpatient and community violence. Aside from the insufficient performance indicators for several of the measures, the study procedures proved to be feasible in both settings. Finally, the study protocol was submitted for publication prior to data analysis, with pre-specified, objective criteria to determine progression to a larger study, and appropriate analyses were undertaken, including bootstrapping of confidence intervals to get an indication for the necessity of a future study in a larger population.

9.6 Limitations

Although the theoretical concepts included in the current study were assumed to be valid, to obtain significant results, these concepts would need to be validly measured. Thus, there is a chance that the null findings relate more to the measures used in the study, as several of them did not demonstrate adequate predictive performance indicators for violent outcomes. Moreover, it could be the case that the measures do not assess the concepts they claim to. However, much attention was given to selecting measures with robust psychometric properties, suggesting the null findings may be a product of the sensitivity of measures, rather than the validity.

Measurement error may have also contributed to the null findings. In the current study, potential sources of measurement error include invalid administration of measures and the response style of participants. Measurement error can result from invalid administration, scoring, or data input, which can introduce artificial variance and lead to invalid interpretations of findings. However, the researcher had previous intensive training on the administration, scoring, and interpretation of neuropsychological assessments. Further, efforts were made to ensure that the testing of individuals was done in quiet, distraction-free rooms and participants were regularly given the option to take breaks. However, it is possible that measurement error still occurred, as the researcher was the sole administrator, scorer, and interpreter of all measures, and reliability was not assessed. Measurement error can also be associated with the response styles of participants, such as their approach to the assessments, namely, motivation. Although theoretically this was controlled for, as informed consent required participants to engage on a voluntary basis rather than a mandatory basis, a lack of motivation was observed in a small proportion of the sample, though this was reliant on observation only. A lack of motivation may have also contributed to the proportion of impairments seen in these samples, though, the attitude most often observed in the participants was a motivation to perform well.

The small sample size may have also contributed to null results, and therefore, impacted the statistical validity of the findings. However, for hypothesis testing, a power analysis was conducted to ensure there was adequate power to run the statistical analyses, though ROC analysis was performed on all measures to obtain their performance indicators, which is an analysis that is influenced by sample size (Singh, 2013), and findings should be interpreted cautiously. Whilst it may have been beneficial to examine the cumulative effects of all independent variables for all hypotheses, the sample size was unlikely to negatively affect the statistical validity of the primary and secondary findings. Though, statistical validity can also be impacted by the distributions of administered measures, and in the current study, several of the measures did not follow a normal distribution. This was assessed before any analyses commenced, and was accounted for by using non-parametric bootstrapping, a method that makes no assumptions about the distribution, similar to ROC analyses, which are also non-parametric. Nonetheless, model estimates, and effect sizes should be interpreted using the BCa confidence intervals for interpretation, as they are likely to contain the true values.

In both samples, only reactive violence occurred, thus, the analysis of reactive and instrumental violence could not be analysed.

Finally, the results may not be generalizable. A limitation often attached to risk assessments is that they perform best in a specific population, with a specifically defined behaviour, and often, a specific setting (e.g., Douglas, Cox, et al., 1999). However, although the generalizability of the results is restricted to men who have been convicted of violent offences, the findings represent inpatient violence, and community violence. Another common limitation in this type of research is that inpatient violence does not generalise out with an institutional setting as patients' behaviour is closely monitored and proxies are often used as outcomes. Though, the other side of the argument suggests that the measurement of community violence may not be complete as incidents of violence are only recorded when the individual is caught.

Whilst both arguments are valid, the combination of including both samples likely contributed to a balance in the current study. Further, while there is no way to know how many violent incidents went unrecorded in the community sample, their participation in criminal justice social work included more supervision than would be given to an individual in the community who does not receive any services. Overall, the outcomes, methodology, and samples included in the current study are generalizable to the current research, and the findings presented here can directly inform the violence risk literature.

9.7 Recommendations

It is recommended that routine neuropsychological screening of forensic populations be implemented into practice, and the results implemented into their treatment plans, and possibly formulations of risk. Understandably, due to the time and resources that neuropsychological assessments require, this may not be possible in all settings, though, if possible, a referral for a neuropsychological examination should be made a priority. As professionals working with vulnerable populations, it is a requirement to ensure that individuals fully understand their situation and conditions.

For researchers, it is recommended that reasoning, inattention, response inhibition, affect recognition, and response monitoring be examined in future studies, as they all emerged as possible important explanatory variables for risk. However, with acknowledgement to the inconsistencies in effect sizes in conjunction with their small magnitude, it is opined that most cognitive measures will not add significant value to risk assessments, thus, it is recommended that the focus shift from using cognitive functioning to predict violence, and instead focus on predicting functional outcomes and ways to optimise treatment gains. It is posited that a focus on treatment gains, rather than future violence, will serve a similar purpose; that is, if an individual is fully benefiting from treatment, they may be less likely to behave violently in the future.

9.8 Conclusions

This feasibility and pilot study was undertaken to test the feasibility of implementing neuropsychological testing into forensic settings, and to pilot a battery of measures to examine their predictive performance for violent outcomes. The findings suggest that a future, larger study is feasible, but small changes would need to be implemented first. However, a future study may only be necessary for some cognitive abilities, as there was little evidence to suggest that most neuropsychological measures were strongly associated with violence risk. ROC analyses demonstrated that measures of reasoning, inattention, response inhibition, affect recognition, and lack of insight in the inpatient sample maintained a large AUC score after controlling for the

HCR-20, however, the base rate was low, which likely inflated the results. Similarly, the community sample also had a low base rate and only the measure of inattention demonstrated incremental validity when controlling for the LSI-R:SV. Overall, lack of insight with better affect recognition in the inpatient sample were the only measures that evidenced their ability to add incremental validity to existing risk assessments in the regression analyses. Based on these findings, it is concluded that while predicting future violence is undoubtedly important, there is little evidence to suggest that the addition of neuropsychological measures will add significant incremental validity to existing risk assessments. As a result, it is recommended that forensic populations, undergo comprehensive neuropsychological assessments to optimize treatment outcomes and to better understand their cognitive strengths of weaknesses.

10 Cognitive Contributors to the Variance in Behaviours Associated with Imminent Situational Aggression: A Pilot Case-Series

Chapter Preface

Chapters 4 through 9 presented a cohesive set of studies examining the relationship between neuropsychological measures and violent and aggressive outcomes over a 6-month follow-up period. The current chapter presents a pilot case-series that utilised the same neuropsychological battery to examine the relationship between the cognitive measures and variance in behaviours that are associated with imminent situational aggression, in a small sample from the State Hospital, a high secure setting. In addition to investigating this relationship, this study was designed with a focus on how the major limitations seen in violence literature may be addressed, such as small sample sizes and low base rates of observational outcomes.

10.1 Introduction

Inpatient violence presents a major concern for psychiatric hospitals and institutions. Twenty-five to 35% of inpatients exhibit violent behaviour during hospitalization, putting both other patients and staff at risk (Arango et al., 1999; Daffern et al., 2003; K. Fisher, 2016). Although this has been evidenced, historically, the focus of violence risk assessments has remained on the offenders' risk of long-term, future violence, namely within the community, and less on their short-term risk of violence within an institutional or hospital setting. More recent research has brought to light the importance of short-term risk assessments, introducing tools such as the Short-Term Assessment of Risk and Treatability (START; Webster et al., 2006), the Dynamic Appraisal of Situational Aggression for Inpatient Violence (DASA-IV) (Ogloff & Daffern, 2006), and the Brøset Violence Checklist (Woods & Almvik, 2002). Researchers have suggested that assessing the risk of violence in the short term is imperative for assessing, managing, and preventing violence in inpatient and institutional settings (Michael Doyle & Logan, 2012).

Whilst many well-known risk assessments encompass static and dynamic risk factors for violence, short-term risk factors are often clinical rather than historical, as they are more likely to vary from day to day (McNiel et al., 2003). Aggressive behaviour can arise from a range of dynamic or clinical factors, including recent violence or threat of violence, physical discomfort or pain, sleep issues, positive symptoms of psychosis, irritability, and lack of therapeutic alliance

(K. Fisher, 2016). In addition to these factors, behavioural and cognitive products of cognitive impairments can also contribute to short term risk; namely, impulsivity, found on the DASA-IV (Ogloff & Daffern, 2006) and lack of insight found on the START (Nicholls, Brink, Desmarais, Webster, & Martin, 2006), however, there is reason to believe that additional cognitive factors may also contribute to short-term risk, and may play a role in the variance in a variety of behaviours which may lead to aggression. For example, as observed in this thesis, forensic populations are characterised by poor cognitive functioning, particularly executive dysfunction, which appear to be more perceptible in the presence of risk factors for cognitive impairment such as TBI, certain mental illness diagnoses, and a history of harmful alcohol and substance misuse, also common in this population. Deficient cognitive abilities are linked to problems with self-regulation and socially inappropriate behaviours, which may manifest as impulsive, disinhibited, and aggressive behaviour (Ogilvie et al., 2011). Further, deficient executive functioning can manifest as reduced self-control and attention control, carelessness, poor planning abilities, goal setting and cognitive flexibility, impulsivity, and increased emotional lability (Jackson et al., 2014; Morgan & Lilienfeld, 2000; Ogilvie et al., 2011; Paschall & Fishbein, 2002). Of particular interest in the current study are poor self-control, emotional lability, impulsive and disinhibited behaviours, and poor self-regulation as these highlight behaviours which may be less predictable, reactive, and more variable from day to day. Most studies which have included the DASA-IV have focused on its psychometric and predictive performance (e.g., Griffith et al., 2013; Nqwaku et al., 2018; Vojt, 2014), however, the current study took a different approach by using the DASA-IV as an outcome measure to capture the variance in patient behaviour over a short period. As the DASA-IV items are risk factors for imminent situational aggression, they are considered behaviours which may precede or be indicative of aggression, rather than being aggressive acts themselves. Thus, the primary focus of this study was not to predict short-term risk *per se* but was instead to determine whether performance on neuropsychological measures can explain the variance in behaviours which are associated with aggression over a 14-day period. It was anticipated that the identification of risk factors which may explain the variance in this behaviour may help to highlight important patient characteristics to assist nursing staff in identifying patients who may be more likely to have variable presentations of behaviour, which may aid in the management and prevention of inpatient violence and aggression.

In addition to assessing cognitive explanations for variable behaviour, the current study was designed with the aim to overcome common methodological limitations of observational research on violent outcomes (see Chapter 3). It is evident that many studies which examine violence in forensic populations suffer from limitations of small sample sizes, poorly operationalised outcomes, and low base rates of observational outcomes. Whilst the recruitment of a large sample in conjunction with the guarantee of a robust base rate of incidents often seems improbable in risk assessment research, it was postulated that the use of a repeated measures design would not only theoretically increase the power in this study to allow for more advanced statistical analyses but would also increase the likelihood of capturing more incidents, increasing the base rate. Moreover, repeated measures designs are highly advantageous over cross-sectional designs, as they allow for the evaluation of with-in participant change across time and increase power while reducing the cost of conducting the research (Guo et al., 2013).

10.1.1 Aims

- To investigate whether performance on neuropsychological measures and scores on clinical measures explain the variance in behaviours indicative of imminent situational aggression in forensic psychiatric inpatients over a short period of time.
- To determine whether the use of a repeated measures design can help overcome methodological limitations of power which result from a small sample size, and a low base rate of incidents.

10.2 Method

10.2.1 Methodological Considerations and Study Design

Chapter 3 of this thesis presented evidence to highlight the major methodological limitations in risk assessment research with recommendations on how these can be acknowledged in future studies. Many of these limitations were considered in the design of the pilot and feasibility study (see Methodological Considerations in Chapter 6), however, some of the weaknesses persisted, namely, small sample size and low-base rates. As emphasized in Chapter 3, a small sample contributes to low power, and low power can decrease the chance of finding the true effect, and may also inflate the magnitude of the effect, both contributing to unreliable findings (Button et al., 2013). Likewise, low base rates of incidents make the accurate prediction of future behaviour more difficult and may tempt researchers to broaden their operationalisations of violence to include additional behaviours or proxies, essentially making the target behaviour larger, but the specificity of the prediction lower. While it is often difficult to recruit large samples in this type of research, it was posited that implementing a repeated measures design would increase the statistical power in a small sample. Repeated measures allow for the examination of within and between subject variance by separating the error variance and lowering the overall random error, increasing power. To acknowledge the base rate limitation, it was theorized that because the observation of actual, physical violence is rare, especially in inpatient settings where behaviours are well managed, perhaps measuring indicators for imminent aggression, rather than violence, may increase the number of recorded incidents. Finally, low base rates of incidents often lead researchers to not only broaden their operationalisation of the outcome, but in cases where a continuous outcome was used and only a small proportion of the sample had a different score than the majority of the sample, the outcome is often treated as binary. For example, this was seen in pilot and feasibility study (Chapters 6-9) where many of the community participants only had one incident during a follow-up period, but a few had more than one. When a continuous outcome is then treated as binary, valuable information may be lost, and interpretations of the data are more inhibited. To address this limitation, a novel statistical analysis using two-part mixed models which analyses the binary and the continuous portions of the outcome in one model was piloted in the current study to examine its utility in situations such as this, and to determine whether this analysis would be of value for future research in this area. Thus, the current study was designed as a repeated measures pilot case-series, which created multilevel data where observations were clustered around patients which allowed for the examination of both within and between patient variance on the outcome, while controlling for individual characteristics. It was hypothesised that this design would help to overcome the limitations in the following ways:

Small Sample Size. In repeated measures design, the repeated observations increase statistical power, thus even with a small sample size, intricate statistical analyses can still be performed.

Low-Base Rate. It was posited that rating patients' behaviour three times a day for 14-days would decrease the chances of patient incidents being missed or overlooked by the State Hospital's formal reporting procedures and would increase the variance of the outcome. Likewise, the outcome measure used to rate behaviours in the current study was chosen because the items on it do not measure *violence*, but instead measure behaviours associated with aggression, which are more likely to be observed on a day-to-day basis. Finally, it was hypothesised that using two-part mixed models would increase our ability to salvage continuous data where there is low variance, due to a low base rate.

10.2.2 Plan of Investigation

The study was conducted at the State Hospital, a high secure hospital, in Carstairs, Scotland (see Chapter 6 for a more in-depth description of the hospital). The proper use of the DASA-IV requires that patient behaviour is rated three times a day to indicate whether there is a risk for inpatient aggression within the following 24 hours (Ogloff & Daffern, 2006). Therefore, the repeated measures design for the current study required nurses to evaluate and rate patient behaviour using the DASA-IV, three times a day for 14-days, and when no rateable behaviour was evident, they were instructed to record a zero.

Pre-Data Collection Approvals. A proposal to amend the main study (presented in chapters 6-9) to include this additional data collection was sent to the State Hospital Research Committee on 31st of May 2018 and received approval on 4th of June 2018. A substantial amendment was then applied for through IRAS, and final approval was granted on 21 September 2018.

Recruitment of Hospital staff. Due to the nature of the data collection, the researcher was required to heavily rely on nursing staff to complete daily ratings of patient behaviours. The researcher liaised with a charge nurse who agreed to assist with engaging ward nurses, and subsequently identified key nurse contacts from each ward who agreed to ensure the ratings were completed each day.

Identification of Participants. Only patients from the State Hospital who participated in the main study (presented in chapters 6-9) were invited to participate. Before patients were approached, their responsible medical officer was contacted to confirm capacity to consent. After confirmation was received, patients were approached, given an information sheet, and were given seven days to decide whether they would like to participate, at which point if they agreed, consent was obtained. Patients were asked to consent to nurses rating their behaviour three times a day for a period of 14 days, and for the researcher to access this data in their electronic record.

10.2.3 Data Collection

Behaviour Ratings. Nursing staff were asked to rate the behaviour of the identified participants three times a day (backshift, morning shift, night shift) for 14 days in a row using the Dynamic Appraisal of Situational Aggression- Inpatient Version (DASA-IV; Ogloff & Daffern, 2006). Each day, the researcher visited the ward in the morning and the afternoon to ensure the ratings were being completed, and reminders were sent out frequently by email.

Cognitive Measures. The cognitive measures comprised the test battery that had already been administered for the purpose of the main study (presented in chapters 6-9). Scores for each measure were used as continuous independent variables. Baseline testing took place between January 2018 and September 2018, and data collection for the current study commenced in October 2018. Cognitive measures and operational definitions are presented in Table 10.1.

Table 10. 1 Operationalisation of Independent Variables (IV)

IV	Measure	Operational Definition
Reasoning	WASI-II	Perceptual Reasoning Index (PRI)[a]
Attention	CPT-3	Number of omission errors[b]
Response Inhibition	Stroop Colour & Word Test	Interference score[a]
Response Monitoring	Modified WCST	Number of perseverative errors[a]
Risk Taking	Iowa Gambling Task	Proportion of cards chosen from disadvantageous card piles[a]
Cognitive Empathy	TASIT	Number of correctly identified feelings[a]
Affect Recognition	TASIT	Number of correctly identified emotions[a]
Lack of Insight	PANSS	Lack of Insight scale[b]

Note. WASI= Wechsler Abbreviated Intelligence Scale, CPT= Conners' Continuous Performance Test; WCST= Wisconsin Card sorting test; TAST= The Awareness of Social Inference Test; PANSS= Positive and Negative Syndrome Scale; [a] lower scores represent higher impairment; [b] higher scores represent higher impairment.

Clinical Measures. Measures of clinical characteristics, including history of alcohol and substance misuse were recorded from available data in patient medical files, and history of severe head injury were coded based on self-report using the OSU-TBI-ID. Clinical measures and operational definitions are presented in Table 10.2.

Table 10. 2 Description of Clinical Risk Factors

Clinical Factor	Measure	Variable Coding
Alcohol Use	Presence of a secondary diagnosis of alcohol dependence or harmful (recreational) use in patient file at baseline.	Presence of diagnosis= 1; Absence of diagnosis= 0
Substance Misuse	Presence of a secondary diagnosis of harmful (recreational) substance use, polysubstance/multiple drug misuse, or drug dependence in patient file at baseline.	Presence of diagnosis= 1; Absence of diagnosis= 0
History of Head injuries	Worst head injury self-reported was rated for level of severity using the OSU-TBI.	0= No head injury; 1= Improbable TBI; 2= Mild TBI; 3= Definite mild TBI; 4= Moderate TBI; 5= Severe TBI

Note. OSU-TBI-ID= Ohio State University Traumatic Brain Injury Identification Method

10.2.4 Outcome Measure

Behaviours Associated with Aggression. Behaviours associated with aggression was operationalised by the DASA-IV (Ogloff & Daffern, 2006). The DASA-IV consists of seven scales: (1) Irritability, (2) impulsivity, (3) unwillingness to follow instructions, (4) sensitive to perceived provocation, (5) easily angered when requests are denied, (6) negative attitudes, and (7) verbal threats. The items are scored 0 if they are absent or 1 if they are present now or have been present in the last 24 hours, and then the scores are summed to obtain a composite risk score. Although there is no cut-off score, other research has used a cut off score of 4 to identify high-risk patients (Kaunomaki, 2015). The ability of the DASA-IV to predict imminent risk for aggression has been compared to the clinical scale of the HCR-20, resulting in good predictive validity, where the DASA-IV demonstrated an AUC of .82 and the HCR-20 clinical scale demonstrated an AUC of .73 (Ogloff and Daffern, 2006). The rationale for choosing the DASA-IV as an outcome measure in the current study was briefly discussed in section 10.2.1, stating that in an attempt to avoid the common limitation of low base rates in observational studies, the outcome measure used to rate behaviours in the current study was chosen because the items in it do not measure *violence* (which is rare and often results in a low base rate), but instead measure behaviours associated with aggression. Although it can be used to 'take the temperature' of a ward, it also has predictive utility for imminent aggression. Moreover, the risk factors included in the DASA-IV, which were used as the outcome in the current study, are closely related to behaviours often observed in patients with cognitive impairments.

10.2.5 Statistical Analysis

Descriptive statistics were reported for all participants using means and standard deviations.

To investigate variance in DASA-IV ratings over time, multilevel modelling (MLM) was used. MLM is increasingly used to analyse single case-series designs (e.g., Shadish et al., 2013), and is useful for the analysis of small-N case-series as it treats each individual as a 'cluster' and allows for the testing of a statistical model based on individual characteristics and a statistical examination of changes in slopes and intercepts over time. Thus, by using MLM, researchers can investigate whether individual characteristics at baseline, results in a change in slope or intercept of the outcome measure. Analysing data overtime in the same individuals often leads to the time points being correlated which violates the independence assumption of ordinary least squares (OLS) regression, thus MLM was deemed a more appropriate analysis for the current study.

Linear MLM were planned for the primary analysis of the current study, however, following data collection it became apparent that the outcome data had an excess of zero ratings, with 13 non-zero ratings (e.g., presence of behaviour) and 701 ratings of zero (absence of behaviour). As a result of the large number of zero ratings, the data were right skewed with the majority of ratings clumping at zero, thus the data were considered to be semi-continuous, that is, data that has a mixture of true zeros and continuously distributed positive values (Farewell et al., 2017). As some of the ratings were ones, the outcome could be transformed into a binary outcome, and logistic mixed models could be used, however, changing a continuous outcome to a binary outcome inhibits the inference that can be made from the model parameters, including making inferences about the intensity, severity, or level of the covariate in relation to the outcome, and can also artificially inflate the variance. Thus, the use of a zero-inflated two-part mixed model was utilised. Zero-inflated two-part mixed models are equipped to handle repeated measures data with a large number of zeros. These models work by separating the zero responses from the non-zero responses, allowing for the examination of the severity or intensity of the outcome by using a linear mixed model with a log transformed outcome analyzing only the non-zero responses, and the odds of the occurrence of the outcome by using a logistic mixed model, by coding all responses as one or zero and treating them as binary, in one analysis. The lognormal distribution is used to describe the probability distribution of the nonzero values, and the binomial distribution is used to describe the binary portion of the model. The outcome in the lognormal model can be interpreted as the typical behaviour rating per time of rating, given that a rating of aggression has occurred, whereas the binary portion can be interpreted as whether or not a behaviour occurred (e.g., Tooze et al., 2002; Walls et al., 2009).

In these models, the random effects are not assumed to be independent, thus the likelihood components cannot be maximized separately. The likelihood is maximized using quasi-Newton optimization of a likelihood approximated by adaptive Guassian quadrature (Tooze et al., 2002). Further, the components from each model share a common parameter, ρ, which reflect possible covariation between the two parts of the model, and marginal effects can be calculated when fixed effects are used in both parts of the model. Interpretation of fixed effect coefficients for these models follow the same interpretation as lognormal regressions or logistic regressions, and standard regression diagnostics, such as Q-Q plots for residuals, to check goodness of fit can be used. The random effects account for the unobserved variability among the patients. In the lognormal part of the model, the random intercept allows some patients to consistently have a high or low mean of nonzero values, and in the logistic part of the model, the random intercept, on the logit scale, allows some patients to have a consistently high or low probability of a nonzero response. The variability of the random effect in the logistic part of the model indicates the variability in the probability of a nonzero response among patients with similar covariate patterns, whereas the random effects in the lognormal part of the model indicates the amount of variability of mean nonzero responses among the patients with similar covariate patterns (Tooze et al., 2002). It is important to note that the linear model reflects only those patients who received a nonzero rating during the follow-up period and is not generalizable across all patients. As this is an analysis that has limited application in violence research, a breakdown of the key equations for the model are not repeated here but can be found in Walls et al. (2009).

Missing data were replaced with a zero after participant notes were cross-checked to ensure there was not a missed incident that should have been rated. Data were entered in long format, where each patient had one row per observation. All cognitive measures were entered into the models as continuous covariates, and behaviour ratings were treated as continuous. Models included an interaction of fixed effects (time and cognitive variables), and a random intercept (patients), resulting in *mixed* effects models. All continuous covariates were centered to have a mean of zero and a standard deviation of one to allow for proper convergence of the models, as well as ease of interpretation of interactions and the intercepts of the fixed effects. Mixed models separate the overall variability of the repeated measures outcome into a fixed and random component. The fixed component measures individual characteristics of the sample and moderators of the outcome overtime (e.g., interaction effects), where the random component estimates the variability between the subjects. All models were fitted using the Gauss-Hermite quadrature by means of the Generalized Linear Mixed Models using Adaptive Gaussian Quadrature (*GLMMAdaptive*) in R (Rizopoulos, 2021). Deviance information criteria was used to assess the unaccounted-for variance, and Alkaike's Information Criteria (AIC) was used to compare models with fixed effects to the intercept only model, where smaller values represent a better model fit.

In the analyses, several models were tested to assess for potential mechanisms of variance over time. First a link function using the family function, *hurdle.lognormal* from the *GLMMAdaptive* package was created to separate data out for each model. For the logistic part, a binary indicator was used ($y > 0$) to allow non-zero responses to be coded as one, and for the lognormal part, $y > 0$ allowed for the identification of all nonzero responses, which were then log-transformed to allow for a reasonably normal distribution. Model 0 (which subsequent models would be built on) evaluated the variance between patients and whether a multi-level model was warranted. Model 1 evaluated whether time explained the variance in the outcome. Cognitive variables, interacting with time, were then added to the model to assess the proportion of variance accounted for over time. Due to the limited sample size, the predictors were analysed individually, for example, model 2 (Reasoning), model 3 (inattention), model 4 (behavioural impulsivity), model 5 (response inhibition), until all predictors were entered, which meant that the effects were not cumulative, and results should be interpreted with caution. Coefficients from the lognormal part of the model can be exponentiated, subtracted from 1 and multiplied by 100 (e.g., exp(coefficient)-1*100) to get the percentage increase or decrease in the outcome for every one unit of change in the predictor, and the coefficients from the logistic part are exponentiated to transform into odds ratios, where a one unit change in the predictor increases or decreases the odds of the patient receiving an aggression rating or not.

10.3 Results

10.3.1 Sample Characteristics

Of the *n*=20 patients who consented to participate in the original pilot and feasibility study, *n*=17 consented to participate in the current study, *n*=2 had been transferred out of the hospital before they could be approached, and *n*=1 declined to participate.

The mean age of the sample was 43.17 (SD = 10.33) years, ranging from 20-60 years old. The primary nationality of the sample was White Scottish (*n*= 12; 71%) and the remaining individuals were White British (*n*= 5; 29%). Seventy-six percent (*n*= 13) of the sample left school with no qualifications, and the remaining 24% (*n*= 4) held an undergraduate degree or had some undergraduate credits.

Forty-seven percent of the sample (*n*= 8) had a primary diagnosis of schizophrenia, 35% (*n*= 6) had a primary diagnosis of schizoaffective disorder, the remaining three participants had primary diagnoses of delusional disorder, bipolar affective disorder with psychosis, and alcohol induced late onset psychotic disorder. The mean length of admission in the State Hospital at baseline was approximately 2.71 year (SD = 3.33) and 5.80 months (SD= 3.17). Most patients were considered restricted patients (88%; *n*= 15), and of the *n*=17 individuals, 47% (*n*= 8) were held at the hospital on a Compulsion Order, 53% (*n*= 9) were on a treatment order. Forty-one percent of the sample (*n*= 7) had an index offence of murder, 18% (*n*= 3) attempted murder, 18% (*n*= 3) culpable homicide, and the remaining 23% (*n*= 4) had index offences of assault and assault to severe injury.

Fifty-three percent of the participants (*n*= 9) self-reported a history of head injury, 6% (*n*= 1) had a secondary diagnosis of a drug dependence disorder, and 18% (*n*= 3) had a secondary diagnosis of an alcohol dependence disorder. Additionally, 12% (*n*= 2) of the sample had a secondary diagnosis of polysubstance misuse, 12% (*n*= 2) had a secondary diagnosis of mental and behaviour disorder due to multiple drug use, and 12% (*n*= 2) had secondary diagnoses of harmful use of drugs and alcohol or harmful use of multiple substances. Last, 35% (*n*= 6) had an ‘other’ diagnosis involving alcohol and substance misuse, namely, mental and behavioural disorder due to harmful alcohol use (12%; *n*= 2), polysubstance misuse (6%; *n*= 1), use of multiple illicit substances (6%; *n*= 1), mental and behaviour disorder due to harmful drug use (6%; *n*= 1), and harmful polysubstance use (6%; *n*= 1).

10.3.2 Behaviour Ratings

Each patient had 42 time points, equalling 714 observations for the entire sample; this resulted in 13 ratings greater than zero and 701 ratings of zero (e.g., absence of behaviour). All 13 ratings were accounted for by only *n*= 4 patients (168 observed time points), referred to from here as the *behaviour group*. Three of the four patients in the behaviour group were white Scottish, and one was white British, three had no education qualifications, and one had an undergraduate degree, all were restricted patients, and all had a schizophrenia spectrum diagnosis, where two had a diagnosis of schizoaffective disorder, one of paranoid schizophrenia, and one of schizophrenia. Descriptive statistics for baseline measures are presented in Table 10.3. Notedly, the behaviour group had poorer performances relative to the no-behaviour group on measures of reasoning, response inhibition, risk taking, and response monitoring, and had a higher lack of insight (Figure 10.1 presents box plots for the mean scores in each group). Table 10.4 provides a frequency table for categorical variables.

Table 10.3 Descriptive Statistics for Baseline Measures

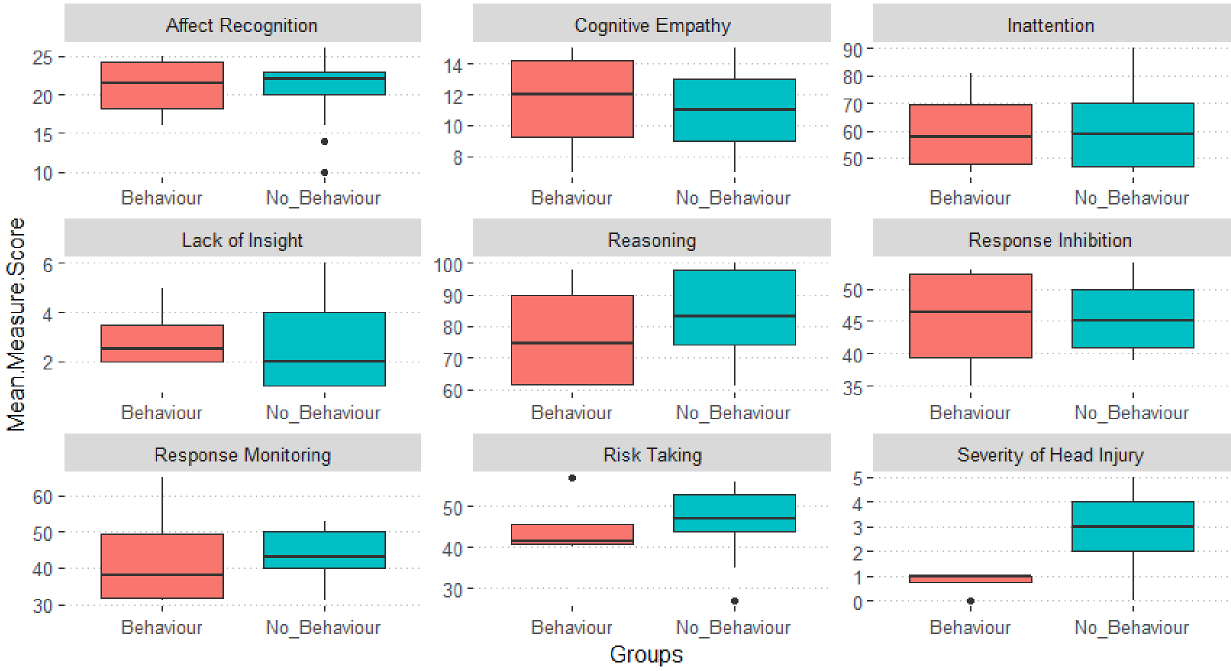
	Behaviour (*n*= 4)		No-Behaviour (*n*= 13)	
Variable	Mean	SD	Mean	SD
Reasoning	77.00	18.50	83.30	13.20
Inattention	60.20	57.50	62.30	17.00
Response Inhibition	45.20	46.50	45.60	5.14
Risk Taking	45.00	41.50	46.20	8.18
Response Monitoring	43.00	38.00	44.00	7.26
Lack of Insight	3.00	2.50	2.54	1.85
Cognitive Empathy	11.50	12.00	10.80	2.51
Affect Recognition	21.00	21.50	20.80	4.72
Severity of Head Injury	0.75	0.50	2.92^	1.38

Note. ^*n*= 12; Normative mean (SD) for reasoning= 100 (50); inattention, response inhibition, risk taking, response monitoring= 50 (10); cognitive empathy= 13.77 (1.16); affect recognition= 24.86 (2.11). Higher scores on measure of inattention, lack of insight, and severity of head injury indicate higher impairment/severity.

Table 10. 4 Frequency Table for Categorical Clinical Risk Factors

	Behaviour	No-Behaviour
Variable	Present	Present
History of Alcohol Use	1 (25%)	2 (15%)
History of Substance Misuse	3 (75%)	5 (33%)

Figure 10. 1 Boxplots for Groups by Measure



Note. Figure 10.1 represents means and confidence intervals for measures in each group.

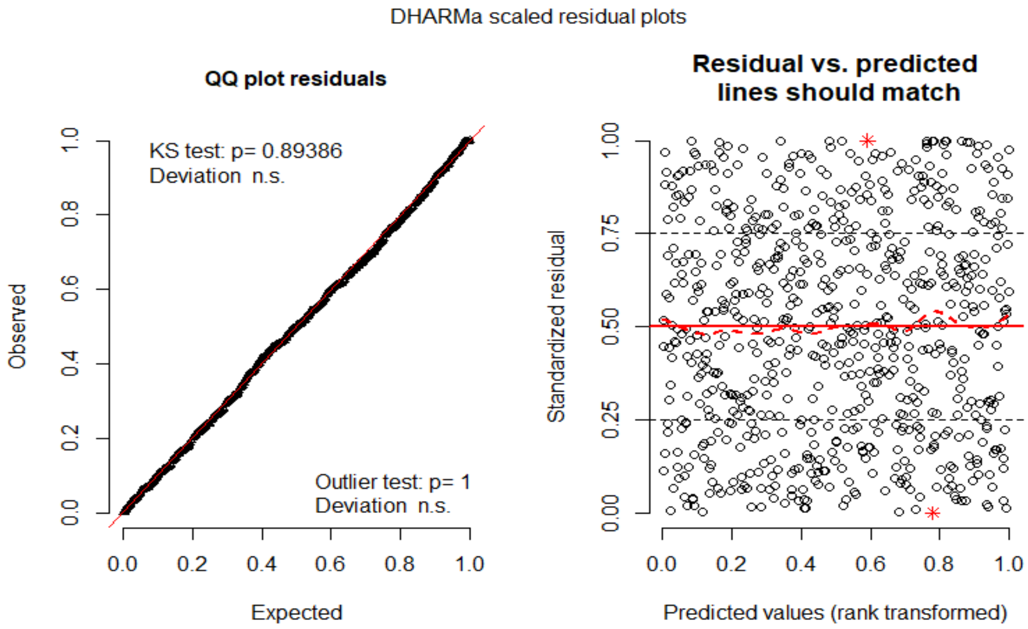
10.3.3 Does Performance on Neuropsychological Measures Explain the Variance in Inpatient Behaviours Associated with Imminent Aggression Overtime?

Null Model. The null model (model 0), with a random intercept only, was utilized to calculate the between subject variance in the outcome through intra-class correlations (ICC; 0.01), and to compare models with fixed effects. The small ICC reflects the separation of all nonzero data from the zero data, as when the ICC was calculated for a logistic mixed model only, it resulted in 0.59, indicating that 59% of the chances of a patient demonstrating behaviour indicative of aggression was explained by between subject differences, though in the two-part model, differences between patients only explained 1% of the variance in behaviour. The initial model was built using only a random intercept in the linear part of the model and was then extended to include a random intercept in the logistic part, providing a significantly better fit (x^2 (2)= 15.87, p<.001), thus random intercepts were included in the linear and logistic parts of all models.

Fixed Effects. After building the null model (Model 0), the effect of time was investigated (Model 1) but did not significantly explain the variance in behaviour, however it did improve model fit and decreased the unaccounted-for variance, as demonstrated through the deviance information criteria, which decreased from 139.50 to 134.14. Following this, each variable and its interaction with time was examined individually. All interactions improved the model fit compared to the null, and decreased the unaccounted-for variance, though not significantly, with model 6 (response inhibition and time) demonstrating the largest decrease, from 139.50 (null model deviance) to 126.52. No covariates significantly accounted for the variance in behaviour over time or increased or decreased the odds of a patient demonstrating behaviour associated with aggression. Further, confidence intervals for the odds ratios included only small, insignificant effects. All model parameters, for all models are presented in Table 10.5.

Testing the Null Model. As all the models decreased the deviance value compared to the null model, additional analyses were conducted to further explore if they provided a significantly better fit for the data over the null. A likelihood ratio test was performed for the model that accounted for the most unexplained variance, namely, response inhibition. To compare models, the AIC values were examined and were interpreted as *smaller is better* where smaller values indicate a better model fit for the data. The addition of response inhibition and time (x^2 (6)= 12.85, p=.04) improved the fit above the null model, as it had lower AIC values, indicating that the addition of response inhibition to the model, provided the best fit for the data in this sample. Model fit was visually examined using Q-Q and scatter plots of simulated scaled residuals, and further indicated that this model fit the data well with no outliers or significant deviations from normality (Figure 10.2). Based on these findings, response inhibition provided the best fit out of all the predictors, however, it did not significantly account for the variance in the behaviour overtime or change the odds of demonstrating behaviour associated with aggression.

Figure 10.2 Residuals Plot for the Response Inhibition model



Note. Figure 10.2 represents scaled residuals for the response inhibition model. Graph was created using the *GLMMAdaptive* and *Dharma* packages in R and is presented with 1000 simulations; *KS test*= Kolmogorov-Smirnov normality test

Table 10.5 Model Parameters for Fixed Effects Examining the Association Between Performance on Neuropsychological Measures and DASA-IV Ratings Over Time

		Fixed Effects (Continuous non-zero outcome)				Zero-Part (Binary outcome)						Nakagawa R^2	
Model		Estimate	SE	*z*	95% CI	Estimate	SE	*z*	OR	OR 95% CI	Deviance	Marginal	Conditional
0	Intercept	0.52	1.13	0.45	[-1.71, 2.75]	5.89	1.42	4.13			139.40		.01
1	Intercept	0.40	0.24	1.62	[-0.08, 0.88]	5.91	1.35	4.38					
	Time	0.04	0.02	-1.60	[-0.00, 0.08]	-0.04	0.02	-1.60	0.96	[0.91, 1.00]	134.14	.15	.17
2	Intercept	0.12	0.79	0.15	[-1.44, 1.68]	5.87	1.32	4.44					
	Time	0.04	0.02	2.16	[0.00, 0.07]	-0.04	0.03	-1.54	0.96	[0.91, 1.01]			
	Reasoning	-0.02	0.02	-1.01	[-0.05, 0.02]	0.03	0.06	0.60	1.04	[0.92, 1.16]			
	Reasoning:Time	0.002	0.001	0.27	[-0.00, 0.00]	0.0005	0.002	0.27	1.00	[0.99, 1.00]	131.04	.28	.31
3	Intercept	0.16	0.98	0.16	[-1.77, 2.09]	5.85	1.30	4.49					
	Time	0.03	0.03	0.94	[-0.03, 0.08]	-0.04	0.02	-1.55	0.96	[0.91, 1.01]			
	Inattention	0.01	0.01	0.94	[-0.01, 0.04]	0.01	0.05	0.25	1.01	[0.91, 1.22]			
	Inattention: Time	-0.002	0.002	-0.72	[-0.00, 0.00]	0.0002	0.002	0.10	1.00	[0.99, 1.00]	131.72	.18	.23
4	Intercept	0.21	0.35	0.58	[-0.48, 0.90]	6.09	1.40	4.33					
	Time	0.05*	0.02	2.61	[0.01, 0.09]	-0.06	0.03	-1.85	0.94	[0.88, 1.00]			
	Response Inhibition	-0.04	0.04	-1.27	[-0.11, 0.02]	0.09	0.33	0.71	1.09	[0.84, 1.43]			
	Response Inhibition: Time	-0.001	0.003	-0.36	[-0.01, 0.00]	-0.006	0.005	-1.29	0.99	[0.98, 1.00]	126.52	.30	.31
5	Intercept	0.03	0.50	0.10	[-0.95, 1.02]	6.03	1.44	4.16					
	Time	0.07	0.05	1.41	[-0.03, 0.18]	-0.04	0.03	-1.60	0.95	[0.90, 1.01]			
	Risk Taking	-0.07	0.07	-0.94	[-0.22, 0.07]	0.06	0.11	0.56	1.06	[0.85, 1.31]			
	Risk Taking: Time	0.007	0.01	0.80	[-0.01, 0.02]	-0.002	0.004	-0.43	0.99	[0.98, 1.01]	132.86	.62	.69
6	Intercept	0.02	3.80	4.36	[-7.43, 7.46]	5.87	1.34	4.36					
	Time	0.05	0.06	0.72	[-0.08, 0.17]	-0.05	0.03	-1.64	0.94	[0.88, 1.01]			
	Response Monitoring	-0.02	0.04	-0.48	[-0.10, 0.06]	0.05	0.08	0.70	1.06	[0.90, 1.23]			

	Resp. Monitoring: Time	0.001	0.005	0.12	[-0.00, 0.01]	-0.002	0.003	-0.74	0.99	[0.99, 1.00]	131.38	.26	.29
7	Intercept	0.11	0.76	0.15	[-1.37, 1.61]	5.99	1.38	4.31					
	Time	0.05*	0.02	2.36	[0.01, 0.09]	-0.04	0.03	-1.43	0.96	[0.91, 1.01]			
	Lack of Insight	-0.23	0.28	-0.79	[-0.79, 0.33]	-0.09	0.53	-0.17	0.91	[0.32, 2.59]			
	Lack of Insight: Time	-0.004	0.02	-1.17	[-0.05, 0.04]	-0.02	0.02	-0.87	0.97	[0.93, 1.02]	128.80	.34	.38
8	Intercept	0.12	1.02	0.12	[-1.97, 2.11]	5.88	1.34	4.37					
	Time	0.02	0.03	0.76	[-0.04, 0.08]	-0.04	0.03	-1.52	0.95	[0.91, 1.01]			
	Cognitive Empathy	-0.04	0.07	-0.52	[-0.17, 0.09]	-0.16	0.29	-0.54	0.85	[0.48, 1.51]			
	Cognitive Empathy: Time	0.005	0.01	0.65	[-0.10, 0.02]	0.0005	0.01	0.06	1.00	[0.98, 1.01]	131.56	.11	.18
9	Intercept	0.05	0.92	0.06	[-1.76, 1.87]	5.97	1.39	4.28					
	Time	0.02	0.03	1.00	[-0.02, 0.08]	-0.04	0.03	-1.59	0.96	[0.91, 1.01]			
	Affect Recognition	-0.04	0.05	-0.82	[-0.15, 0.06]	-0.04	0.19	-0.19	0.96	[0.65, 1.42]			
	Affect Recognition: Time	0.004	0.01	0.73	[-0.00, 0.02]	0.001	0.01	0.11	1.00	[0.98, 1.01]	131.72	.16	.25

Note. Continuous, nonzero outcome fixed effects include only the patients who had a positive behaviour rating (*n*=4); *SE*= standard error; *95% CI*= 95% confidence intervals; *OR*= Odds Ratio; *OR 95% CI*= confidence intervals for the odds ratio; *Nakagawa* R^2= model fit (*marginal* =model fit for fixed effects; *conditional*= model fit for full model); *Resp. Monitoring*= Response Monitoring,.
*95% CI do not cross the null line

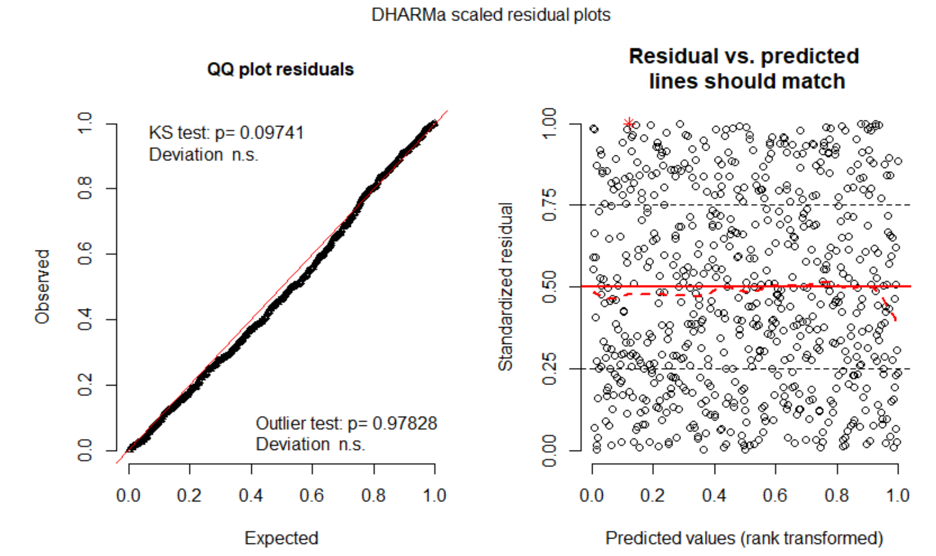
10.3.4 Do Clinical Factors Explain the Variance in Inpatient Behaviour Associated with Imminent Aggression Overtime?

Null Model. As the same dependent variable was investigated for the exploratory analyses, the parameters for models zero (null model) and one (effect of time) remained the same.

Fixed Effects. All the models accounted for a portion of the unexplained variance and improved model fit over the null model, however, all confidence intervals crossed the null line. The interaction between severity of head injury and time in the linear mixed model accounted for the most unexplained variance decreasing it from 139.4 to 120.8. Model parameters for all clinical factor models are presented in Table 10.6. The confidence intervals for the odds ratios were variable ranging from negligible to large.

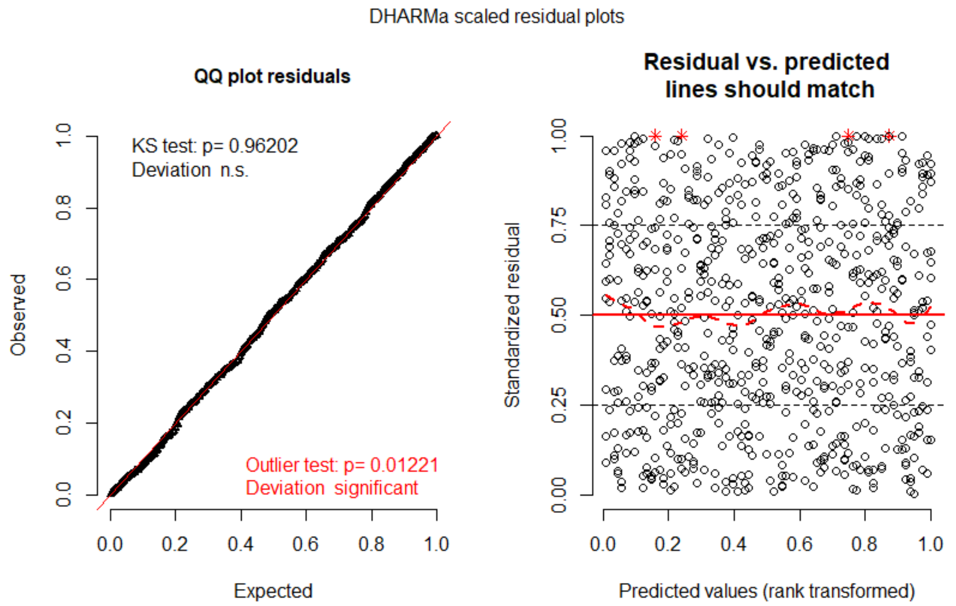
Testing the Null Model. The model that accounted for the most unexplained variance and improved model fit was severity of head injury, which was compared to the null model using a likelihood ratio test. According to the likelihood ratio test, there was a significant improvement made to the null model by adding severity of head injury and time to the model (x^2 (6) = 19.01, *p*= .004), indicating that the addition of this covariate significantly improved the accuracy of the model, though there was likely not enough power to for coefficients to reach significance. The Q-Q and scatter plots indicated that this model fit the data well with no outliers or significant deviations from normality (Figure 10.3). The model for substance use was also statistically significant (x^2 (6)= 12.56, *p*= .05) and the model for alcohol use (x^2 (6)= 11.39, *p*= .08) was approaching significance. The Q-Q and scatter plot for the substance use model indicated that the model fit the data well, however, there were significant outliers (Figure 10.4).

Figure 10.3 Residuals Plots for the Severity of Head Injury Model



Note. Figure 10.3 represents scaled residuals for the response inhibition model. Graph was created using the *GLMMAdaptive* and *Dharma* packages in R and is presented with 1000 simulations; *KS test*= Kolmogorov-Smirnov normality test

Figure 10. 4 Residuals Plots for the Substance Use Model



Note. Figure 10.4 represents scaled residuals for the response inhibition model. Graph was created using the *GLMMAdaptive* and *Dharma* packages in R and is presented with 1000 simulations; *KS test*= Kolmogorov-Smirnov normality test

Table 10.6 Model Parameters for Fixed Effects Examining the Association Between Clinical Measures and DASA-IV Ratings Over Time

		Fixed Effects (Continuous non-zero outcome)				Zero-Part (Binary outcome)						Nakagawa R^2	
Model		Estimate	SE	*z*	95% CI	Estimate	SE	*z*	OR	OR 95% CI	Deviance	Marginal	Conditional
0	Intercept	0.52	1.13	0.45	[-1.71, 2.75]	5.89	1.42	4.13			139.4		.01
1	Intercept	0.40	0.24	1.62	[-0.08, 0.88]	5.91	1.35	4.38					
	Time	0.04	0.02	-1.60	[-0.00, 0.08]	-0.04	0.02	-1.60	0.96	[0.91, 1.00]	134.14	.15	.17
2	Intercept	-0.71	1.20	-0.59	[-3.07, 1.64]	8.15	2.49	3.27					
	Time	0.06	0.06	0.99	[-0.06, 0.17]	-0.16	0.11	-1.38	0.85	[0.68, 1.07]			
	Substance Use	1.03	0.95	1.08	[-0.83, 2.90]	-3.18	2.33	-1.37	0.04	[0.00, 3.99]			
	Substance Use: Time	-0.01	0.06	-0.11	[-0.13, 0.11]	0.13	0.11	1.11	1.14	[0.90, 1.43]	128.82	.40	.42
3	Intercept	0.38	0.59	0.65	[-0.77, 1.53]	5.89	1.47	3.99					
	Time	0.05	0.02	2.62	[0.01, 0.09]	-0.03	0.03	-1.02	0.97	[0.92, 1.02]			
	Alcohol Use	-0.98	0.95	-1.04	[-2.83, 0.87]	1.68	2.63	0.63	5.38	[0.03, 935.9]			
	Alcohol Use: Time	0.01	0.06	0.14	[-0.11, 0.13]	-0.14	0.12	-1.12	0.87	[0.67, 1.11]	128.7	.36	.37
4	Intercept	1.89	1.36	1.39	[-0.77, 4.56]	6.48	1.62	3.98					
	Time	0.04	0.09	0.41	[-0.14, 0.22]	0.01	0.06	0.09	1.01	[0.87, 1.15]			
	Severity of Head Injury	1.01	0.95	1.06	[-0.85, 2.88]	1.57	0.84	1.85	4.81	[0.91, 25.37]			
	Severity of HI: Time	-0.01	0.06	-0.14	[-0.13, 0.11]	0.03	0.04	0.73	1.03	[0.94, 1.12]	120.8	.66	.63

Note. Continuous, nonzero outcome fixed effects include only the patients who had a positive behaviour rating (*n*=4); *SE*= standard error; *95% CI*= 95% confidence intervals; *OR*= Odds Ratio; *OR 95% CI*= confidence intervals for the odds ratio; *Nakagawa* R^2= model fit (*marginal* =model fit for fixed effects; *conditional*= model fit for full model); *Severity of HI*= Severity of Head Injury

10.3.5 Results of Pilot Research Questions

The findings did not provide evidence to support either of the research questions. Thus, the implementation of a repeated measures design and the use of an outcome measure of behaviours leading to situational aggression rather than violence, did not increase the number of recorded incidents, and therefore did not increase the variance in the outcome in this case.

10.4 Discussion

10.4.1 Main Findings

Using two-part mixed models for semi-continuous data, the association between cognitive and clinical risk factors and the variance in behaviour over time, was investigated. Overall, all 95% confidence intervals for the interactions in all models crossed the null line suggesting uncertainty in the true estimates and effects. Further, there was limited evidence to suggest that performance on cognitive measures or scores on clinical measures will significantly explain the variance in behaviour over time and increase the odds of a patient demonstrating behaviour associated with aggression in a future study. However, the addition of response inhibition and severity of head injury to the null model demonstrated positive findings as they both improved the model fit, though neither measure explained the variance in behaviour over and above time, suggesting that there is likely to be another explanation for the occurrence of the behaviour.

10.4.2 Performance on Cognitive Measures and Behaviours Associated with Aggression

Upon examination of the boxplots in Figure 10.1 for the neuropsychological measures, there were no notable differences between the behaviour and no-behaviour groups as indicated by the overlap in confidence intervals between the groups. However, the results are likely to be biased because of the small sample size, skewness of the data, and due to the magnitude of difference between the observations in each group, where the behaviour group had 168 observations and the non-behaviour group had 546 observations. Mean scores on neuropsychological measures ranged from below average to low average and the mean for cognitive empathy was within the impaired range for the no-behaviour group, indicating that this small sample is mostly representative of wider forensic samples which often demonstrate low average to impaired functioning on similar measures (e.g., O'Reilly et al 2015). Although low means were evidenced, a possible explanation for these null findings is that performance on cognitive measures moderates the relationship between another patient characteristic and behaviour associated with aggression over time, or that they have a cumulative relationship with other variables on the outcome, though, due to the small sample size, this could not be investigated. Moreover, it is likely that the study did not have enough statistical power to differentiate the groups based on measure performances, the primary outcome did not have enough variance, or that cognitive abilities truly do not explain behaviour variability.

The improvement of the null model after the addition of response inhibition suggests that perhaps in a larger sample, these findings would be significant, or that they may cumulatively account for the variance in behaviours associated with aggression, however, the small odds ratio and the 95% confidence intervals suggest that the magnitude is insignificant. Moreover, the estimate for the interaction in the linear model is interpreted by exponentiating it, subtracting from 1 and multiplying it by 100 to get the percentage increase or decrease in the outcome for every one unit of change in the predictor. Thus, for every one unit decrease in response inhibition scores, there is a 1% increase in behaviour indicative of aggression which is also negligible. Taken together, the hypothesis was not fully supported by the data, however, given that the model for response inhibition improved the fit over the null model, even if it was an artifact of methodological limitations, the null hypothesis cannot be fully accepted, thus, it may be of interest for some to further examine this in a larger sample.

10.4.3 Clinical Risk Factors and Behaviour Associated with Aggression

The analysis of clinical risk factors provided little evidence of a strong association with indicators for behaviours related to risk of aggression. However, severity of head injury had the largest decrease in unexplained variance and improved the null model, and similarly, the models for historical substance and alcohol use were approaching significance for improving the null model, and whilst the confidence intervals crossed the null line for all estimates and effects, except for time, these findings are in line with similar research. For example, clinical risk factors such as substance misuse (e.g., Fernández-Serrano et al., 2011), harmful alcohol use or dependence (e.g., José Fernández-Serrano et al., 2010), and head injury (e.g., Muccigrosso et al., 2016) contribute to cognitive impairments, suggesting that they also contribute to more variable behaviour. Though, the null hypothesis, that is, clinical factors have no effect on the variance in behaviours associated with aggression risk, cannot be ruled out based on these findings. Like the cognitive measures, these measures all boasted confidence intervals that ranged from negligible to large, indicating uncertainty in where the true findings sit. Further, although severity of head injury accounted for the most unexplained variance, the unexplained variance remained high at 120.8, indicating that there are additional mechanisms that explain the variance of behaviour that were not examined in this study. Overall, the model for severity of head injury not only provided the best fit as indicated by Nakagawa R^2, but also explained the most variance as indicated by the deviance information criteria making it the best explanation for variations in behaviour in the current study. In line with past research, head injuries and TBI may result in impairments that inhibit an individual's ability to comply with correctional authorities, in addition to tending to their activities of daily living (Slaughter et al., 2003), consequently affecting their treatment as well as their ability to be released. The association between head injury and increased aggressive and antisocial behaviours has been well established in the literature (e.g., Brown et al., 2019; Fazel et al., 2011; Williams et al., 2018), and is discussed in depth in Chapter 2 of this thesis. Similarly, the neuropsychological effects of historical substance misuse and the association with aggressive and violent behaviour have been thoroughly discussed in the literature (e.g., Chandler et al., 2009; Hoaken & Stewart, 2003), and are also discussed in depth in Chapter 2 of this thesis.

10.4.4 Pilot and Null Findings

The current study was designed with the intention to overcome the methodological limitations that are often apparent in forensic inpatient research, namely, small sample size and a low base rate of incidents. To combat these, a repeated-measures design was implemented to, (1) increase statistical power (while allowing for a smaller sample size) and (2) to overcome a low base rate of incidents by observing and rating patients 42 times in a two-week period, thus introducing more time points to theoretically decrease the opportunity for incidents to go unrated or missed. Moreover, the DASA-IV was carefully chosen as an outcome as it was specifically designed to measure inpatient behaviour. However, while the repeated measures did increase statistical power, the low base rate of incidents persisted. Hence, it is apparent that it is not the study design, or the number of times patients' behaviour is measured or how it is measured, but it is likely due to the wellness of the individual patients and the stability of their environment. For example, the patients who participated in the current study were well enough to consent, were abstinent from drugs and alcohol before and during the time of the study and were settled in a stable environment with minimal de-stabilizers. Additional postulations have been made to explain this, for example, a patient displaying irritable behaviour may be observed by nurses as having a bad day and a DASA-IV rating may not seem appropriate; also nurses who have known patients for a long period of time may not see patient behaviour as an indicator for later aggressive and thus may not rate it as such (e.g., observer bias); it is also possible that if a patient became irritable or displayed another behaviour encompassed by a DASA-IV item, that it went unnoticed; and lastly, as with many studies of inpatient aggression and violence, institutionalised behaviour may have played a role, that is, patients may have adapted their behaviour to the services (Chow & Priebe, 2013), and may not react in the same way to certain situations in hospital as they would in the community. Research has also suggested that psychiatric nurses are often well versed in noticing patients who are at a greater risk for inpatient violence and may therefore act before an incident has the chance to occur (Ogloff & Daffern, 2006). Moreover, the DASA-IV was designed to capture change in patients' behaviour (Ogloff & Daffern, 2006), thus, if a patient's baseline behaviour is irritable, for example, unless the irritability increased, the behaviour is unlikely to receive a rating. Another possible explanation is, as per hospital policy, nurses at the State Hospital are only required to rate patient behaviour on the DASA-IV when the patient is either new, and therefore their behaviour variation is unknown, or when the patient has shown a new pattern of behaviour indicative of risk during their time at the hospital. Of the n= 17 patients who participated in this study, only n= 2 were required to have their behaviour rated daily before consenting for the study, suggesting that the baseline behaviour of the remainder of this sample may have been much less variable than might be expected. Further, the seven items on the DASA-IV do not measure violent behaviour, instead they signify the presence of behaviours that are likely to lead to situational aggression (Ogloff & Daffern, 2006), which lead to the hypothesis that perhaps capturing these behaviours would broaden the target behaviour, increasing the base rate, though, this was not the case. Based on these findings, it appears that inpatient settings may not be suitable for the examination of variance in patient behaviour, at least over a short period of time. Future research in this area should focus on longer follow-up periods, and perhaps patients who are not yet clinically stable and settled into their environment.

10.4.5 Limitations

This study has several noteworthy limitations. For instance, the study took place in an inpatient forensic high secure hospital consequently constraining generalisability of the findings, however, obtaining the ratings on the same patients three times a day would have been difficult had the study taken place in the community. Further, the entire sample was made up of men, it is possible therefore that women's behaviour may have varied more, and their performance on cognitive measures may have affected behaviour associated with aggression differently.

Whilst the use of two-part mixed models allowed the clustering of repeated measures data to increase power for statistical analyses, it is still likely that a larger sample would have provided more robust estimates of the model parameters, and the lack of robust findings is perhaps the result of insufficient statistical power to identify covariates which explain a proportion of the variance in behaviour over time, as well as magnitude of difference between patients who presented behaviour associated with aggression and those who did not. Further, findings are biased, as the only continuous outcomes that were analysed were nonzero and were only for the participants who demonstrated rateable behaviour, in this study, that was only n=4 patients.

Furthermore, this analysis was performed in this study to demonstrate its utility in studies of violence risk to increase specificity of findings by obtaining model estimates for severity of an outcome, as well as the odds of the outcome occurring, though, due to the low base rate in this study, it is likely that the models were overly complex for this data, thus results should be interpreted cautiously. This is likely to be the explanation for the statistical significance of the head injury and substance misuse models, as all confidence intervals for individual variables in the model crossed the null line. Last, the null results may be a product of a flaw effect in the study design as the data were skewed and had low variance, however, these results are generally in line with the findings from the previous study in this thesis.

The study also took place over a short period, which was also likely to have impacted to the low base rate of rateable behaviour. Research on short-term risk assessments has suggested that monitoring patients who are clinically stable over a short period is neither as useful nor realistic as longer follow-up periods, especially when it comes to picking up psychiatric symptomology (Grevatt et al., 2004). However, obtaining three ratings per patient, per day for 14 days involved considerable efforts from all parties. A longer period may have resulted in significant missing data, as well as substantial burden on ward staff.

10.4.6 Implications and Recommendations for Similar Research

Limited clinical implications can be made based on these findings due to the sample size, however, deficient response inhibition, severe head injuries, and historical substance misuse may be indicators of less regulated behaviours which may be indicators for imminent aggression, though these are not necessarily new findings. For example, impulsive behaviour is a product of deficient response inhibition, and is an established risk factor for violence, similar to substance misuse, though TBI is not a formal risk factor on existing measures. Though, based on these findings, and bearing in mind the findings from Chapter 8 of this thesis, there is little evidence to suggest that performance on cognitive measures explain a significant proportion of the variance in violent and aggressive behaviours. There are a few methodological implications which can be made based on this study. For instance, it is apparent that the limitations outlined in Chapter 3 of this thesis are difficult to overcome even when they are carefully and critically considered in the development stages of a study. Perhaps the *glass ceiling* that risk assessments have hit are a product of these limitations, and study quality has also hit a *glass ceiling.* However, some aspects of this research remain in the control of researchers, for example, conducting prospective studies rather than cross-sectional, repeated measures when only a small sample can be obtained, and the inclusion of a daily measure of aggression or of functional outcomes to supplement the observational outcomes to create a better chance of robust statistical analyses even in the case of a low base rate of incidents. Additionally, the use of two-part mixed models for analysis is novel in this field and allows inferences to be drawn on the outcome as both continuous and binary, rather than sacrificing measures of severity or intensity due to low base rates or a large number of zeros and using it as a binary outcome only. As a result, these analyses may improve the specificity of findings in studies which examine violence as an outcome. Despite biases in the linear part of the model, this analysis appears promising for similar multilevel research, especially in pilot studies where estimates of effects are desired. Finally, it is well known that human behaviour is difficult to predict, thus, if researchers continue to critique studies of low quality, critically evaluate their own methodologies, and aim to be specific and thorough in their definitions and measures of outcomes, as well as innovative in study designs and analyses, it is likely that the research in this field will slowly increase in quality.

10.5 Conclusion

The findings of this study imply that in an unchanging environment, abstinent of drugs and alcohol, and with stability of mental health, performance on neuropsychological measures and scores on clinical measures may have little effect on the variance of inpatient behaviour. As these results lack generalisability to community and outpatient settings, it is posited that they would differ in these settings where drugs and alcohol are more likely to be available, and other de-stabilisers might be present, though this was not the case in the pilot and feasibility study presented in Chapters 6-9 in this thesis. Albeit response inhibition and severe head injury emerged as possibly important explanatory variables for variance in inpatient behaviour, however, the methodological limitations of the study constricted the robustness of these findings, thus a future, larger study may be warranted to further examine these relationships to gain a more accurate understanding of the underlying mechanisms related to the variance in violence risk.

11 General Discussion

Chapter Preface

This chapter concludes the thesis by summarising the major findings from the preceding chapters. Following a final discussion about the main research question for this thesis, clinical and research implications are discussed, with recognition to the limitations of the research in this thesis. Finally, several recommendations are offered based on the implications of the findings, in addition to the scope for future research, and the way forward in this field.

11.1 Thesis Aims

The assessment of violence risk and the management of offenders have improved considerably in the last few decades; however, violence risk assessments appear to have hit a 'glass ceiling' with most predicting outcomes with only moderate effect sizes. Respectively, there has been an increase in published literature examining the association between cognitive impairments and violence and aggression, which has contributed to the postulation that perhaps the addition of cognitive risk factors for violence may have the ability to improve the predictive performance of risk assessments and break through the glass ceiling. Whilst efforts have been made to determine whether this is the case, the apparent need for innovation in this field has seemingly overshadowed some of the fundamentals in research, introducing the same methodological weaknesses into the neuropsychology research that have plagued the general risk assessment field of research since before the 1980's. Methodological weaknesses such as poorly defined outcomes, the wide range of available neuropsychological assessments in combination with overlaps between measured concepts, cross-sectional studies, a lack of transparency in reporting findings, and the lack of published null findings have resulted in mostly small, heterogenic effect sizes with wide confidence intervals and a general lack of certainty regarding the true relationship between cognitive abilities and violence risk.

Despite the uncertainty perpetuated by the methodological limitations, there is strong theoretical support for the relationship between cognitive impairments and violence risk, based partially on the scope of cognitive impairments which often characterise forensic populations, and more so on the behavioural products of cognitive dysfunction often making it difficult for individuals to desist from offending and to function in a structured *normative* society. This thesis examined the ability of neuropsychological measures to explain violent outcomes, and to add incremental validity to contemporary risk assessments. The central aim was to identify neuropsychological domains relevant for inclusion in a structured professional judgement tool that could be used in parallel with existing tools to increase both their accuracy in predicting risk and their utility in treatment planning. It was hypothesised that the inclusion of these specific assessments would improve the accuracy of clinical judgements and formulations of risk, increasing the ability to accurately target treatments; and to adapt treatments to maximise their efficacy. Operationalising the aim resulted in a selective research strategy which followed the example of the National Institute of Mental Health's MATRICS initiative's development of the MATRICS Consensus Cognitive Battery (MCCB) which sought to identify not only cognitive domains identified by the research literature, but also those that, in a relatively new field, may have yet to be examined sufficiently through expert consensus (Nuechterlein et al., 2008). This work led to four pieces of novel research. The novel research offers evidence on the correlations between disaggregated cognitive abilities and violent outcomes in violent, forensic populations, the first expert ranked list of important cognitive risk factors to consider for violence risk assessment, initial findings for the predictive and incremental validity of neuropsychological measures for violent and similar outcomes in two distinct violent populations, as well as a first look at neuropsychological explanations for variance in daily behaviour of forensic inpatients which may be indicative of aggression. In addition to these findings, the current thesis examined the literature on a conceptual level using a bottom-up approach to identify areas where researchers can make improvements to increase the quality of future research in this area.

11.2 Summary of Findings

Meta-Analyses. The systematic review and meta-analysis in Chapter 4 critiqued and synthesized the extant literature on the cognitive differences between violent and non-violent offenders, and the strength of the relationship between neuropsychological measures and violent outcomes. With the inclusion of 42 studies, three out of 12 measures homogenously and significantly differentiated violent and non-violent offenders (e.g., reasoning: g= 0.26, impulsivity: g= 0.28, and expressive speech: g= 0.85) and three out of five measures (e.g., impulsivity: r= .27, attention: r= -.16, and insight: r= -.17) significantly correlated with violence risk prospectively. The between group portion of the review evidenced that the largest effect sizes came from studies which did not use narrow inclusion criteria, and those which used the LNNB. However, the vast range of measures used across the included studies was posited to have contributed to the heterogeneity and wide range of effect sizes (e.g., g= -0.76 to 0.85), as well as the small sample sizes and magnitude of difference between groups, perhaps increasing Type I error. The association portion of the review suggested that measures of attention, impulsivity, and lack of insight would independently add approximately 3-7% incremental variance to a risk assessment, which is unlikely to be useful for improving their predictive accuracy. Given the state and poor quality of much of the literature, it was difficult to draw definitive conclusions, though it was postulated that most measures of cognitive abilities do not have the ability to add significant incremental validity to risk assessments. As a result of the multidimensional underlying structure of many neuropsychological measures, with several different cognitive abilities interacting to explain one given performance (Duggan & Garcia-Barrera, 2015; Karr et al., 2018), and their reputation for task impurity, it was recommended for the field to reach a consensus on which cognitive measures are the most robust in measuring cognitive risk factors for violence, and similarly, that a consensus be reached on the cognitive domains which are most important to violence risk. It was further recommended that, to move this field of research forward, transparent, high quality, replicable studies using the same neuropsychological measures in various forensic samples, are crucial.

Delphi Study. Chapter 5 demonstrated the use of the Delphi method to obtain a list of cognitive abilities considered relevant for inclusion in a violence risk assessment from a panel of experts. The three round Delphi method proved successful, funneling 41 separable cognitive domains into a ranked list of 10, namely, poor inhibitory control, behavioural impulsivity, risk taking, affective empathy, response monitoring, affect recognition, cognitive insight, reasoning, cognitive empathy, and clinical insight. With the inclusion of researchers and clinicians, this study also highlighted the possible lag between research and clinical practice and emphasized the differences of opinion between the two groups.

Assessing the Cognitive Contributors to Violence. The feasibility and pilot study presented in chapters 6-9 implemented the evidence obtained from the meta-analysis and Delphi study to build a neuropsychological battery to pilot in two samples of violent offenders, and to test the feasibility of implementing a neuropsychological assessment into practice to aid in the identification and assessment of cognitive impairments relevant to violence risk. The examination of participant scores on baseline measures indicated that there were only significant between group differences with large effect sizes on measures of cognitive empathy (d= 0.91) and understanding simple sarcasm (d= 1.05) where the inpatient group performed more poorly. Although it was hypothesised that the inpatient sample would perform significantly more poorly on all measures relative to the community group due to their additional risk factors for cognitive impairment, such as psychosis, this was not the case. This finding is similar to that found in the between-groups portion of the review and may be explained by subclinical differences rather than clinical differences, as when scores were transformed into z-scores to obtain the proportion of participants with impairments, several measures evidenced high proportions of impairments relative to normative scores. This result not only emphasized how misleading group means can be, but also the pervasive cognitive impairments observed in violent offenders with and without major mental illness.

Feasibility Findings. The feasibility findings suggested that the criteria for the tolerability of the neuropsychological measures and the completion rate of measures were met. The criterion for the recruitment rate was partially met; however, amendments were made to recruitment procedures while the study was ongoing to reimburse the community participants for their time which proved to be successful and is recommended for future studies of this nature. Moreover, the members of the social work team were unable to record the reasons why service-users were excluded and were also unable to offer participation to all clients, a limitation which needs to be addressed in future, similar studies. ROC curve analyses demonstrated that measures of reasoning, inattention, response inhibition, and affect recognition were sufficiently accurate, sensitive, and specific to prospective violence, and measures of reasoning, cognitive empathy, and lack of insight had sufficient performance indicators for retrospective violence in the inpatient sample. In the community sample, only the measure of inattention was sufficient for prospective violence, and no measures were sufficient for retrospective violence. Incremental validity was examined using the independent variance (i.e., standardised residuals) of neuropsychological measures, only on measures which had an AUC score $\geq .70$, which resulted in reasoning, inattention, response inhibition, and affect recognition remaining significant, after controlling for the HCR-20, for prospective violence, and only lack of insight for retrospective violence. In the community sample, only the measure of inattention was analysed in this way, which remained significant after controlling for the LSI-R:SV, however, there was only one prospective violent incident in the LSI-R:SV subsample, and two for prospective violence in the inpatient sample, which likely inflated effects. The feasibility findings suggest that the methods and procedures are feasible for a future study with minor changes, though not all measures demonstrated sufficient performance indicators for predicting violence, especially in the community sample. This finding may indicate that the domains that these measures assess are not sufficiently related to violence and may need to be reconsidered should a future study move forward. Additionally, in the case of a future, larger study, it is recommended that the implementation of different or additional measures also be considered.

Primary and Secondary Findings. Primary and secondary findings in the inpatient sample suggested that the addition of affect recognition and lack of insight to the HCR-20 explained an additional 18% of the variance in prospective violence severity, and both variables remained significant after controlling for head injuries with LOC, number of head injuries, severity of head injuries, and alcohol and substance misuse. Positive emotion recognition with lack of insight also explained a significant proportion of the variance in violence severity (prospective), and both variables remained significant in the presence of head injuries with LOC, number of head injuries, severity of head injuries, alcohol and substance misuse, and the HCR-20, explaining an additional 22% of the variance over and above the HCR-20. Whilst these findings appear to be promising, the direction of affect recognition and positive emotion recognition were both in the opposite direction than expected, suggesting that better affect recognition is associated with more severe violence. Although explanations have been made for this with reference to the literature, this finding is in stark contrast to the hypothesis that impairments in cognitive abilities will relate to violent outcomes. As these findings may be an artifact of the study limitations, related to the potentially non-representative sample size, or that the measure lacks sufficient sensitivity and specificity; the nature of the relationship between violence severity and affect recognition deserves further investigation. Last, response inhibition was a significant explanatory variable for a *perceived risk* outcome, operationalised as the mean number of days participants were on elevated levels, and individually explained 8% of the variance.

The data from the community sample did not support any of the hypotheses. In exploratory analyses, response monitoring emerged as significant variable associated with retrospective antisocial behaviour, however, after controlling for the LSI-R:SV, it was no longer significant. This is likely to be an artifact of the study, as the LSI-R:SV was only available in a subset of n= 18 participants, and response monitoring only remained significant in the model with the entire sample of n= 31, and not in the model with the subsample even without controlling for the LSI-R:SV. However, response monitoring may be an important variable to explain antisocial behaviour and warrants further investigation. Finally, inattention explained a significant proportion of the variance in self-reported impulsivity, which served as a proxy measure for violence risk as impulsivity is a well-established risk factor for violence, individually explaining 11% of the variance.

When the two samples were combined to examine self-reported impulsivity, only cognitive empathy emerged as a significant explanatory variable, however, it was in the opposite direction than expected, suggesting that better cognitive empathy was associated with higher impulsivity. This may also be a product of the study limitations; however, it may be that the TASIT is not a suitable measure for research of this nature, as this is the first study to examine its predictive validity for violence risk.

Taking the presented findings, the range of small to large effect sizes found in this study are mostly consistent with those found in the in the meta-analysis. Moreover, the variation in effects demonstrated in both studies further showcases the significant heterogeneity in this field, and the need for more replicability of high-quality studies using the same predictors, criterion measures, and prospective designs across similar settings and populations.

Case-Series Findings. The final study in the thesis was a pilot case-series, designed with careful consideration of the common methodological weaknesses in this field of research, with the aim of explaining the variance in behaviours that indicate imminent situational aggression. While no variables explained the variance over and above time, the models for response inhibition, severity of head injury, and history of substance misuse explained the most unaccounted for variance (i.e., deviance), and fit the data well, suggesting they permit further investigation. Although efforts were made to overcome the methodological limitations, they persisted. The findings from the study indicated that in a stable environment, such as a secure hospital, scores on neuropsychological measures had little effect on the variance of inpatient behaviour over a short period of time.

11.3 Do the Addition of Cognitive Measures to Risk Assessments Increase Their Predictive Accuracy?

Taken together, this thesis highlights the glass ceiling effect in violence risk assessments and offers a strong theoretical and empirical hypothesis for how the glass ceiling might be broken. Further, criticisms of risk assessment research were carefully considered in the design stages of all studies with the aim to overcome these limitations, in order to find valid and robust results. However, even in the presence of careful planning and consideration of methods and procedures, the findings of this thesis lend little evidence to suggest that the addition of neuropsychological measures to risk assessments will significantly increase their predictive accuracy. This conclusion is based on the findings that several measures were not associated with violent outcomes in either sample, in addition to the inconsistency in effects found in the meta-analysis. As the findings go against the central hypothesis and the strong theoretical support, these results are somewhat surprising. As stated earlier, it may be that these domains were not relevant to violence risk, or that the measures were not sufficiently sensitive or specific to the outcomes, in which case, a future, larger study is necessary. However, most of the effect sizes in the meta-analysis were small, and likewise, the effect sizes in the main empirical study reached a large magnitude for only specific variables, suggesting that these findings are consistent with similar literature. Further, bootstrapped confidence intervals often evidenced a wide range of estimates and effects even in a population of n= 1,000 individuals, further evidencing the heterogeneity seen in cognitive abilities, and violent populations. While it may be worth further investigating the few measures which were significant, namely, affect recognition, positive affect recognition, lack of insight, response monitoring, reasoning, response inhibition, and inattention in a larger study, it is opined that the limitations will persist, and the findings will not significantly differ. Furthermore, translating research into practice requires robust, consistent evidence to convince services of the benefits of implementing new ideas, which this field of research is presently lacking. Thus, if researchers in this field wish to hold on to this hypothesis and one day implement this into practice, the research quality must improve, and high-quality studies must be replicated to increase the strength and consistency of the evidence. Further, collaborative efforts between researchers and disciplines are essential to move this field forward.

Although the thesis findings did not support the central hypothesis, it cannot be concluded that the implementation of neuropsychological assessments for use alongside of violence risk assessments is not warranted. Highlighted in the following sections are several ways in which assessing cognitive functions may be helpful to clinicians, services, and the patients themselves, assisting in identifying cognitive strengths and weaknesses, tailored treatment styles, identifying capacity to make decisions, as well as fitness to plead. In addition to these benefits, it should be acknowledged that risk assessment is broad and includes measures for different groups of offenders, different types of violence and offending, and a wide range of risk and protective factors, all of which individually differ in their ability to predict risk, and for some, to inform treatment and management. Thus, it may be that the prediction of risk using neuropsychological measures might look different than existing risk assessments where instead of directly predicting violence risk, they may aid in the identification of those who are less likely to benefit from treatment programmes as a result of their cognitive functioning, and therefore, those who may be more likely to recidivate or remain in a high secure setting for a longer time period. Moreover, the lack of effects found in the current thesis may suggest that cognitive impairments might be moderated by certain variables which weaken the strength of the relationship between cognitive measures and violence, such as being in a treatment setting or having access to treatment, being on psychotropic medications, and the community sample being closely monitored. A theoretical assumption can be made based on the literature covered in Chapter 2 that current use of drugs, alcohol, and an increase in symptoms of psychosis may have increased the relationship between cognitive measures and violence risk.

11.4 Implications for Research and Routine Practice

The above findings highlight several implications for research and practice.

State of the Literature. The systematic review and meta-analysis revealed that the literature in this field lacks methodological rigour and continues to be dominated by small samples and poorly defined outcomes which has likely contributed to the heterogeneity and general uncertainty which surrounds the effect sizes. Moreover, aside from a few measures, the small effect sizes and null findings were overwhelming and emphasize the need for researchers to take a step back and to re-examine this literature from the bottom up to improve future studies, starting with clearly defining the basic concepts and reaching a consensus on the neuropsychological measures which are most relevant to violence risk to begin curtailing the heterogeneity, and improving research quality. Although violence and cognitive functioning in violent populations are both heterogenous, it is expected that with more collaborative efforts to increase the quality of research and replicability, that the heterogeneity will decrease enough to allow for more complete conclusions to be drawn.

Low AUC Scores. Characteristics that make good predictability of risk assessment tools include good validity (e.g., the tool measures what it is meant to measure) reliability, especially, inter-rater reliability (e.g., consistency between different raters on the same measure), the risk factors should be empirically and theoretically derived, and should be significantly associated with violence. In addition to this, risk measures should have good discrimination based on sensitivity and 1-specificity, and calibration (e.g., how the predicted risk of the tool matches the observed risk) (Fazel, 2019). With acknowledgment to these key performance indicators and characteristics, it is well known that the HCR-20 does not meet all of these criteria, and nor does the LSI-R, for example, the risk factors in the HCR-20 were based on expert opinion rather than being empirically derived, and the LSI-R does not include some of the strongest predictors for violence and offending, such as age and gender (Fazel, 2019), which may explain the poor predictive ability of these tools seen in the current study. However, upon reflection of the low AUC scores seen in the current study, and the criticisms of low AUC scores in the violence risk literature, it may be helpful to draw parallels between the violence risk literature and the literature on medical models of risk to put them into perspective. For instance, AUC scores for medical models of risk for breast cancer, for example, are low to moderate, with a 2018 paper reporting values ranging from .60 to .70 (Gail & Pfeiffer, 2018), scores which are currently deemed too low when assessing for violence risk. However, the authors highlight that although it would be desirable to have higher values, there are several benefits of AUC scores, as even when they are low, they can still tell some of the story. For instance, risk models can provide individuals with a realistic estimate of the risk, as well as putting the risk into perspective, they can identify important dynamic risk factors which may provide perspectives on reducing risks by avoiding the identified factors (e.g., drinking alcohol and smoking cigarettes), and they can identify certain genes or mutations that can inform patients on preventative care. In addition to this, these models can be used to inform future preventative intervention trials and can identify women who are high-risk for developing cancer to be screened for prevalent breast cancer in an effort to take preventative measures (Gail & Pfeiffer, 2018). Thus, in comparison to the violence risk literature, the medical literature appears to value lower AUC scores in a way that the violence literature does not, such that the medical literature appears to view these scores as a steppingstone to more accurate models, using them to inform treatments and trials, whereas the violence literature views them more as a failure of the assessment and research. As the issue with low predictive accuracy surrounding violence risk assessments has persisted for several years, it may be more productive for the field to shift their view to match that of the medical literature to inform higher quality research to identify risk factors for violence that may have more utility and predictability than those which are being used now.

Deficient Cognitive Functioning and Treatment Benefits. The large proportion of the individuals sampled with borderline and impaired cognitive functions in this thesis is noteworthy as intact cognitive abilities are required for individuals to fully engage in treatment programmes, and research has demonstrated that deficient cognitive abilities account for a significant proportion of the variance in an individual's ability to benefit from treatment programmes (Fishbein et al., 2009; Fishbein & Sheppard, 2006; Green, 1996; O'Reilly et al., 2016; Richter et al., 2018). Thus, it is recommended that comprehensive neuropsychological assessments are routinely conducted in forensic settings to fully understand the cognitive strengths and weaknesses of the offender, to aid in the development of treatment plans that can be individually tailored to the needs of offenders to ensure they receive the necessary support, and their strengths can be leveraged to optimize treatments. This is a notion shared by the Risk-Needs-Responsivity (RNR) model, where *responsivity* encourages the use of cognitive-behavioural treatments, using an approach which is tailored to the offender's ability and learning style (Bonta & Andrews, 2017). Moreover, Cheng and colleagues (2019) advise that the assessment of executive functions prior to interventions locates and quantifies the degree of impairment, which is imperative, as while some interventions such as CBT address the overlap between executive functions and criminogenic needs, clinicians often do not have a clear indication of the degree of impairment. These authors further cautioned that failing to efficiently identify impairments in executive functioning prior to intervention may inhibit treatment progress and deplete resources (Cheng et al., 2019).

Recent high-quality studies have demonstrated encouraging findings for the benefits of focusing on and remediating cognitive impairments. For instance, a 2017 systematic review on cognitive remediation and social cognition training for violence in schizophrenia suggests that therapies which target social cognition and executive functions, positively impacted patients' aggressive attitudes and physical assaults (Darmedru et al., 2017). Similarly, a 2018 single blind randomized controlled trial of cognitive remediation examined the effectiveness of a low support training programme for working memory in patients with psychosis, and findings indicated that patients who completed the training demonstrated significant increases in both neuropsychological and social functioning after a two-week follow-up period, and subsequently at 3–6-month follow-ups (Donohoe et al., 2018). While studies which preceded this one had already demonstrated positive findings to support the association between cognitive remediation and improved neuropsychological and social functioning (e.g., Wykes et al., 2011), this was the first study to demonstrate the efficacy of a low-support treatment. The authors concluded that in outpatients who are mostly stable but have chronic psychosis, the effects found in the study were comparable to the more intense cognitive remediation trainings that require 1:1 support, suggesting it as a viable method for use in patients where psychological support may be lacking. Most recently, O'Reilly and colleagues (2019) conducted a randomized-controlled trial to examine the effectiveness of cognitive remediation training in a population of forensic patients diagnosed with schizophrenia and schizoaffective disorder who had a mean composite score on the MCCB that was three standard deviations below the non-clinical mean, a level of impairment that the authors associated with moderate intellectual disability. The trial findings were encouraging, demonstrating significant improvements in visual and working memory after 8 months, as well as a significant mediation effect between cognitive benefits and improved functioning, and increased movement of patients to lower security units (O'Reilly et al., 2019).

Deficient Cognitive Functioning and Decisional Capacity. The proportion of these samples with cognitive impairments also has implications for researchers recruiting forensic populations. For example, as it is now known that executive dysfunction is overrepresented in forensic populations, the association between executive dysfunction and decisional capacity raises questions on the capacity of individuals with executive impairments to give valid informed consent to participate in research. A 2007 meta-analysis examining the neuropsychological abilities required for decisional capacity found that specific cognitive abilities including executive functioning, and episodic and working memory are important to both the individual and general dimensions of decisional capacity (Palmer & Savla, 2007). Moreover, as the capacity to consent to research or treatment likely requires a range of intact cognitive functions (Marson, 2001), the authors suggested that clinicians and researchers should be aware of the potential for impaired decisional capacity in individuals with cognitive impairments, regardless of the specific form of the impairment. Moreover, there is a significant predictive relationship between severity of cognitive impairments, and impaired capacity to make decisions (Palmer et al., 2012).

Deficient Cognitive Functioning and Fitness to Plead (FtP). Considering the impact that cognitive impairments have on decisional capacity, it would be imprudent not to mention their effect on fitness to stand trial. The Scottish criminal justice system includes processes to make allowance for those defendants who are considered, by reason of mental disorder, to be incapable of participating in a trial. The Criminal Procedure (Scotland) Act 1995 (CP(S) Act) highlights that a person is unfit for trial if it is established that the individual is incapable, by reason of a mental or physical condition, to participate effectively in a trial. The test of whether a person may be 'fit to plead' focuses upon the key questions of whether a defendant is able to understand the nature of the charge against them, the need to enter a plea to a charge, the effects of such a plea, the purpose of a trial, along with their ability to instruct a legal representative, and to follow the course of a trial (Douds & Haut, 2015). Whilst international legal jurisdictions differ regarding specific thresholds and requirements when assessing FtP, they generally relate to an individual's capacity to understand and instruct counsel that allows them to meaningfully participate in the court process (White et al., 2014). However, as cognitive abilities play a crucial role in a defendant's ability to understand, comprehend, and assist counsel (Douds & Haut, 2015), it appears that the only way to ensure that defendants can participate in a court process at a satisfactory level is to correctly assess their cognitive abilities. Studies examining the relationship between cognition and FtP have largely focused on the importance of language, memory, attention and executive functioning (Kirkish & Sreenivasan, 1999; Martell, 1992; Nestor et al., 1999), as well as intelligence level, evidencing that lower intellectual functioning is associated with an increased risk of being unfit to plead (Pirelli et al., 2011). Based on these findings, it is recommended that neuropsychological assessments be considered to identify FtP in at risk member of these populations.

11.5 Limitations

In addition to the limitations of studies presented in individual chapters of this thesis, there exist some limitations that span multiple study chapters.

Generalizability. Methodological limitations which affected the generalizability of findings included small sample sizes, all male samples, and most participants being of white Scottish or white British descent. However, the pilot/feasibility study included both community and inpatient violent offenders, demonstrating findings in two distinct settings and populations, which increased the generalizability of findings. Moreover, given that the aims of the thesis were to examine violence, the results are not generalizable to non-violent populations. The multidimensional use of neuropsychological measures, and the ability for specific domains to pick up other impairments may have also limited generalizability, as similar studies may employ different operational definitions for domains, as was observed in the meta-analysis of this thesis. Similarly, the operational definitions of violence may also inhibit generalizability due to the vast variety used across the literature. However, operational definitions for independent and dependent variables were carefully considered in these studies, and the common limitation of using non-specific definitions for violence were conceded prior to choosing definitions which fit the goals and aims of each study. Though, while this may have increased the specificity of the findings, the use of narrow definitions may have also contributed to the null findings in the thesis. Albeit it is argued that increased specificity in the operationalisation of violent outcomes is more meaningful for the further understanding of the association between cognitive impairment and violence, and while many results were null, it is expected that these studies will contribute to the growth in quality research in the field, rather than adding to the heterogeneity currently observed across similar studies.

Sample Size. In addition to limiting generalizability, the small sample sizes also limited statistical power, and the ability to examine cumulative relationships between all variables to investigate primary and secondary hypotheses. The small sample size may have also played a role in the low base rate of recorded offences for many of the outcomes in this thesis, and while an effort was made to correct this in the final study, it remained a limitation. Consequently, not all measured cognitive domains were analysed in this thesis, leaving the possibility that some may have demonstrated a strong relationship with the outcomes. Moreover, the AUC analyses were likely to be affected by low power, and need to be interpreted cautiously.

Conceptual Limitations. A conceptual limitation of this thesis was the use of latent variables. As cognitive impairments are not observable, neuropsychological assessments measure theoretical variables which are thought to relate to the function of interest. A consequence of this is the potential for task impurity, suggesting that the assessments used in this study may not have been specifically measuring the variables of interest. However, the psychometric properties of the neuropsychological assessments included in this thesis were carefully considered, and all of them have demonstrated sufficient reliability and validity across various samples and settings. In addition, while, arguably, purer tests of functions may be available for the purposes of lab experiments, the researcher chose to use those which are available to clinicians as the eventual clinical utility/replicability of the thesis findings was a priority. Last, the failure to include functional outcomes was also a limitation of this thesis, as the findings could not speak to the potential for the measured cognitive impairments to hinder functional outcomes or treatment benefits.

11.6 Translating Research into Practice

Implementing research into practice is a long process and may largely depend on the needs and priorities of the clinical teams, patients, and programmes. Due to the time and resources required to conduct neuropsychological assessments, there may be little uptake on the recommendations of this thesis to routinely assess offenders alongside risk assessments, particularly in social work or similar community settings. However, this notion presents an opportunity for future research to develop electronically self-administered and self-scored screening batteries to aid in the identification and assessment of cognitive impairment to inform treatment plans, as well as individuals' capacity to make important decisions, such as consenting to research. Based on the findings from this thesis, it is anticipated that a carefully constructed neuropsychological battery that is easy and quick to administer, has clinical utility (i.e., normative data), and is self-scoring, would be considered exceptionally valuable to clinicians and health care workers alike. A battery such as this could aspire to inform professionals quickly, reliably, validly, and efficiently on the location and magnitude of impairments, allowing for objective and informed decisions to be made on ways to best support and treat their patients or clients. However, prior to developing this battery, further research is required to demarcate which domains would prove most useful for improving treatment efficacy, and measures may need to be adapted for use on an electronic interface and tested to ensure they remain reliable and valid, similar to those which already exist such as the Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition, 2019) and Q-Interactive (Pearson, 2015).

11.7 Recommendations

1. In consideration of the inherent heterogeneity in cognitive functioning, violence, and neuropsychological measures, it is recommended that researchers begin replicating high quality studies to further understand the specific sources of heterogeneity, and to determine whether uniformity in findings is possible across this type of research.
2. Based on all the findings in this thesis, it is recommended that researchers begin to shift their focus from examining the utility of neuropsychological measures to improve violence risk assessments, and instead examine the utility of these measures to optimize treatment benefits in violent offenders, which will work toward the same end goal, which is to reduce recidivism. To complete this, it is recommended that the investigation of the cognitive abilities which impede individuals from fully benefitting from treatment take priority in research, by examining their relationship with treatment outcomes.
3. The final recommendation is based on the findings from the pilot and feasibility study and suggests that forensic services implement routine neuropsychological assessments alongside risk assessments to inform risk formulations, tailored treatments of offenders, assessment of decisional capacity, and fitness to plead.

a. To optimize the implementation of routine neuropsychological assessments in services that are lacking the necessary resources and skills, it is recommended that future research focus on building an electronic, self-administered and scored, neuropsychological battery to aid in developing targeted treatment plans.
b. As the proposed electronic battery is yet to be developed, and with acknowledgement that out with clinics and hospitals, routine neuropsychological assessments may not be possible, it is recommended that,
    i. Organizations offer training and educational courses to professionals working with these populations to aid in their understanding of the cognitive deficits commonly seen, and the impact they have on the daily lives of the individuals they are caring for. These courses should also offer exercises and practical skills which can be easily translated to the service-user to aid in their support.
    ii. Practical training and skills programmes should also be made available to offenders to encourage success, such as methods for keeping organized (e.g., time management, scheduling, attending appointments), setting attainable goals, how to work toward achieving goals, coping with symptoms of mental illness and adversities, how to ask for help, and how and where to find resources in the community to further assist them.

11.8 Conclusions

This thesis aimed to examine the potential for neuropsychological measures to improve the predictive accuracy of violence risk assessments in two forensic samples. Although many of the findings failed to support the hypotheses, lack of insight, affect recognition, reasoning, response inhibition, inattention, and response monitoring may merit further investigation in similar populations in relation to violence risk, though, the current findings reveal limited convincing evidence to support the central hypothesis that neuropsychological measures will add significant incremental value to violence risk assessments. However, the findings do not denote that neuropsychological assessments in the presence of risk assessments have no utility, instead it is argued that their use remains vital for informing risk assessments by providing specific treatment targets, identifying cognitive strengths which can be leveraged to optimize treatments, and identifying cognitive weaknesses that can be used to assist in designing interventions that are compatible to individuals offenders. Likewise, the use of these assessments in parallel with risk assessments may add an additional dimension to risk prediction, where rather than directly predicting violence risk, they can aid in the identification of those who are less likely to benefit from standard treatments, and therefore those who may also be more likely to recidivate or require a higher secure placement. In conclusion, for researchers, the way forward is to focus on designing and conducting high quality studies with careful consideration of operational definitions of key variables, to maximize efforts to overcome the obstacles that perpetuate the limitations in this type of research, and to begin focusing on the relationship between cognitive impairments and treatment benefits to reduce recidivism. Equally, practicing clinicians and professionals should implement the routine use of neuropsychological measures to aid in the identification and assessment of cognitive impairments to inform and optimize treatment plans and formulations of risk.

References

Aas, M., Dazzan, P., Fisher, H. L., Morgan, C., Morgan, K., Reichenberg, A., Zanelli, J., Fearon, P., Jones, P. B., Murray, R. M., & Pariante, C. M. (2011). Childhood trauma and cognitive function in first-episode affective and non-affective psychosis. *Schizophrenia Research*, *129*(1), 12–19. https://doi.org/10.1016/j.schres.2011.03.017

Aas, M., Steen, N. E., Agartz, I., Aminoff, S. R., Lorentzen, S., Sundet, K., Andreassen, O. A., & Melle, I. (2012). Is cognitive impairment following early life stress in severe mental disorders based on specific or general cognitive functioning? *Psychiatry Research*, *198*(3), 495–500. https://doi.org/10.1016/j.psychres.2011.12.045

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry*, *13*. https://doi.org/10.1186/1471-244X-13-197

Abu-Akel, A., & Abushua'leh, K. (2004). "Theory of mind" in violent and nonviolent patients with paranoid schizophrenia. *Schizophrenia Research*, *69*(1), 45–53. https://doi.org/10.1016/S0920-9964(03)00049-5

Addington, J., & Addington, D. (1997). Substance Abuse and Cognitive Functioning in Schizophrenia. *Journal of Psychiatry and Neuroscience*, *22*(2), 99–104. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1188831/

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, *11*(2), 231–239. https://doi.org/10.1016/S0959-4388(00)00202-6

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction. *The Counseling Psychologist*, *34*(3), 341–382. https://doi.org/10.1177/0011000005285875

Alia-Klein, N., O'Rourke, T. M., Goldstein, R. Z., & Malaspina, D. (2007). Insight into illness and adherence to psychotropic medications are separately associated with violence severity in a forensic sample. *Aggressive Behavior*, *33*(1), 86–96. https://doi.org/10.1002/ab.20170

Allen, D. N., Goldstein, G., & Aldarondo, F. (1999). Neurocognitive dysfunction in patients diagnosed with schizophrenia and alcoholism. *Neuropsychology*, *13*(1), 62–68. https://doi.org/10.1037/0894-4105.13.1.62

Allen, J., Fonagy, P., & Bateman, A. (2008). *Mentalizing in clinical practice*. https://appi.org/

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Ancín, I., Cabranes, J. A., Santos, J. L., Sánchez-Morla, E., & Barabash, A. (2013). Executive deficits: A continuum schizophrenia-bipolar disorder or specific to schizophrenia? *Journal of Psychiatric Research*, *47*(11), 1564–1571. https://doi.org/10.1016/j.jpsychires.2013.07.008

Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory-Revised*. Multi-Health Systems. https://mhs.com/

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for Effective Rehabilitation. *Criminal Justice and Behavior*, *17*(1), 19–52. https://doi.org/10.1177/0093854890017001004

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The risk-need-responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behavior*, *38*(7), 735–755. https://doi.org/10.1177/0093854811406356

Ang, G. K., & Pridmore, S. (2009). Theory of mind and psychiatry: An introduction. *Australasian Psychiatry*, *17*(2), 117–122. https://doi.org/10.1080/10398560802375982

Arango, C., Barba, A. C., Gonzalez-Salvador, T., & Ordonez, A. C. (1999). Violence in Inpatients With Schizophrenia: A Prospective Study. *Schizophrenia Bulletin*, *25*(3), 493–503. https://doi.org/10.1093/oxfordjournals.schbul.a033396

Arciniegas, D. B., & Wortzel, H. S. (2014). Emotional and Behavioral Dyscontrol After Traumatic Brain Injury. In *Psychiatric Clinics of North America* (Vol. 37, Issue 1, pp. 31–53). https://doi.org/10.1016/j.psc.2013.12.001

Armstrong, J. (2001). *Long-range forecasting*. http://www.forecastingprinciples.com/files/pdf/LR01-jsa.pdf

Aromataris, E., & Munn, Z. (2018). Joanna Briggs Institute Reviewer's Manual. *The Joanna Briggs Institute*, *299*. https://doi.org/https://doi.org/10.46658/JBIMES-20-01

Association of Directors of Social Work. (2018). *Media Briefing: Criminal Justice*. https://socialworkscotland.org

Azouvi, P., Vallat-Azouvi, C., Joseph, P. A., Meulemans, T., Bertola, C., Le Gall, D., Bellmann, A., Roussel, M., Coyette, F., Krier, M., Franconie, C., Bindschadler, C., Diouf, M., & Godefroy, O. (2016). Executive functions deficits after severe traumatic brain injury: The GREFEX study. *Journal of Head Trauma Rehabilitation*, *31*(3), E10–E20. https://doi.org/10.1097/HTR.0000000000000169

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). Statistics in practice How to perform a meta-analysis with R: a practical tutorial. *Evidenced Based Mental Health*, *22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Barrasso-Catanzaro, C., & Eslinger, P. J. (2016). Neurobiological Bases of Executive Function and Social-Emotional Development: Typical and Atypical Brain Changes. *Family Relations*, *65*(1), 108–119. https://doi.org/10.1111/fare.12175

Bass, S., & Nussbaum, D. (2010). Decision making and aggression in forensic psychiatric inpatients. *Criminal Justice and Behavior*, *37*(4), 365–383. https://doi.org/10.1177/0093854809360043

Batty, R., Rossell, S., Francis, A. J., & Ponsford, J. (2013). Psychosis following traumatic brain injury. *Brain Impairment*, *14*(1), 21. https://doi.org/https://doi.org/0.1017/BrImp.2013.10

Baumeister, R. F., & Heatherton, T. F. (1996). Self-Regulation Failure: An Overview. *Psychological Inquiry*, *7*(1), 1–15. https://doi.org/10.1207/s15327965pli0701_1

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. In *Nature Neuroscience* (Vol. 8, Issue 11, pp. 1458–1463). https://doi.org/10.1038/nn1584

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Beers, S. R., & De Bellis, M. D. (2002). Neuropsychological function in children with maltreatment-related posttraumatic stress disorder. *American Journal of Psychiatry*, *159*(3), 483–486. https://doi.org/10.1176/appi.ajp.159.3.483

Beggs, S. M., & Grace, R. C. (2008). Psychopathy, intelligence, and recidivism in child molesters: Evidence of an Interaction Effect. *Journals.Sagepub.Com*, *35*(6), 683–695.

https://doi.org/10.1177/0093854808314786

Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: A prospective study in two maximum-security correctional institutions. *Journal of Forensic Psychiatry*, *11*(1), 167–175. https://doi.org/10.1080/095851800362445

Bennett, T., Holloway, K., & Farrington, D. (2008). The statistical association between drug misuse and crime: A meta-analysis. *Aggression and Violent Behavior*, *13*(2), 107–118. https://doi.org/10.1016/j.avb.2008.02.001

Bjork, J. M., & Grant, S. J. (2009). Does traumatic brain injury increase risk for substance abuse? *Journal of Neurotrauma*, *26*(7), 1077–1082. https://doi.org/10.1089/neu.2008.0849

Bjørkly, S. (1995). Prediction of aggression in psychiatric patients: A review of prospective prediction studies. *Clinical Psychology Review*, *15*(6), 475–502. https://doi.org/10.1016/0272-7358(95)00016-I

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, *14*(4), 698–718. https://doi.org/10.1016/j.concog.2005.06.004

Bock, E. M., & Hosser, D. (2014). Empathy as a predictor of recidivism among young adult offenders. *Psychology, Crime and Law*, *20*(2), 101–115. https://doi.org/10.1080/1068316X.2012.749472

Boer, D. R., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). Manual for the Sexual Violence Risk - 20. In *Vancouver BC The British Columbia Institute Against Family Violence*. https://www.bcifv.org/

Bogner, J., & Corrigan, J. D. (2009). Reliability and Predictive Validity of the Ohio State University TBI Identification Method With Prisoners. *Journal of Head Trauma Rehabilitation*, *24*(4), 279–291. https://doi.org/10.1097/HTR.0b013e3181a66356

Bonta, J., & Andrews, D. A. (2017). *The Psychology of Criminal Conduct - 6th Edition*. Routledge. https://www.routledge.com/The-Psychology-of-Criminal-Conduct/Bonta-Andrews/p/book/9781138935778

Bora, E., Yucel, M., & Pantelis, C. (2009). Theory of mind impairment in schizophrenia: Meta-analysis. In *Schizophrenia Research* (Vol. 109, Issues 1–3, pp. 1–9). Elsevier. https://doi.org/10.1016/j.schres.2008.12.020

Bora, E., Yücel, M., & Pantelis, C. (2010). Neurocognitive markers of psychosis in bipolar disorder: A meta-analytic study. *Journal of Affective Disorders*, *127*(1–3), 1–9. https://doi.org/10.1016/j.jad.2010.02.117

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Converting Among Effect Sizes. In *Introduction to Meta-Analysis* (pp. 45–49). John Wiley & Sons. https://doi.org/10.1002/9780470743386.ch7

Borum, R., Bartel, P., & Forth, A. (2006). *SAVRY: Structured assessment of violence risk in youth*. Psychological Assessment Resources. https://www.parinc.com/Products/Pkey/390

Braga, A. A., Weisburd, D., & Turchan, B. (2018). Focused Deterrence Strategies and Crime Control: An Updated Systematic Review and Meta-Analysis of the Empirical Evidence. *Criminology and Public Policy*, *17*(1), 205–250. https://doi.org/10.1111/1745-9133.12353

Brand, M., Recknor, E. C., Grabenhorst, F., & Bechara, A. (2007). Decisions under ambiguity and decisions under risk: Correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *Journal of Clinical and Experimental Neuropsychology*, *29*(1), 86–99. https://doi.org/10.1080/13803390500507196

Brimigion, R. P. (2014). Cognitive differences between violent and non-violent adolescent male offenders as a variable of violence risk [Massachusetts School of Professional Psychology]. In *ProQuest Dissertations and Theses*. http://search.proquest.com/openview/e13ad3c70b44220da1dced51eace00e5/1?pq-origsite=gscholar&cbl=18750&diss=y&casa_token=_3xDzY7pXHcAAAAA:zZ_RPV_CTXwaePulwB2oEPaNh86DjA9Reg97CEgOW5_T7MA35z2XB6wQeK5j9iaLzLSkSCaibQ

Brower, M. C., & Price, B. H. (2001). Neuropsychiatry of frontal lobe dysfunction in violent and criminal behaviour: a critical review. *Journal of Neurology, Neurosurgery & Psychiatry*, *71*(6), 7. https://doi.org/10.1136/jnnp.71.6.720

Brown, S., O’Rourke, S., & Schwannauer, M. (2019). Risk factors for inpatient violence and self-harm in forensic psychiatry: the role of head injury, schizophrenia and substance misuse. *Brain Injury*, *33*(3), 313–321. https://doi.org/10.1080/02699052.2018.1553064

Brugman, S., Lobbestael, J., von Borries, A. K. L., Bulten, B. E. H., Cima, M., Schuhmann, T., Dambacher, F., Sack, A. T., & Arntz, A. (2016). Cognitive predictors of violent incidents in forensic psychiatric inpatients. *Psychiatry Research*, *237*, 229–237. https://doi.org/10.1016/j.psychres.2016.01.035

Brüne, M. (2005). Emotion recognition, “theory of mind,” and social behavior in schizophrenia. *Psychiatry Research*, *133*(2–3), 135–147. https://doi.org/10.1016/j.psychres.2004.10.007

Bryant, E. T., Scott, M. L., Golden, C. J., & Tori, C. D. (1984). Neuropsychological deficits, learning disability, and violent behavior. *Journal of Consulting and Clinical Psychology*, *52*(2), 323–324. https://doi.org/10.1037//0022-006x.52.2.323

Buchanan, A. (2008). Risk of violence by psychiatric patients: beyond the “actuarial versus clinical” assessment debate. In *Psychiatric Services* (Vol. 59, Issue 2, pp. 184–190). American Psychiatric Association. https://doi.org/10.1176/ps.2008.59.2.184

Buffington-Vollum, J., Edens, J. F., Johnson, D. W., & Johnson, J. K. (2002). Psychopathy as a predictor of institutional misbehavior among sex offenders: A prospective replication. *Criminal Justice and Behavior*, *29*(5), 497–511. https://doi.org/10.1177/009385402236730

Burgess, P. W. (2004). Theory and methodology in executive function research. In *Methodology of Frontal and Executive Function* (pp. 79–113). Routledge. https://doi.org/10.4324/9780203344187-8

Busch, K. G., Zagar, R., Hughes, J. R., Arbit, J., & Bussell, R. E. (1990). Adolescents who kill. *Journal of Clinical Psychology*, *46*(4), 472–485. https://doi.org/10.1002/1097-4679(199007)46:4<472::AID-JCLP2270460416>3.0.CO;2-F

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. https://doi.org/10.1038/nrn3475

Caggiano, A. D. (2000). Identifying Violent-Toward-Staff Juvenile Delinquents via the Millon Adolescent Clinical Inventory and Neuropsychological Measures. *Journal of Offender Rehabilitation*, *32*(1–2), 1–2. https://doi.org/10.1300/J076v32n01_06

Cagigas, X. E., & Manly, J. J. (2014). Cultural neuropsychology: The new norm. In *Clinical neuropsychology: A pocket handbook for assessment (3rd ed.).* (pp. 132–156). American Psychological Association. https://doi.org/10.1037/14339-008

Cahill, B. S., Coolidge, F. L., Segal, D. L., Klebe, K. J., Marle, P. D., & Overmann, K. A. (2012). Prevalence of ADHD and Its Subtypes in Male and Female Adult Prison Inmates. *Behavioral Sciences & the Law*, *30*(2), 154–166. https://doi.org/10.1002/bsl.2004

Cambridge Cognition. (2019). *CANTAB® [Cognitive assessment software]*. www.cantab.com

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105. https://doi.org/10.1037/h0046016

Campbell, M. A., French, S. A., & Gendreau, P. (2007). *Assessing the utility of risk assessment tools and personality measures in the prediction of violent revidivism for adult offenders*. Public Safety Canada. https://www.securitepublique.gc.ca/cnt/rsrcs/pblctns/rsk-ssmnt-tls/rsk-ssmnt-tls-eng.pdf

Cardinal, R. N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Networks*, *19*(8), 1277–1301. https://doi.org/10.1016/j.neunet.2006.03.004

Carroll, J. B. (1993). Human Cognitive Abilities. In *Human Cognitive Abilities*. Cambridge University Press. https://doi.org/10.1017/cbo9780511571312

Chan, H. C. O., & Chui, W. H. (2012). Psychological Correlates of Violent and Non-violent Hong Kong Juvenile Probationers. *Behavioral Sciences & the Law*, *30*(2), 103–120. https://doi.org/10.1002/bsl.2003

Chandler, R. K., Fletcher, B. W., & Volkow, N. D. (2009). Treating drug abuse and addiction in the criminal justice system: Improving public health and safety. In *JAMA - Journal of the American Medical Association* (Vol. 301, Issue 2, pp. 183–190). American Medical Association. https://doi.org/10.1001/jama.2008.976

Chanraud, S., Martelli, C., Delain, F., Kostogianni, N., Douaud, G., Aubin, H. J., Reynaud, M., & Martinot, J. L. (2007). Brain morphometry and cognitive performance in detoxified alcohol-dependents with preserved psychosocial functioning. *Neuropsychopharmacology*, *32*(2), 429–438. https://doi.org/10.1038/sj.npp.1301219

Chen, H., Cohen, P., & Chen, S. (2010). Communications in Statistics-Simulation and Computation® How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Taylor & Francis*, *39*(4), 860–864. https://doi.org/10.1080/03610911003650383

Cheng, J., O'Connell, M. E., & Wormith, J. S. (2019). Bridging Neuropsychology and Forensic Psychology: Executive Function Overlaps With the Central Eight Risk and Need Factors. *International Journal of Offender Therapy and Comparative Criminology*, *63*(4), 558–573. https://doi.org/10.1177/0306624X18803818

Cherpitel, C. J. (2007). Alcohol and injuries: A review of international emergency room studies since 1995. In *Drug and Alcohol Review* (Vol. 26, Issue 2, pp. 201–214). https://doi.org/10.1080/09595230601146686

Chow, W. S., & Priebe, S. (2013). Understanding psychiatric institutionalization: A conceptual review. *BMC Psychiatry*, *13*(1), 169. https://doi.org/10.1186/1471-244X-13-169

Cochran, W. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101–129. https://onlinelibrary.wiley.com/journal/15410420

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. https://doi.org/https://doi.org/10.4324/9780203771587

Coid, J., Kallis, C., Doyle, M., Shaw, J., & Ullrich, S. (2015). Identifying causal risk factors for violence among discharged patients. *PLoS ONE*, *10*(11). https://doi.org/10.1371/journal.pone.0142493

Coid, J., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Farrington, D., & Rogers, R. (2011). Most items in structured risk assessment instruments do not predict violence. *Journal of Forensic Psychiatry and Psychology*, *22*(1), 3–21. https://doi.org/10.1080/14789949.2010.495990

Coid, J., Yang, M., Ullrich, S., Zhang, T., Sizmur, S., Roberts, C., Farrington, D. P., & Rogers, R. D. (2009). Gender Differences in Structured Risk Assessment: Comparing the Accuracy of Five Instruments. *Journal of Consulting and Clinical Psychology*, *77*(2), 337–348. https://doi.org/10.1037/a0015155

Conners, C. (2015). *Connors Continuous Performance Test*. Multi-Health Systems. https://mhs.com/

Cooper, L., Liberman, D., Tucker, D., Nuechterlein, K. H., Tsuang, J., & Barnett, H. L. (1999). Neurocognitive Deficits in the Dually Diagnosed With Schizophrenia and Cocaine Abuse. *Psychiatric Rehabilitation Skills*, *3*(2), 231–245. https://doi.org/10.1080/10973439908408386

Cornell, D. G., & Wilson, L. A. (1992). The PIQ & VIQ discrepancy in violent and nonviolent delinquents. *Journal of Clinical Psychology*, *48*(2), 256–261. https://doi.org/10.1002/1097-4679(199203)48:2<256::AID-JCLP2270480219>3.0.CO;2-4

Corrigan, J. D., & Bogner, J. (2007). Screening and Identification of TBI. *Journal of Head Trauma Rehabilitation*, *22*(6), 315–317.

https://doi.org/10.1097/01.HTR.0000300226.67748.3e

Corrigan, J. D., Bogner, J., & Holloman, C. (2012). Lifetime history of traumatic brain injury among persons with substance use disorders. *Brain Injury*, *26*(2), 139–150. https://doi.org/10.3109/02699052.2011.648705

Corrigan, J. D., Bogner, J., Mellick, D., Bushnik, T., Dams-O'Connor, K., Hammond, F. M., Hart, T., & Kolakowsky-Hayner, S. (2013). Prior history of traumatic brain injury among persons in the Traumatic Brain Injury Model Systems National Database. *Archives of Physical Medicine and Rehabilitation*, *94*(10), 1940–1950. https://doi.org/10.1016/j.apmr.2013.05.018

Crews, F. T., & Boettiger, C. A. (2009). Impulsivity, frontal lobes and risk for addiction. *Pharmacology Biochemistry and Behavior*, *93*(3), 237–247. https://doi.org/10.1016/j.pbb.2009.04.018

Cuffel, B. J., Shumway, M., Chouljian, T. L., & Macdonald, T. (1994). A longitudinal study of substance use and community violence in schizophrenia. *Journal of Nervous and Mental Disease*, *182*(12), 704–708. https://doi.org/10.1097/00005053-199412000-00005

D., C. P., Harris, R., & Lamb, R. (1984). The Encyclopedic Dictionary of Psychology. *The American Journal of Psychology*, *97*(4), 628. https://doi.org/10.2307/1422179

Daffern, M., Mayer, M. M., & Martin, T. (2003). A preliminary investigation into patterns of aggression in an Australian forensic psychiatric hospital. *Journal of Forensic Psychiatry and Psychology*, *14*(1), 67–84. https://doi.org/10.1080/1478994031000074306

Dagnan, D., Chadwick, P., & Proudlove, J. (2000). Toward an assessment of suitability of people with mental retardation for cognitive therapy. *Cognitive Therapy and Research*, *24*(6), 627–636. https://doi.org/10.1023/A:1005531226519

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and human brain*. Penguin Books. https://www.penguin.co.uk/

Daniel, B., & Scott, J. (2018). *50 years of social work 1968-2018*. https://socialworkscotland.org//wp-content/uploads/2018/09/August-2018-50th-Anniversary-report-FINAL.pdf

Darmedru, C., Demily, C., & Franck, N. (2017). Cognitive remediation and social cognitive training for violence in schizophrenia: a systematic review. *Psychiatry Research*, *251*, 266–274. https://doi.org/10.1016/j.psychres.2016.12.062

De Bellis, M. D., Hooper, S. R., Spratt, E. G., & Woolley, D. P. (2009). Neuropsychological findings in childhood neglect and their relationships to pediatric PTSD. *Journal of the International Neuropsychological Society*, *15*(6), 868–878. https://doi.org/10.1017/S1355617709990464

De Vogel, V., & De Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime and Law*, *12*(3), 321–336. https://doi.org/10.1080/10683160600569029

De Vogel, V., De Ruiter, C., Van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior*, *28*(3), 235–251. https://doi.org/10.1023/B:LAHU.0000029137.41974.eb

De Vogel, V., de Vries Robbé, M., de Ruiter, C., & Bouman, Y. H. A. (2011). Assessing protective factors in forensic psychiatric practice: Introducing the SAPROF. *International Journal of Forensic Mental Health*, *10*(3), 171–177. https://doi.org/10.1080/14999013.2011.600230

Dejong, J., Virkkunen, M., & Linnoila, M. (1992). Factors associated with recidivism in a criminal population. *Journal of Nervous and Mental Disease*, *180*(9), 543–550. https://doi.org/10.1097/00005053-199209000-00001

Denson, T. F. (2021). Breaking the Cycle of Violent Crime and Punishment: The Promise of Neuronormalization. *Social Issues and Policy Review*, *15*(1), 237–276. https://doi.org/10.1111/sipr.12076

Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in U.S. correctional settings. *Psychological Services*, *13*(3), 206–222. https://doi.org/10.1037/ser0000075

Dewall, C. N., Anderson, C. A., & Bushman, B. J. (2011). The General Aggression Model: Theoretical Extensions to Violence. *Psychology of Violence*, *1*(3), 245. https://doi.org/10.1037/a0023842

Dinn, W. M., Gansler, D. A., Shattuck Hospital, L., & Fulwiler, C. (2009). Brain Dysfunction and Community Violence in Patients with Major Mental Illness MOCZYNSKI. *Criminal Justice and Behavior*, *36*(2), 117–136. https://doi.org/10.1177/0093854808327507

Dolan, M. (2012). The neuropsychology of prefrontal function in antisocial personality disordered offenders with varying degrees of psychopathy. *Psychological Medicine*, *42*(8), 1715–1725. https://doi.org/10.1017/S0033291711002686

Dolan, M., & Doyle, M. (2000). Violence risk prediction: Clinical and actuarial measures and the role of the Psychopathy Checklist. In *British Journal of Psychiatry* (Vol. 177, Issue OCT., pp. 303–311). Cambridge University Press. https://doi.org/10.1192/bjp.177.4.303

Dolan, M., & Fullam, R. (2006). Face affect recognition deficits in personality-disordered offenders: Association with psychopathy. *Psychological Medicine*, *36*(11), 1563–1569. https://doi.org/10.1017/S0033291706008634

Dolan, M., & Park, I. (2002). The neuropsychology of antisocial personality disorder. *Psychological Medicine*, *32*(3), 417–427. https://doi.org/10.1017/S0033291702005378

Donohoe, G., Dillon, R., Hargreaves, A., Mothersill, O., Castorina, M., Furey, E., Fagan, A. J., Meaney, J. F., Fitzmaurice, B., Hallahan, B., McDonald, C., Wykes, T., Corvin, A., & Robertson, I. H. (2018). Effectiveness of a low support, remotely accessible, cognitive remediation training programme for chronic psychosis: Cognitive, functional and cortical outcomes from a single blind randomised controlled trial. *Psychological Medicine*, *48*(5), 751–764. https://doi.org/10.1017/S0033291717001982

Douds, F., & Haut, F. (2015). Evolution and devolution: Scottish legislation relevant to people with intellectual disabilities within mental health and forensic services. *Journal of Intellectual Disabilities and Offending Behaviour*, *6*(3–4), 127–136. https://doi.org/10.1108/JIDOB-08-2015-0017

Douglas, K. S., Cox, D. N., & Webster, C. D. (1999). Violence risk assessment: Science and practice. In *Legal and Criminological Psychology* (Vol. 4, Issue 2, pp. 149–184). John Wiley & Sons, Ltd. https://doi.org/10.1348/135532599167824

Douglas, K. S., Guy, L. S., & Hart, S. D. (2009). Psychosis as a Risk Factor for Violence to Others: A Meta-Analysis. *Psychological Bulletin*, *135*(5), 679–706. https://doi.org/10.1037/a0016311

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). HCR-20 (Version 3): Assessing risk for violence. *Mental Health Law and Policy Institute, Simon Fraser University*. http://hcr-20.com/

Douglas, K. S., Hart, S. D., Webster, C. D., Belfrage, H., Guy, L. S., & Wilson, C. M. (2014). Historical-Clinical-Risk Management-20, Version 3 (HCR-20 V3): Development and

Overview. *International Journal of Forensic Mental Health*, *13*(2), 93–108. https://doi.org/10.1080/14999013.2014.906519

Douglas, K. S., & Ogloff, J. R. (2003). Violence by Psychiatric Patients: The Impact of Archival Measurement Source on Violence Base Rates and Risk Assessment Accuracy. *Canadian Journal of Psychiatry*, *48*(11), 734–740. https://doi.org/10.1177/070674370304801105

Douglas, K. S., Ogloff, J. R. P., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services*, *54*(10), 1372–1379. https://doi.org/10.1176/appi.ps.54.10.1372

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology*, *67*(6), 917–930. https://doi.org/10.1037//0022-006x.67.6.917

Douglas, K. S., Yeomans, M., & Boer, D. P. (2005). Comparative Validity Analysis of Multiple Measures of Violence Risk in a Sample of Criminal Offenders. *Criminal Justice and Behavior*, *32*(5), 479–510. https://doi.org/10.1177/0093854805278411

Doyle, M., & Dolan, M. (2006). Predicting community violence from patients discharged from mental health services. *British Journal of Psychiatry*, *189*(DEC.), 520–526. https://doi.org/10.1192/bjp.bp.105.021204

Doyle, M., Dolan, M., & McGovern, J. (2002). The validity of North American risk assessment tools in predicting in-patient violent behaviour in England. *Legal and Criminological Psychology*, *7*(2), 141–154. https://doi.org/10.1348/135532502760274756

Doyle, Michael, & Logan, C. (2012). Operationalizing the Assessment and Management of Violence Risk in the Short-term. *Behavioral Sciences & the Law*, *30*(4), 406–419. https://doi.org/10.1002/bsl.2017

Duggan, E. C., & Garcia-Barrera, M. A. (2015). Executive functioning and intelligence. In *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts* (pp. 435–458). Springer New York. https://doi.org/10.1007/978-1-4939-1562-0_27

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, *33*(3), 261–268. https://doi.org/10.1016/0028-3932(94)00124-8

Duwors, R. M. C. (1998). An exploratory study of the relationship between neuropsychological deficit and selected personality disorders, impulsivity and violence. In *Dissertation Abstracts International: Section B: The Sciences and Engineering* (Vol. 58, Issues 8-B, p. 4442). https://elibrary.ru/item.asp?id=5521094

Edens, J. F., Skeem, J. L., & Douglas, K. S. (2006). Incremental validity analyses of the violence risk appraisal guide and the psychopathy checklist: Screening version in a civil psychiatric sample. *Assessment*, *13*(3), 368–374. https://doi.org/10.1177/1073191105284001

Edwards, D. W., Scott, C. L., Yarvis, R. M., Paizis, C. L., & Panizzon, M. S. (2003). Impulsiveness, impulsive aggression, personality disorder, and spousal violence. *Violence and Victims*, *18*(1), 3–14. https://doi.org/10.1891/vivi.2003.18.1.3

Ekinci, O., & Ekinci, A. (2013). Association between insight, cognitive insight, positive symptoms and violence in patients with schizophrenia. *Nordic Journal of Psychiatry*, *67*(2), 116–123. https://doi.org/10.3109/08039488.2012.687767

Ellis, P. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press. www.cambridge.org/9780521142465

Etain, B., Mathieu, F., Henry, C., Raust, A., Roy, I., Germain, A., Leboyer, M., & Bellivier, F. (2010). Preferential association between childhood emotional abuse and bipolar disorder. *Journal of Traumatic Stress*, *23*(3), 376–383. https://doi.org/10.1002/jts.20532

Farewell, V. T., Long, D. L., Tom, B. D. M., Yiu, S., & Su, L. (2017). Two-Part and Related Regression Models for Longitudinal Data. *Annual Review of Statistics and Its Application*, *4*(1), 283–315. https://doi.org/10.1146/annurev-statistics-060116-054131

Farrington, D., Loeber, R., Stallings, R., & Homish, D. (2012). Early risk factors for young homicide offenders and victims. In M. DeLisi & P. Conis (Eds.), *Violent offenders: Theory, research, policy, and practice* (pp. 143–160). Jones & Bartlett Learning. https://www.jblearning.com/

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fazel, S. (2019). The Scientific Validity of Current Approaches to Violence and Criminal Risk

Assessment. In J. W. de Keijser, J. V. Roberts, & J. Ryberg (Eds.), *Predictive Sentencing* (first, pp. 197–212). Hart. https://doi.org/10.5040/9781509921447.ch-011

Fazel, S., Bains, P., & Doll, H. (2006). Substance abuse and dependence in prisoners: A systematic review. In *Addiction* (Vol. 101, Issue 2, pp. 181–191). https://doi.org/10.1111/j.1360-0443.2006.01316.x

Fazel, S., Gulati, G., Linsell, L., Geddes, J. R., & Grann, M. (2009). Schizophrenia and Violence: Systematic Review and Meta-Analysis. *PLoS Medicine*, *6*(8), e1000120. https://doi.org/10.1371/journal.pmed.1000120

Fazel, S., Langström, N., Hjern, A., Grann, M., & Lichtenstein, P. (2009). Schizophrenia, substance abuse, and violent crime. *JAMA - Journal of the American Medical Association*, *301*(19), 2016–2023. https://doi.org/10.1001/jama.2009.675

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *BMJ (Online)*, *345*(7868). https://doi.org/10.1136/bmj.e4692

Fefer, J. P., De-Urioste Stone, S., Daigle, J., & Silka, L. (2016). Using the Delphi Technique to Identify Key Elements for Effective and Sustainable Visitor Use Planning Frameworks. *SAGE Open*, *6*(2). https://doi.org/10.1177/2158244016643141

Feichtinger, D. (2007). *Cognitive flexibility in a forensic population* [Pacific University]. https://commons.pacificu.edu/work/a5e67390-c674-4a8f-8722-36fd3e5329a0

Fernández-Serrano, M. J., Pérez-García, M., & Verdejo-García, A. (2011). What are the specific vs. generalized effects of drugs of abuse on neuropsychological performance? In *Neuroscience and Biobehavioral Reviews* (Vol. 35, Issue 3, pp. 377–406). Pergamon. https://doi.org/10.1016/j.neubiorev.2010.04.008

Field, A., Miles, J., & Field, Z. (2012). Discovering statistics using R. *Choice Reviews Online*, *50*(04), 50-2114-50–2114. https://doi.org/10.5860/choice.50-2114

Fishbein, D. (2000). Neuropsychological function, drug abuse, and violence: A conceptual framework. *Criminal Justice and Behavior*, *27*(2), 139–159. https://doi.org/10.1177/0093854800027002001

Fishbein, D., Krupitsky, E., Flannery, B. A., Langevin, D. J., Bobashev, G., Verbitskaya, E., Augustine, C. B., Bolla, K. I., Zvartau, E., Schech, B., Egorova, V., Bushara, N., & Tsoy,

M. (2007). Neurocognitive characterizations of Russian heroin addicts without a significant history of other drug use. *Drug and Alcohol Dependence*, *90*(1), 25–38. https://doi.org/10.1016/j.drugalcdep.2007.02.015

Fishbein, D., & Sheppard, M. (2006). Assessing the Role of Neuropsychological Functioning in Inmates' Treatment Response. *Pdfs.Semanticscholar.Org*, 1–104. https://pdfs.semanticscholar.org/1826/2c89bc9f9ed4cb5624ff0d209e15c9224e3f.pdf

Fishbein, D., Sheppard, M., Hyde, C., Hubal, R., Newlin, D., Serin, R., Chrousos, G., & Alesci, S. (2009). Deficits in behavioral inhibition predict treatment engagement in prison inmates. *Law and Human Behavior*, *33*(5), 419–435. https://doi.org/10.1007/s10979-008-9163-7

Fisher, H. L., Craig, T. K., Fearon, P., Morgan, K., Dazzan, P., Lappin, J., Hutchinson, G., Doody, G. A., Jones, P. B., McGuffin, P., Murray, R. M., Leff, J., & Morgan, C. (2011). Reliability and comparability of psychosis patients' retrospective reports of childhood abuse. *Schizophrenia Bulletin*, *37*(3), 546–553. https://doi.org/10.1093/schbul/sbp103

Fisher, H. L., Jones, P. B., Fearon, P., Craig, T. K., Dazzan, P., Morgan, K., Hutchinson, G., Doody, G. A., McGuffin, P., Leff, J., Murray, R. M., & Morgan, C. (2010). The varying impact of type, timing and frequency of exposure to childhood adversity on its association with adult psychotic disorder. *Psychological Medicine*, *40*(12), 1967–1978. https://doi.org/10.1017/S0033291710000231

Fisher, K. (2016). Inpatient Violence. In *Psychiatric Clinics of North America* (Vol. 39, Issue 4, pp. 567–577). W.B. Saunders. https://doi.org/10.1016/j.psc.2016.07.005

Frith, C. D. (2004). Schizophrenia and theory of mind. *Psychological Medicine*, *34*(3), 385–389. https://doi.org/10.1017/S0033291703001326

Fujii, D., & Ahmed, I. (2002). Characteristics of Psychotic Disorder Due to Traumatic Brain Injury: An analysis of case studies in the literature. *Journal of Neuropsychiatry and Clinical Neurosciences*, *14*(2), 130–140. https://doi.org/10.1176/jnp.14.2.130

Fujii, D., Ahmed, I., & Hishinuma, E. (2004). A neuropsychological comparison of psychotic disorder following traumatic brain injury, traumatic brain injury without psychotic disorder, and schizophrenia. *Journal of Neuropsychiatry and Clinical Neurosciences*, *16*(3), 306–314. https://doi.org/10.1176/jnp.16.3.306

Fullam, R. S., & Dolan, M. (2008). Executive function and in-patient violence in forensic patients with schizophrenia. *British Journal of Psychiatry*, *193*(3), 247–253. https://doi.org/10.1192/bjp.bp.107.040345

Gail, M., & Pfeiffer, R. (2018). Breast cancer risk model requirements for counseling, prevention, and screening. *Journal of the National Cancer Institute*, *110*(9), 994–1002. https://academic.oup.com/jnci/article-abstract/110/9/994/4912416

Gamer, M., Lemon, J., & Singh, I. F. P. (2012). *CRAN - Package irr* (0.84). Cran. http://www.r-project.org

García-Crespo, Á., Colomo-Palacios, R., Soto-Acosta, P., & Ruano-Mayoral, M. (2010). Information Systems Management A Qualitative Study of Hard Decision Making in Managing Global Software Development Teams A Qualitative Study of Hard Decision Making in Managing Global Software Development Teams. *Information Systems Management*, *27*(3), 247–252. https://doi.org/10.1080/10580530.2010.493839

Garrington, C., & Boer, D. P. (2020). Structured Professional Judgement in Violence Risk Assessment. In J. Worminth, L. Craig, & T. Hogue (Eds.), *The Wiley Handbook of What Works in Violence Risk Management* (pp. 145–162). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119315933.ch7

Geen, R. (1998). *Human Aggression: Theories, Research, and Implications for Social Policy* (R. Geen & E. Donnerstein (eds.); 2nd Editio). Academic Press. https://www.elsevier.com/books-and-journals/academic-press

Geier, C. F. (2013). Adolescent cognitive control and reward processing: Implications for risk taking and substance use. *Hormones and Behavior*, *64*(2), 333–342. https://doi.org/10.1016/j.yhbeh.2013.02.008

Glover, A. J. J., Nicholson, D. E., Hemmati, T., Bernfeld, G. A., & Quinsey, V. L. (2002). A comparison of predictors of general and violent recidivism among high-risk federal offenders. *Criminal Justice and Behavior*, *29*(3), 235–249. https://doi.org/10.1177/0093854802029003001

Goldberg, T. E., & Gold, J. M. (1995). Neurocognitive functioning in patients with schizophrenia: an overview. *Psychopharmacology, the Fourth Generation of Progress. New York, NY: Raven Press Ltd*, 1245–1257. https://www.academia.edu/download/30304806/ch48_657-670.pdf

Golden, C. J. (1978). Stroop Color and Word Test: A manual for clinical and experimental uses.

Chicago: Stoelting, 1–46. https://nsuworks.nova.edu/cps_facbooks/47/

Golden, C. J., Jackson, M. L., Peterson-Rohne, A., & Gontkovsky, S. T. (1996). Neuropsychological correlates of violence and aggression: A review of the clinical literature. *Aggression and Violent Behavior*, *1*(1), 3–25. https://doi.org/10.1016/1359-1789(95)00002-X

Goldstein, H., & Higgins-D'allessandro, A. (2001). Empathy and attachment in relation to violent vs. non-violent offense history among jail inmates. *Journal of Offender Rehabilitation*, *32*(4), 31–53. https://doi.org/10.1300/J076v32n04_03

Goldstein, P. J. (1985). The drugs/violence nexus: A tripartite conceptual framework. *Journal of Drug Issues*, *15*(4), 493–506. https://doi.org/10.1177/002204268501500406

Goldstein, R. Z., Leskovjan, A. C., Hoff, A. L., Hitzemann, R., Bashan, F., Khalsa, S. S., Wang, G. J., Fowler, J. S., & Volkow, N. D. (2004). Severity of neuropsychological impairment in cocaine and alcohol addiction: Association with metabolism in the prefrontal cortex. *Neuropsychologia*, *42*(11), 1447–1458. https://doi.org/10.1016/j.neuropsychologia.2004.04.002

Grafman, J., Schwab, K., Warden, D., Pridgen, A., Brown, H. R., & Salazar, A. M. (1996). Frontal lobe injuries, violence, and aggression: A report of the Vietnam head injury study. *Neurology*, *46*(5), 1231–1238. https://doi.org/10.1212/wnl.46.5.1231

Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial Assessment of Risk for Violence. *Criminal Justice and Behavior*, *27*(1), 97–114. https://doi.org/10.1177/0093854800027001006

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404–411. https://doi.org/10.1037/h0059831

Gray, N. S., Hill, C., McGleish, A., Timmons, D., MacCulloch, M. J., & Snowden, R. J. (2003). Prediction of violence and self-harm in mentally disordered offenders: A prospective study of the efficacy of HCR-20, PCL-R, and psychiatric symptomatology. *Journal of Consulting and Clinical Psychology*, *71*(3), 443–451. https://doi.org/10.1037/0022-006X.71.3.443

Green, M. F. (1996). What are the functional consequences of neurocognitive deficits in schizophrenia? In *American Journal of Psychiatry* (Vol. 153, Issue 3, pp. 321–330).

American Psychiatric Association. https://doi.org/10.1176/ajp.153.3.321

Green, M. F., Horan, W. P., & Lee, J. (2019). Nonsocial and social cognition in schizophrenia: current evidence and future directions. In *World Psychiatry* (Vol. 18, Issue 2, pp. 146–161). Blackwell Publishing Ltd. https://doi.org/10.1002/wps.20624

Greenfield, R., & Valliant, P. M. (2007). Moral reasoning, executive function, and personality in violent and nonviolent adult offenders. *Psychological Reports*, *101*(1), 323–333. https://doi.org/10.2466/PR0.101.1.323-333

Gretton, H. (1998). *Psychopathy and recidivism in adolescence: A ten-year retrospective follow-up* [University of British Columbia]. https://open.library.ubc.ca/collections/831/831/items/1.0088959

Grevatt, M., Thomas-Peter, B., & Hughes, G. (2004). Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient violence? In *Journal of Forensic Psychiatry and Psychology* (Vol. 15, Issue 2, pp. 278–292). https://doi.org/10.1080/1478994032000199095

Griffith, J. J., Daffern, M., & Godber, T. (2013). Examination of the predictive validity of the Dynamic Appraisal of Situational Aggression in two mental health units. *International Journal of Mental Health Nursing*, *22*(6), 485–492. https://doi.org/10.1111/inm.12011

Grych, J., & Hamby, S. (2014). Advancing the measurement of violence: Challenges and opportunities. *Psychology of Violence*, *4*(4), 363–368. https://doi.org/10.1037/a0037886

Gunn, J., & Robertson, G. (1976). Drawing a criminal profile. *British Journal of Criminology*, *16*(2), 156–160. https://doi.org/10.1093/oxfordjournals.bjc.a046713

Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC Medical Research Methodology*, *13*(1), 100. https://doi.org/10.1186/1471-2288-13-100

Gupta, U. G., & Clarke, R. E. (1996). Theory and applications of the Delphi technique: a bibliography (1975-1994). *Technological Forecasting and Social Change*, *53*(2), 185–211. https://doi.org/10.1016/S0040-1625(96)00094-7

Guy, L. S., Douglas, K. S., & Hart, S. D. (2015). Risk assessment and communication. In B. L. Cutler & P. A. Zapf (Eds.), *APA handbook of forensic psychology, Vol. 1: Individual and situational influences in criminal and civil contexts.* (pp. 35–86). American Psychological Association. https://doi.org/10.1037/14461-003

Haarsma, G., Davenport, S., White, D. C., Ormachea, P. A., Sheena, E., & Eagleman, D. M. (2020). Assessing Risk Among Correctional Community Probation Populations: Predicting Reoffense With Mobile Neurocognitive Assessment Software. *Frontiers in Psychology*, *10*, 2926. https://doi.org/10.3389/fpsyg.2019.02926

Haig, B. D. (2014). Investigating the psychological world: Scientific method in the behavioral sciences. In *Investigating the Psychological World: Scientific Method in the Behavioral Sciences*. MIT Press. https://doi.org/10.5860/choice.185617

Hancock, M., Tapscott, J. L., & Hoaken, P. N. S. (2010). Role of executive dysfunction in predicting frequency and severity of violence. *Aggressive Behavior*, *36*(5), 338–349. https://doi.org/10.1002/ab.20353

Hanczar, B., Hua, J., Sima, C., Weinstein, J., Bittner, M., & Dougherty, E. R. (2010). Small-sample precision of ROC-related estimates. *Bioinformatics*, *26*(6), 822–830. https://doi.org/10.1093/bioinformatics/btq037

Hanson, W., & Ramani, N. (1988). Technology forecasting: a hydroelectric company experience. *Technology Management Publication*, *1*(3), 266–270. https://www.tandfonline.com/loi/urtm20

Haque, Q., & Webster, C. D. (2013). Structured professional judgement and sequential redirections. *Criminal Behaviour and Mental Health*, *23*(4), 241–251. https://doi.org/10.1002/cbm.1886

Haque, Q., & Webster, C. D. (2019). Why is it so difficult to achieve accuracy in predictions of violence? In D. Eaves, C. Webster, Q. Haque, & J. Eaves-Thalken (Eds.), *Risk Rules: A practical guide to structured professional judgement and violence prevention* (1st ed., pp. 65–78). Pavilion Publishing and Media Ltd. https://www.pavpub.com/mental-health/risk-rules

Harrington, L., Siegert, R. J., & McClure, J. (2005). Theory of mind in schizophrenia: A critical review. *Cognitive Neuropsychiatry*, *10*(4), 249–286. https://doi.org/10.1080/13546800444000056

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent Recidivism of Mentally Disordered Offenders: The Development of a Statistical Prediction Instrument. *Criminal Justice and Behavior*, *20*(4), 315–335. https://doi.org/10.1177/0093854893020004001

Harris, G. T., Rice, M. E., Quinsey, V. L., & Cormier, C. A. (2015). *Violent offenders:*

Appraising and managing risk (3rd ed.). (Vol. 51, Issue 3). American Psychological Association. https://doi.org/10.1037/14572-000

Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumière, M. L., Boer, D., & Lang, C. (2003). A Multisite Comparison of Actuarial Risk Instruments for Sex Offenders. *Psychological Assessment*, *15*(3), 413–425. https://doi.org/10.1037/1040-3590.15.3.413

Harris, S. T., Oakley, C., & Picchioni, M. (2013). Quantifying violence in mental health research. *Aggression and Violent Behavior*, *18*(6), 695–701. https://doi.org/10.1016/j.avb.2013.07.022

Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. In *Legal and Criminological Psychology* (Vol. 3, Issue 1, pp. 121–137). John Wiley and Sons Ltd. https://doi.org/10.1111/j.2044-8333.1998.tb00354.x

Hart, S. D. (2011). Complexity, uncertainty, and the reconceptualization of violence risk assessment. In *Victims and Offenders: Chapters on psychology and law* (pp. 57–69). Politeia (Original work published in 2004). https://www.politeia.co.uk/

Hart, S. D., Douglas, K. S., & Guy, L. S. (2016). The Structured Professional Judgement Approach to Violence Risk Assessment. In *The Wiley Handbook on the Theories, Assessment and Treatment of Sexual Offending* (pp. 643–666). John Wiley & Sons Singapore Pte. Ltd. https://doi.org/10.1002/9781118574003.wattso030

Hatton, C. (2002). Psychosocial interventions for adults with intellectual disabilities and mental health problems: A review. *Journal of Mental Health*, *11*(4), 357–374. https://doi.org/10.1080/09638230020023732

Hawkins, K. A., & Trobst, K. K. (2000). Frontal lobe dysfunction and aggression: Conceptual issues and research findings. *Aggression and Violent Behavior*, *5*(2), 147–157. https://doi.org/10.1016/S1359-1789(98)00033-0

Hays, J. R., Solway, K. S., & Schreiner, D. (1978). Intellectual characteristics of juvenile murderers versus status offenders. *Psychological Reports*, *43*(1), 80–82. https://doi.org/10.2466/pr0.1978.43.1.80

Heaton, R. K. (1981). *Wisconsin Card Sorting Test manual*. Psychological Assessment Resources. https://www.parinc.com/Products/Pkey/478

Heilbrun, K. (2009). Evaluation for Risk of Violence in Adults. In *Evaluation for Risk of Violence in Adults*. Oxford University Press.

https://doi.org/10.1093/med:psych/9780195369816.001.0001

Heilbrun, K., Marczyk, G., DeMatteo, D., Zillmer, E., Harris, J., & Jennings, T. (2003). Principles of Forensic Mental Health Assessment. *Assessment*, *10*(4), 329–343. https://doi.org/10.1177/1073191103258591

Higgins, & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions. *The Cochrane Collaboration*. www.cochrane-hanbook.org

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. In *British Medical Journal* (Vol. 327, Issue 7414, pp. 557–560). BMJ Publishing Group. https://doi.org/10.1136/bmj.327.7414.557

Hoaken, P. N. S., & Stewart, S. H. (2003). Drugs of abuse and the elicitation of human aggressive behavior. *Addictive Behaviors*, *28*(9), 1533–1554. https://doi.org/10.1016/j.addbeh.2003.08.033

Hoffer, T., Hargreaves-Cormany, H., Muirhead, Y., & Meloy, J. R. (2018). Meloy's Bimodal Theory of Affective (Reactive) and Predatory (Instrumental) Violence. In *Violence in Animal Cruelty Offenders* (pp. 21–23). Springer. https://doi.org/10.1007/978-3-319-91038-3_7

Hoptman, M. J. (2003). Neuroimaging Studies of Violence and Antisocial Behavior. *Journal of Psychiatric Practice*, *9*(4), 265–278. https://doi.org/10.1097/00131746-200307000-00002

Houdé, O., & Borst, G. (2015). Evidence for an inhibitory-control theory of the reasoning brain. *Frontiers in Human Neuroscience*, *9*(MAR), 148. https://doi.org/10.3389/fnhum.2015.00148

Hounsome, J., Whittington, R., Brown, A., Greenhill, B., & McGuire, J. (2018). The Structured Assessment of Violence Risk in Adults with Intellectual Disability: A Systematic Review. *Journal of Applied Research in Intellectual Disabilities*, *31*(1), e1–e17. https://doi.org/10.1111/jar.12295

Howard, R. C., Khalifa, N., & Duggan, C. (2014). Antisocial personality disorder comorbid with borderline pathology and psychopathy is associated with severe violence in a forensic sample. *Journal of Forensic Psychiatry and Psychology*, *25*(6), 658–672. https://doi.org/10.1080/14789949.2014.943797

Howieson, D. (2019). Current limitations of neuropsychological tests and assessment procedures. *Clinical Neuropsychologist*, *33*(2), 200–208.

https://doi.org/10.1080/13854046.2018.1552762

Hughes, K., Bellis, M. A., Hardcastle, K. A., Sethi, D., Butchart, A., Mikton, C., Jones, L., & Dunne, M. P. (2017). The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. *The Lancet Public Health*, *2*(8), e356–e366. https://doi.org/10.1016/S2468-2667(17)30118-4

Jackson, C. J., Loxton, N. J., Harnett, P., Ciarrochi, J., & Gullo, M. J. (2014). Original and revised reinforcement sensitivity theory in the prediction of executive functioning: A test of relationships between dual systems. *Personality and Individual Differences*, *56*(1), 83–88. https://doi.org/10.1016/j.paid.2013.08.024

Janes, S., O'rourke, S., & Schwannauer, M. (2021). Assessing the Cognitive Contributors to Violence: A Pilot and Feasibility Study Protocol. *Social Science Protocols*, 1–16. https://doi.org/10.7565/ssp.2021.v4.5213

Jaworski, F., Dubertret, C., Adès, J., & Gorwood, P. (2011). Presence of co-morbid substance use disorder in bipolar patients worsens their social functioning to the level observed in patients with schizophrenia. *Psychiatry Research*, *185*(1–2), 129–134. https://doi.org/10.1016/j.psychres.2010.06.005

Jones, H. (1992). Neuropsychology of violence. *Forensic Reports*, *5*, 221–233.

José Fernández-Serrano, M., Pérez-García, M., Schmidt Río-Valle, J., & Verdejo-García, A. (2010). Neuropsychological consequences of alcohol and drug abuse on different components of executive functions. *Journals.Sagepub.Com*, *24*(9), 1317–1332. https://doi.org/10.1177/0269881109349841

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kassambara, A. (2017). *Machine Learning Essentials: Practical Guide in R* (1st ed., Vol. 1). STHDA. http://www.sthda.com

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Kendler, K. S., Ohlsson, H., Mezuk, B., Sundquist, K., & Sundquist, J. (2016). A Swedish

National Prospective and Co-relative Study of School Achievement at Age 16, and Risk for Schizophrenia, Other Nonaffective Psychosis, and Bipolar Illness. *Schizophrenia Bulletin*, *42*(1), 77–86. https://doi.org/10.1093/schbul/sbv103

Kennedy, T. D., Burnett, K. F., & Edmonds, W. A. (2011). Intellectual, behavioral, and personality correlates of violent vs. non-violent juvenile offenders. *Aggressive Behavior*, *37*(4), 315–325. https://doi.org/10.1002/ab.20393

Khan, R. A., & Brandenburger, T. (2014). *Performance Assessment of Binary Classifier with Visualization [R package ROCit version 2.1.1]*. Comprehensive R Archive Network (CRAN). https://cran.r-project.org/package=ROCit

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352–358. https://doi.org/10.1037/h0043688

Kirkish, P., & Sreenivasan, S. (1999). Neuropsychological assessment of competency to stand trial evaluations: a practical conceptual model. *Journal of the American Academy of Psychiatry and the Law*. https://psycnet.apa.org/record/1999-10991-008

Klepfisz, G., Daffern, M., & Day, A. (2017). Understanding protective factors for violent reoffending in adults. *Aggression and Violent Behavior, 32*, 80–87. https://doi.org/10.1016/j.avb.2016.12.001

Koenen, K. C., Moffitt, T. E., Caspi, A., Taylor, A., & Purcell, S. (2003). Domestic violence is associated with environmental suppression of IQ in young children. *Development and Psychopathology*, *15*(2), 297–311. https://doi.org/10.1017/S0954579403000166

Kravariti, E., Morgan, K., Fearon, P., Zanelli, J. W., Lappin, J. M., Dazzan, P., Morgan, C., Doody, G. A., Harrison, G., Jones, P. B., Murray, R. M., & Reichenberg, A. (2009). Neuropsychological functioning in first-episode schizophrenia. *British Journal of Psychiatry*, *195*(4), 336–345. https://doi.org/10.1192/bjp.bp.108.055590

Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. In *Nature Reviews Neuroscience* (Vol. 6, Issue 9, pp. 691–702). Nat Rev Neurosci. https://doi.org/10.1038/nrn1747

Kroese, Biza Stenfert. (2005). Cognitive-behaviour therapy for people with learning disabilities: Conceptual and contextual issues. In B.S. Kroese, D. Dagnan, & K. Loumidis (Eds.), *Cognitive-Behaviour Therapy for People with Learning Disabilities* (pp. 1–15). Routledge.

https://doi.org/10.4324/9780203977200-9

Kroner, D. G., & Loza, W. (2001). Evidence for the efficacy of self-report in predicting nonviolent and violent criminal recidivism. *Journal of Interpersonal Violence*, *16*(2), 168–177. https://doi.org/10.1177/088626001016002005

Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior*, *28*(4), 471–489. https://doi.org/10.1177/009385480102800405

Kuhns, J. B., & Clodfelter, T. A. (2009). Illicit drug-related psychopharmacological violence: The current understanding within a causal context. In *Aggression and Violent Behavior* (Vol. 14, Issue 1, pp. 69–78). Pergamon. https://doi.org/10.1016/j.avb.2008.11.001

Kuin, N. C., Masthoff, E. D. M., Munafò, M. R., & Penton-Voak, I. S. (2017). Perceiving the evil eye: Investigating hostile interpretation of ambiguous facial emotional expression in violent and non-violent offenders. *PLoS ONE*, *12*(11). https://doi.org/10.1371/journal.pone.0187080

Kumari, V., Aasen, I., Taylor, P., Ffytche, D. H., Das, M., Barkataki, I., Goswami, S., O'Connell, P., Howlett, M., Williams, S. C. R., & Sharma, T. (2006). Neural dysfunction and violence in schizophrenia: An fMRI investigation. *Schizophrenia Research*, *84*(1), 144–164. https://doi.org/10.1016/j.schres.2006.02.017

Laakso, M. P., Gunning-Dixon, F., Vaurio, O., Repo-Tiihonen, E., Soininen, H., & Tiihonen, J. (2002). Prefrontal volumes in habitually violent subjects with antisocial personality disorder and type 2 alcoholism. *Psychiatry Research - Neuroimaging*, *114*(2), 95–102. https://doi.org/10.1016/S0925-4927(02)00005-7

LaDuke. (2015). *Can neuropsychology inform violence risk assessment?* [Drexel University]. https://idea.library.drexel.edu/islandora/object/idea%3A6618/datastream/OBJ/download/Can_neuropsychology_inform_violence_risk_assessment_.pdf

LaDuke, C., DeMatteo, D., Heilbrun, K., Gallo, J., & Swirsky-Sacchetti, T. (2017). The Neuropsychological Assessment of Justice-Involved Men: Descriptive Analysis, Preliminary Data, and a Case for Group-Specific Norms. *Archives of Clinical Neuropsychology*, *32*(8), 929–942. https://doi.org/10.1093/arclin/acx042

Lagerberg, T. V., Larsson, S., Sundet, K., Hansen, C. B., Hellvin, T., Andreassen, O. A., & Melle, I. (2010). Treatment delay and excessive substance use in bipolar disorder. *Journal* *of Nervous and Mental Disease*, *198*(9), 628–633. https://doi.org/10.1097/NMD.0b013e3181ef3ef4

Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, *73*(5), 467–482. https://doi.org/10.1016/j.techfore.2005.09.002

Langlands, R. L., Jorm, A. F., Kelly, C. M., & Kitchener, B. a. (2008). First aid for depression: A Delphi consensus study with consumers, carers and clinicians. *Journal of Affective Disorders*, *105*(1–3), 157–165. https://doi.org/10.1016/j.jad.2007.05.004

Larrabee, G. J. (2018). Psychometric Foundations of Neuropsychological Assessment. In J. Morgan & J. Ricker (Eds.), *Textbook of Clinical Neuropsychology* (2nd ed., pp. 22–38). Taylor & Francis. https://doi.org/10.4324/9781315271743-3

Lee, A. H., & DiGiuseppe, R. (2018). Anger and aggression treatments: a review of meta-analyses. *Current Opinion in Psychology*, *19*, 65–74. https://doi.org/10.1016/j.copsyc.2017.04.004

Leistico, A. M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, *32*(1), 28–45. https://doi.org/10.1007/s10979-007-9096-6

Lejoyeux, M., Feuché, N., Loi, S., Solomon, J., & Adès, J. (1998). Impulse-control disorders in alcoholics are related to sensation seeking and not to impulsivity. *Psychiatry Research*, *81*(2), 149–155. https://doi.org/10.1016/S0165-1781(98)00103-6

Lejuez, C. W., Aklin, W. M., Jones, H. A., Strong, D. R., Richards, J. B., Kahler, C. W., & Read, J. P. (2003). The Balloon Analogue Risk Task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology*, *11*(1), 26–33. https://doi.org/10.1037/1064-1297.11.1.26

Lejuez, C. W., Richards, J. B., Read, J. P., Kahler, C. W., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84. https://doi.org/10.1037/1076-898X.8.2.75

Levi, M., Nussbaum, D., & Rich, J. (2010). Neuropsychological and personality characteristics of predatory, irritable, and nonviolent offenders: Support for a typology of criminal human aggression. *Criminal Justice and Behavior*, *37*(6), 633–655.

https://doi.org/10.1177/0093854810362342

Levy, B., & Weiss, R. D. (2010). Neurocognitive impairment and psychosis in bipolar I disorder during early remission from an acute episode of mood disturbance. *Journal of Clinical Psychiatry*, *71*(2), 201–206. https://doi.org/10.4088/JCP.08m04663yel

Lim, L., Radua, J., & Rubia, K. (2014). Gray matter abnormalities in childhood maltreatment: A voxelwise metaanalysis. *American Journal of Psychiatry*, *171*(8), 854–863. https://doi.org/10.1176/appi.ajp.2014.13101427

Lineweaver, T. T., Bondi, M. W., Thomas, R. G., & Salmon, D. P. (1999). A normative study of Nelson's (1976) modified version of the Wisconsin Card Sorting Test in healthy older adults. *Clinical Neuropsychologist*, *13*(3), 328–347. https://doi.org/10.1076/clin.13.3.328.1745

Liu, J., & Wuerker, A. (2005). Biosocial bases of aggressive and violent behavior - Implications for nursing studies. *International Journal of Nursing Studies*, *42*(2), 229–241. https://doi.org/10.1016/j.ijnurstu.2004.06.007

Lodewijks, H. P. B., Doreleijers, T. A. H., de Ruiter, C., & Borum, R. (2008). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry*, *31*(3), 263–271. https://doi.org/10.1016/j.ijlp.2008.04.009

Loeber, R., Burke, J. D., & Pardini, D. A. (2009). Development and Etiology of Disruptive and Delinquent Behavior. *Annual Review of Clinical Psychology*, *5*(1), 291–310. https://doi.org/10.1146/annurev.clinpsy.032408.153631

Lumley, T. (2020). *CRAN - Package leaps*. https://cran.r-project.org/web/packages/leaps/index.html

Lynch, C. (2004). Psychotherapy for persons with mental retardation. *Mental Retardation*, *42*(5), 399–405. https://doi.org/10.1352/0047-6765(2004)42<399:PFPWMR>2.0.CO;2

MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. Gorfein & C. MacLeod (Eds.), *Inhibition in cognition.* (pp. 3–23). American Psychological Association. https://doi.org/10.1037/11587-001

Majorek, K., Wolfkühler, W., Küper, C., Saimeh, N., Juckel, G., & Brüne, M. (2009). "Theory of Mind" and Executive Functioning in Forensic Patients with Schizophrenia. *Journal of Forensic Sciences*, *54*(2), 469–473. https://doi.org/10.1111/j.1556-4029.2008.00966.x

Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: Journal of Research and Treatment*, *22*(2), 191–217. https://doi.org/10.1177/1079063210366039

Marceau, R., Meghani, R., & Reddon, J. R. (2008). Neuropsychological assessment of adult offender. *Journal of Offender Rehabilitation*, *47*(1–2), 41–73. https://doi.org/10.1080/10509670801940409

Marson, D. C. (2001). Aging, Neuropsychology, and Cognition Loss of Financial Competency in Dementia: Conceptual and Empirical Approaches. *Aging, Neuropsychology, and Cognition*, *8*(3), 164–181. https://doi.org/10.1076/anec.8.3.164.827

Martell, D. A. (1992). Forensic neuropsychology and the criminal law. *Law and Human Behavior*, *16*(3), 313–336. https://doi.org/10.1007/BF01044772

Mazza, M., Mandelli, L., Di Nicola, M., Harnic, D., Catalano, V., Tedeschi, D., Martinotti, G., Colombo, R., Bria, P., Serretti, A., & Janiri, L. (2009). Clinical features, response to treatment and functional outcome of bipolar disorder patients with and without co-occurring substance use disorder: 1-year follow-up. *Journal of Affective Disorders*, *115*(1–2), 27–35. https://doi.org/10.1016/j.jad.2008.08.019

McDermott, B. E., Edens, J. F., Quanbeck, C. D., Busse, D., & Scott, C. L. (2008). Examining the role of static and dynamic risk factors in the prediction of inpatient violence: Variable- and person-focused analyses. *Law and Human Behavior*, *32*(4), 325–338. https://doi.org/10.1007/s10979-007-9094-8

McDonald, S. (2013). Impairments in social cognition following severe traumatic brain injury. In *Journal of the International Neuropsychological Society* (Vol. 19, Issue 3, pp. 231–246). Cambridge University Press. https://doi.org/10.1017/S1355617712001506

McDonald, S., Flanagan, S., Rollins, J., & Klinch, J. (2003). TASIT: A New Clinical Tool for Assessing Social Perception A... : The Journal of Head Trauma Rehabilitation. *Head Trauma Rehabilitation*, *18*(3), 219–238. https://journals.lww.com/headtraumarehab/Abstract/2003/05000/TASIT__A_New_Clinical_Tool_for_Assessing_Social.1.aspx

McEwen, B. S. (2012). Brain on stress: How the social environment gets under the skin. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(SUPPL.2), 17180–17185. https://doi.org/10.1073/pnas.1121254109

McKee, S. (2004). *Disinhibition as an underlying factor in the violent and recidivistic behaviour of a sample of mentally disordered offenders.* [Your University (Canada)]. https://elibrary.ru/item.asp?id=6719447

McKenna, H. P. (1994). The Delphi technique: a worthwhile research approach for nursing? *Journal of Advanced Nursing*, *19*(6), 1221–1225. https://doi.org/10.1111/j.1365-2648.1994.tb01207.x

McKinlay, A., & Albicini, M. (2016). Prevalence of traumatic brain injury and mental health problems among individuals within the criminal justice system. *Concussion*, *1*(4), CNC25. https://doi.org/10.2217/cnc-2016-0011

McNiel, D. E., Gregory, A. L., Lam, J. N., Binder, R. L., & Sullivan, G. R. (2003). Utility of decision support tools for assessing acute risk of violence. *Journal of Consulting and Clinical Psychology*, *71*(5), 945–953. https://doi.org/10.1037/0022-006X.71.5.945

Meijers, J., Harte, J. M., Meynen, G., & Cuijpers, P. (2017). Differences in executive functioning between violent and non-violent offenders. *Psychological Medicine*, *47*(10), 1784–1793. https://doi.org/10.1017/S0033291717000241

Meloy, J. R. (2006). Empirical basis and Forensic Application of Affective and Predatory Violence. *Australian & New Zealand Journal of Psychiatry*, *40*(6–7), 539–547. https://doi.org/10.1080/j.1440-1614.2006.01837.x

Mezzacappa, E., Kindlon, D., & Earls, F. (2001). Child abuse and performance task assessments of executive functions in boys. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(8), 1041–1048. https://doi.org/10.1111/1469-7610.00803

Mills, J. F., Kroner, D. G., & Hemmati, T. (2007). The validity of violence risk estimates: An issue of item performance. In *Psychological Services* (Vol. 4, Issue 1, pp. 1–12). https://doi.org/10.1037/1541-1559.4.1.1

Mitchell, & Beech, A. R. (2011). Towards a neurobiological model of offending. *Clinical Psychology Review*, *31*(5), 872–882. https://doi.org/10.1016/j.cpr.2011.04.001

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to

Complex "'Frontal Lobe'" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E. (1993). Adolescence-Limited and Life-Course-Persistent Antisocial Behavior: A Developmental Taxonomy. *Psychological Review*, *100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, *8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. 23–26. https://uk.sagepub.com/en-gb/eur/home

Monahan, J. (1984). The prediction of violent behavior: Toward a second generation of theory and policy. *American Journal of Psychiatry*, *141*(1), 10–15. https://doi.org/10.1176/ajp.141.1.10

Monahan, J. (1988). Risk assessment of violence among the mentally disordered: Generating useful knowledge. *International Journal of Law and Psychiatry*, *11*(3), 249–257. https://doi.org/10.1016/0160-2527(88)90012-X

Monahan, J. (2000). Violence Risk Assessment : Scientific Validity and Evidentiary Admissibility. *Wash. & Lee L. Rev.901*, *57*(3), 14. https://scholarlycommons.law.wlu.edu/wlulr/vol57/iss3/8

Monahan, J., & Skeem, J. L. (2014). The evolution of violence risk assessment. *CNS Spectrums*, *19*(5), 419–424. https://doi.org/10.1017/S1092852914000145

Monahan, J., & Steadman, H. (1994). Designing a new generation of risk assessment research. In J. Monahan & H. Steadman (Eds.), *Violence and mental disorder : developments in risk assessment* (pp. 297–318). University of Chicago Press. https://press.uchicago.edu/

Monahan, J., & Steadman, H. J. (1996). Violent Storms and Violent People: How Meteorology Can Inform Risk Communication in Mental Health Law. *American Psychologist*, *51*(9), 931–938. https://doi.org/10.1037/0003-066X.51.9.931

Mooney, C. Z., & Duval, R. D. (1994). Bootstrapping: A Nonparametric Approach to Statistical Inference. *Journal of the American Statistical Association*, *89*(427), 1150. https://doi.org/10.2307/2290969

Morgan, & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, *20*(1), 113–136. https://doi.org/10.1016/S0272-7358(98)00096-8

Morris, Z. S., wooding, S., & Grant, J. (2011). The answer is 17 years, what is the question: Understanding time lags in translational research. *Journal of the Royal Society of Medicine*, *104*(12), 510–520. https://doi.org/10.1258/jrsm.2011.110180

Morrissey, C., Hogue, T., Mooney, P., Allen, C., Johnston, S., Hollin, C., Lindsay, W. R., & Taylor, J. L. (2007). Predictive validity of the PCL-R in offenders with intellectual disability in a high secure hospital setting: Institutional aggression. *Journal of Forensic Psychiatry and Psychology*, *18*(1), 1–15. https://doi.org/10.1080/08990220601116345

Moulden, H. M. (2009). Social competence and sexual aggression: Social intelligence, cognitive distortions, and victim empathy in men who sexually offend against children. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *70*(4-B), 2580. https://doi.org/10.20381/RUOR-13034

Muccigrosso, M. M., Ford, J., Benner, B., Moussa, D., Burnsides, C., Fenn, A. M., Popovich, P. G., Lifshitz, J., Walker, F. R., Eiferman, D. S., & Godbout, J. P. (2016). Cognitive deficits develop 1 month after diffuse brain injury and are exaggerated by microglia-associated reactivity to peripheral immune challenge. *Brain, Behavior, and Immunity*, *54*, 95–109. https://doi.org/10.1016/j.bbi.2016.01.009

Müller, J. L., Sommer, M., Döhnel, K., Weber, T., Schmidt-Wilcke, T., & Hajak, G. (2008). Disturbed prefrontal and temporal brain function during emotion and cognition interaction in criminal psychopathy. *Behavioral Sciences & the Law*, *26*(1), 131–150. https://doi.org/10.1002/bsl.796

National Institute of Health. (2014). *Study quality assessment tools: Quality assessment tool for observational cohort and cross-sectional studies*. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

National Research Council. (1989). *Improving Risk Communication*. National Academies Press. https://doi.org/10.17226/1189

Navalta, C. P., Polcari, A., Webster, D. M., Boghossian, A., & Teicher, M. H. (2006). Effects of childhood sexual abuse on neuropsychological and cognitive function in college women. *Journal of Neuropsychiatry and Clinical Neurosciences*, *18*(1), 45–53. https://doi.org/10.1176/jnp.18.1.45

Nazmie, I. F., Nebi, M. R., Zylfije, H., & Bekim, H. (2013). Poor executive functioning associated with the risk of aggressive behavior recidivism in the forensic community in schizophrenic patients. *International Journal of BioMedicine*, *3*(2), 94–99. https://riccem.org/international-journal-of-biomedicine-ijbm/

Nelson, H. E. (1976). A Modified Card Sorting Test Sensitive to Frontal Lobe Defects. *Cortex*, *12*(4), 313–324. https://doi.org/10.1016/S0010-9452(76)80035-4

Nestor, P. G., Daggett, D., Haycock, J., & Price, M. (1999). Competence to stand trial: A neuropsychological inquiry. *Law and Human Behavior*, *23*(4), 397–412. https://doi.org/10.1023/A:1022339130582

NHS Quality Improvements Scotland. (2010). *Intensive psychiatric care units*. http://www.healthcareimprovementscotland.org/idoc.ashx?docid=f83b114e-62eb-46f0-81bd-e981ead76ef6&version=-1

NICE Clinical Guideline 176. (2014). *Head injury: Assessment and early management*. https://www.nice.org.uk/Guidance/CG176.

Nicholls, T. L., Ogloff, J. R. P., & Douglas, K. S. (2004). Assessing risk for violence among male and female civil psychiatric patients: the HCR-20, PCL:SV, and VSC. *Behavioral Sciences & the Law*, *22*(1), 127–158. https://doi.org/10.1002/bsl.579

Nigel, S. M., Dudeck, M., Otte, S., Knauer, K., Klein, V., Bottcher, T., Maas, C., Vasic, N., & Streb, J. (2018). Psychopathy, the Big Five and empathy as predictors of violence in a forensic sample of substance abusers. *Journal of Forensic Psychiatry & Psychology*, *29*(6), 882–900. https://doi.org/http://dx.doi.org/10.1080/14789949.2018.1439993

Nqwaku, M., Draycott, S., Aldridge-Waddon, L., Bush, E.-L., Tsirimokou, A., Jones, D., & Puzzo, I. (2018). Predictive power of the DASA-IV: Variations in rating method and timescales. *International Journal of Mental Health Nursing*, *27*(6), 1661–1672. https://doi.org/10.1111/inm.12464

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., Essock, S., Fenton, W. S., Frese, F. J., Gold, J. M., Goldberg, T., Heaton, R. K., Keefe, R. S. E., Kraemer, H., Mesholam-Gately, R., Seidman, L. J., Stover, E., Weinberger, D. R., Young, A. S., … Marder, S. R. (2008). The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203–213. https://doi.org/10.1176/appi.ajp.2007.07010042

O’Reilly, K., Donohoe, G., Coyle, C., O’Sullivan, D., Rowe, A., Losty, M., McDonagh, T., McGuinness, L., Ennis, Y., Watts, E., Brennan, L., Owens, E., Davoren, M., Mullaney, R., Abidin, Z., & Kennedy, H. G. (2015). Prospective cohort study of the relationship between neuro-cognition, social cognition and violence in forensic patients with schizophrenia and schizoaffective disorder. *BMC Psychiatry*, *15*(1). https://doi.org/10.1186/s12888-015-0548-0

O’Reilly, K., Donohoe, G., O’Sullivan, D., Coyle, C., Corvin, A., O’Flynn, P., O’Donnell, M., Galligan, T., O’Connell, P., & Kennedy, H. G. (2019). A randomized controlled trial of cognitive remediation for a national cohort of forensic patients with schizophrenia or schizoaffective disorder. *BMC Psychiatry*, *19*(1), 27. https://doi.org/10.1186/s12888-019-2018-6

O’Reilly, K., O’Connell, P., Donohoe, G., Coyle, C., O’Sullivan, D., Azvee, Z., Maddock, C., Sharma, K., Sadi, H., McMahon, M., & Kennedy, H. G. (2016). Anticholinergic burden in schizophrenia and ability to benefit from psychosocial treatment programmes: A 3-year prospective cohort study. *Psychological Medicine*, *46*(15), 3199–3211. https://doi.org/10.1017/S0033291716002154

O’Rourke, S. (2013). Risk assessment and management with clients with cognitive impairment. In C. Logan & L. Johnstone (Eds.), *Managing clinical risk: A guide to effective practice. Issues in forensic practice* (pp. 183–198). Willan. https://openlibrary.org/publishers/Willan_Publishing_(UK)

Oatley, K. (2004). Emotional intelligence and the intelligence of emotions. *Psychological Inquiry*, *15*(3), 216–222. http://www.jstor.com/stable/20447230

Ogilvie, J. M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. *Criminology*, *49*(4), 1063–1107. https://doi.org/10.1111/j.1745-9125.2011.00252.x

Ogloff, J. R. P., & Davis, M. R. (2020). From Predicting Dangerousness to Assessing and Managing Risk for Violence. In *The Wiley Handbook of What Works in Violence Risk Management* (pp. 79–98). Wiley. https://doi.org/10.1002/9781119315933.ch4

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, *42*(1), 15–29. https://doi.org/10.1016/j.im.2003.11.002

Online Surveys. (2018). *Online Surveys*. Jisc. https://onlinesurveys.ac.uk.

Palmer, B. W., Salva, G. N., & Harmell, A. L. (2012). Healthcare decision-making capacity. In G. J. Demakis (Ed.), *National Academy of Neuropsychology series in evidence-based practices. Civil capacities in clinical neuropsychology: Research findings and practical applications* (pp. 69–94). Oxford University Press. https://psycnet.apa.org/record/2011-27638-004

Palmer, B. W., & Savla, G. N. (2007). The association of specific neuropsychological deficits with capacity to consent to research or treatment. In *Journal of the International Neuropsychological Society* (Vol. 13, Issue 6, pp. 1047–1059). Cambridge University Press. https://doi.org/10.1017/S1355617707071299

Pardini, D. A., Raine, A., Erickson, K., & Loeber, R. (2014). Lower amygdala volume in men is associated with childhood aggression, early psychopathic traits, and future violence. *Biological Psychiatry*, *75*(1), 73–80. https://doi.org/10.1016/j.biopsych.2013.04.003

Parrott, D. J., & Giancola, P. R. (2007). Addressing "The criterion problem" in the assessment of aggressive behavior: Development of a new taxonomic system. In *Aggression and Violent Behavior* (Vol. 12, Issue 3, pp. 280–299). Pergamon. https://doi.org/10.1016/j.avb.2006.08.002

Parry-Jones, B. L., Vaughan, F. L., & Miles Cox, W. (2006). Traumatic brain injury and substance misuse: A systematic review of prevalence and outcomes research (1994–2004). *Neuropsychological Rehabilitation*, *16*(5), 537–560. https://doi.org/10.1080/09602010500231875

Paschall, M. J., & Fishbein, D. H. (2002). Executive cognitive functioning and aggression: A public health perspective. In *Aggression and Violent Behavior* (Vol. 7, Issue 3, pp. 215–235). Pergamon. https://doi.org/10.1016/S1359-1789(00)00044-6

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*, 768–774. http://www.impulsivity.org/measurement/bis11

Pearson. (2015). *Welcome to Q-Interactive*. https://qiactive.com/

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: An integrated review of human literature. *Psychopharmacology*, *214*(1), 55–70. https://doi.org/10.1007/s00213-010-2009-2

Pencer, A., & Addington, J. (2003). Substance use and cognition in early psychosis. *Journal of*

Psychiatry & Neuroscience : JPN, *28*(1), 48–54. http://jpn.ca/

Perez, C. M., & Widom, C. S. (1994). Childhood victimization and long-term intellectual and academic outcomes. *Child Abuse and Neglect*, *18*(8), 617–633. https://doi.org/10.1016/0145-2134(94)90012-4

Pettinati, H. M., O'Brien, C. P., & Dundon, W. D. (2013). Current status of co-occurring mood and substance use disorders: A new therapeutic target. *American Journal of Psychiatry*, *170*(1), 23–30. https://doi.org/10.1176/appi.ajp.2012.12010112

Pirelli, G., Gottdiener, W. H., & Zapf, P. A. (2011). A Meta-Analytic review of competency to stand trial research. *Psychology, Public Policy, and Law*, *17*(1), 1–53. https://doi.org/10.1037/a0021713

Pitel, A. L., Rivier, J., Beaunieux, H., Vabret, F., Desgranges, B., & Eustache, F. (2009). Changes in the Episodic Memory and Executive Functions of Abstinent and Relapsed Alcoholics Over a 6-Month Period. *Alcoholism: Clinical and Experimental Research*, *33*(3), 490–498. https://doi.org/10.1111/j.1530-0277.2008.00859.x

Pochon, J.-B., Levy, R., Poline, J.-B., Crozier, S., Lehéricy, S., Pillon, B., Deweer, B., Le Bihan, D., & Dubois, B. (2001). The Role of Dorsolateral Prefrontal Cortex in the Preparation of Forthcoming Actions: an fMRI Study. *Cerebral Cortex*, *11*(3), 260–266. https://doi.org/10.1093/cercor/11.3.260

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (1990). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Lindell & Whit-Ney*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Ponsford, J., Draper, K., & Schönberger, M. (2008). Functional outcome 10 years after traumatic brain injury: Its relationship with demographic, injury severity, and cognitive and emotional status. *Journal of the International Neuropsychological Society*, *14*(2), 233–242. https://doi.org/10.1017/S1355617708080272

Ponsford, J., Whelan-Goodinson, R., & Bahar-Fuchs, A. (2007). Alcohol and drug use following traumatic brain injury: A prospective study. *Brain Injury*, *21*(13–14), 1385–1392. https://doi.org/10.1080/02699050701796960

Porteus, S. D. (1965). *Porteus Maze Tests: Fifty years' application*. Pacific Books. https://psycnet.apa.org/record/1966-00881-000

Power, M., Fell, G., & Wright, M. (2013). Principles for high-quality, high-value testing. *Evidence-Based Medicine*, *18*(1), 5–10. https://doi.org/10.1136/eb-2012-100645

Pulay, A. J., Dawson, D. A., Hasin, D. S., Goldstein, R. B., Ruan, W. J., Pickering, R. P., Huang, B., Chou, S. P., & Grant, B. F. (2008). Violent behavior and DSM-IV psychiatric disorders: Results from the national epidemiologic survey on alcohol and related conditions. *Journal of Clinical Psychiatry*, *69*(1), 12–22. https://doi.org/10.4088/JCP.v69n0103

Quinsey, V. L. (1981). The long term management of the mentally disordered offender. In S. J. Hucker & M. B. Webster (Eds.), *Mental disorder and criminal responsibility* (pp. 137–155). Buttersworths. https://www.researchgate.net/publication/273764357_The_long_term_management_of_the_mentally_disordered_offender

Quinsey, V. L., Harris, G., Rice, M., & Cormier, C. (2000). Violent Offenders: Appraising and Managing Risk. *Psychiatric Services*, *51*(3), 395–396. https://doi.org/10.1176/appi.ps.51.3.395

Quinsey, V. L., Harris, G. T., Rice, M., & Comier, C. (1998). *Violent offenders: Managing and appraising risk*. American Psychological Association. https://www.apa.org/pubs/books/4316068

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.r-project.org/

Raine, A. (2002). Biosocial studies of antisocial and violent behavior in children and adults: A review. *Journal of Abnormal Child Psychology*, *30*(4), 311–326. https://doi.org/10.1023/A:1015754122318

Raine, A., Ishikawa, S., Arce, E., Lencz, T., Knuth, K., Bihrle, S., LaCasse, L., & Colletti, P. (2004). Hippocampal structural asymmetry in unsuccessful psychopaths. *Biological Psychiatry*, *55*(2), 185–191. https://doi.org/10.1016/S0006-3223(03)00727-3

Raine, A., & Yang, Y. (2006). Neural foundations to moral reasoning and antisocial behavior. *Social Cognitive and Affective Neuroscience*, *1*(3), 203–213. https://doi.org/10.1093/scan/nsl033

Rankin, K. P., Kramer, J. H., & Miller, B. L. (2005). Patterns of Cognitive and Emotional Empathy in Frontotemporal Lobar Degeneration. *Cognitive and Behavioral Neurology*, *18*(1), 28–36. https://doi.org/10.1097/01.wnn.0000152225.05377.ab

Reichenberg, A., & Harvey, P. D. (2007). Neuropsychological Impairments in Schizophrenia: Integration of Performance-Based and Brain Imaging Findings. *Psychological Bulletin*, *133*(5), 833–858. https://doi.org/10.1037/0033-2909.133.5.833

Reinharth, J., Reynolds, G., Dill, C., & Serper, M. (2014). Cognitive predictors of violence in schizophrenia: a meta-analytic review. *Schizophrenia Research: Cognition*, *1*(2), 101–111. https://doi.org/10.1016/j.scog.2014.06.001

Reitan, R. M. (1971). Trail making test results for normal and brain-damaged children. *Perceptual and Motor Skills*, *33*(2), 575–581. https://doi.org/10.2466/pms.1971.33.2.575

Rice, M. E., & Harris, G. T. (2016). The Sex Offender Risk Appraisal Guide. In *Sexual Offending* (pp. 471–488). Springer New York. https://doi.org/10.1007/978-1-4939-2416-5_21

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: The violence risk appraisal guide-revised (VRAG-R). *Psychological Assessment*, *25*(3), 951–965. https://doi.org/10.1037/a0032878

Rice, M. E., Harris, G. T., Lang, C., & Cormier, C. (2006). Violent sex offenses: How are they best measured from official records? *Law and Human Behavior*, *30*(4), 525–541. https://doi.org/10.1007/s10979-006-9022-3

Richter, M. S., O'Reilly, K., O'Sullivan, D., O'Flynn, P., Corvin, A., Donohoe, G., Coyle, C., Davoren, M., Higgins, C., Byrne, O., Nutley, T., Nulty, A., Sharma, K., O'Connell, P., & Kennedy, H. G. (2018). Prospective observational cohort study of 'treatment as usual' over four years for patients with schizophrenia in a national forensic hospital. *BMC Psychiatry*, *18*(1), 289. https://doi.org/10.1186/s12888-018-1862-0

Ridderinkhof, K. R., De Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. H. (2002). Alcohol Consumption Impairs Detection of Performance Errors in Mediofrontal Cortex. *Science*, *298*(5601), 2209–2211. https://doi.org/10.1126/science.1076929

Rimmer, M. L. (1998). Executive functioning and problem-solving ability in youthful offenders: A neuropsychological assessment [California School of Professional Psychology]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering: Vol. .59* (Issues 4-B). https://elibrary.ru/item.asp?id=5499910

Risk Management Authority. (2011). *Framework for Risk Assessment Management and*

Evaluation: FRAME. 4–86. https://www.rma.scot/wp-content/uploads/2018/02/FRAME_policy.pdf

Rizopoulos, D. (2021). *Generalized Linear Mixed Models using Adaptive Gaussian Quadrature — GLMMadaptive* (0.8-0). CRAN. https://drizopoulos.github.io/GLMMadaptive/reference/GLMMadaptive.html

Robertson, G., Taylor, P. J., & Gunn, J. C. (1987). Does violence have cognitive correlates? *British Journal of Psychiatry*, *151*(JULY), 63–68. https://doi.org/10.1192/bjp.151.1.63

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*. https://doi.org/10.1186/1471-2105-12-77

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (2nd ed., Vol. 621, pp. 231–244). Russell Sage Foundation. https://www.russellsage.org/

Rössler, W., Ajdacic-Gross, V., Müller, M., Rodgers, S., Haker, H., & Hengartner, M. P. (2015). Assessing sub-clinical psychosis phenotypes in the general population - A multidimensional approach. *Schizophrenia Research*, *161*(2–3), 194–201. https://doi.org/10.1016/j.schres.2014.11.033

Ruscio, J. (2008). A Probability-Based Measure of Effect Size: Robustness to Base Rates and Other Factors. *Psychological Methods*, *13*(1), 19. https://doi.org/10.1037/1082-989X.13.1.19

Rusticus, S. A., & Lovato, C. Y. (2014). Impact of sample size and variability on the power and type I error rates of equivalence tests: A simulation study. *Practical Assessment, Research and Evaluation*, *19*(11), 1–10. https://doi.org/10.7275/4s9m-4e81

Sachdev, P., Smith, J. S., & Cathcart, S. (2001). Schizophrenia-like psychosis following traumatic brain injury: A chart-based descriptive and case-control study. *Psychological Medicine*, *31*(2), 231–239. https://doi.org/10.1017/S0033291701003336

Salas-Wright, C. P., Vaughn, M. G., Reingle Gonzalez, J. M., Fu, Q., & Clark Goings, T. (2016). Attacks Intended to Seriously Harm and Co-occurring Drug Use Among Youth in the United States. *Substance Use & Misuse*, *51*(13), 1681–1692. https://doi.org/10.1080/10826084.2016.1191516

Sapolsky, R. M. (2004). The frontal cortex and the criminal justice system.

Royalsocietypublishing.Org, *359*(1451), 1787–1796. https://doi.org/10.1098/rstb.2004.1547

Schmidt, R. C. (1997). Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences*, *28*(3), 763–774. https://doi.org/10.1111/j.1540-5915.1997.tb01330.x

Schmidt, R. C., Lyytinen, K., Keil, M., & Cule, P. (2001). Identifying Software Project Risks: An International Delphi Study. *Journal of Management Information Systems*, *17*(4), 5–36. https://doi.org/10.1080/07421222.2001.11045662

Schwalbe, C. S. (2008). A Meta-Analysis of Juvenile Justice Risk Assessment Instruments. *Criminal Justice and Behavior*, *35*(11), 1367–1381. https://doi.org/10.1177/0093854808324377

Scottish Government. (2016). *Inpatient Census 2016*. https://www.gov.scot/publications/inpatient-census-2016-part-1-mental-health-learning-disability-inpatient/pages/7/

Sedgwick, O., Young, S., Baumeister, D., Greer, B., Das, M., & Kumari, V. (2017). Neuropsychology and emotion processing in violent individuals with antisocial personality disorder or schizophrenia: The same or different? A systematic review and meta-analysis. *The Australian and New Zealand Journal of Psychiatry*, *51*(12), 1178–1197. https://doi.org/10.1177/0004867417731525

Seidman, L. J., Cherkerzian, S., Goldstein, J. M., Agnew-Blais, J., Tsuang, M. T., & Buka, S. L. (2013). Neuropsychological performance and family history in children at age 7 who develop adult schizophrenia or bipolar psychosis in the New England Family Studies. *Psychological Medicine*, *43*(1), 119–131. https://doi.org/10.1017/S0033291712000773

Seidman, L. J., & Mirsky, A. F. (2017). Evolving notions of schizophrenia as a developmental neurocognitive disorder. In *Journal of the International Neuropsychological Society* (Vol. 23, Issues 9-10 Special Issue, pp. 881–892). https://doi.org/10.1017/S1355617717001114

Serper, M. R. (2011). Aggression in schizophrenia. In *Schizophrenia Bulletin* (Vol. 37, Issue 5, pp. 897–898). Oxford Academic. https://doi.org/10.1093/schbul/sbr090

Serper, M. R., Bergman, A., Copersino, M. L., Chou, J. C. Y., Richarme, D., & Cancro, R. (2000). Learning and memory impairment in cocaine-dependent and comorbid schizophrenic patients. *Psychiatry Research*, *93*(1), 21–32. https://doi.org/10.1016/S0165-1781(99)00122-5

Sevy, S., Kay, S. R., Opler, L. A., & Van Praag, H. M. (1990). Significance of cocaine history in schizophrenia. *Journal of Nervous and Mental Disease*, *178*(10), 642–648. https://doi.org/10.1097/00005053-199010000-00005

Shadish, W. R., Kyse, E. N., & Rindskopf, D. M. (2013). Analyzing data from single-case designs using multilevel models: New applications and some agenda items for future research. *Psychological Methods*, *18*(3), 385–405. https://doi.org/10.1037/a0032964

Shannon, C., Douse, K., McCusker, C., Feeney, L., Barrett, S., & Mulholland, C. (2011). The association between childhood trauma and memory functioning in schizophrenia. *Schizophrenia Bulletin*, *37*(3), 531–537. https://doi.org/10.1093/schbul/sbp096

She, P., & Stapleton, D. (2006). A Review of Disability Data for the Institutional Population: Research Brief. In *K. Lisa Yang and Hock E. Tan Institute on Employment and Disability Collection*. https://digitalcommons.ilr.cornell.edu/edicollect/1205

Sheffield, J. M., Kandala, S., Burgess, G. C., Harms, M. P., & Barch, D. M. (2016). Cingulo-opercular Network Efficiency Mediates the Association Between Psychotic-like Experiences and Cognitive Ability in the General Population. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *1*(6), 498–506. https://doi.org/10.1016/j.bpsc.2016.03.009

Sheffield, J. M., Karcher, N. R., & Barch, D. M. (2018). Cognitive Deficits in Psychotic Disorders: A Lifespan Perspective. In *Neuropsychology Review* (Vol. 28, Issue 4, pp. 509–533). Springer New York LLC. https://doi.org/10.1007/s11065-018-9388-2

Shiroma, E. J., Ferguson, P. L., & Pickelsimer, E. E. (2010). Prevalence of Traumatic Brain Injury in an Offender Population: A Meta-Analysis. *Journal of Correctional Health Care*, *16*(2), 147–159. https://doi.org/10.1177/1078345809356538

Shook, J. J., Vaughn, M. G., & Salas-Wright, C. P. (2013). Exploring the Variation in Drug Selling Among Adolescents in the United States. *Journal of Criminal Justice*, *41*(6), 365–374. https://doi.org/10.1016/j.jcrimjus.2013.07.008

Shumlich, E. J., Reid, G. J., Hancock, M., & N. S. Hoaken, P. (2019). Executive Dysfunction in Criminal Populations: Comparing Forensic Psychiatric Patients and Correctional Offenders. *International Journal of Forensic Mental Health*, *18*(3), 243–259. https://doi.org/10.1080/14999013.2018.1495279

Simonsen, C., Sundet, K., Vaskinn, A., Birkenaes, A. B., Engh, J. A., Faerden, A., Jonsdottir, H., Ringen, P. A., Opjordsmoen, S., Melle, I., Friis, S., & Andreassen, O. A. (2011). Neurocognitive Dysfunction in Bipolar and Schizophrenia Spectrum Disorders Depends on History of Psychosis Rather Than Diagnostic Group. *Schizophrenia Bulletin*, *37*(1), 73–83. https://doi.org/10.1093/schbul/sbp034

Singh, J. P. (2012). The history, development, and testing of forensic risk assessment tools. In *Handbook of Juvenile Forensic Psychology and Psychiatry* (pp. 215–225). Springer US. https://doi.org/10.1007/978-1-4614-0905-2_14

Singh, J. P. (2013). Predictive Validity Performance Indicators in Violence Risk Assessment: A Methodological Primer. *Behavioral Sciences and the Law*, *31*(1), 8–22. https://doi.org/10.1002/bsl.2052

Singh, J. P., Bjorkly, S., & Fazel, S. (2016). International Perspectives on Violence Risk Assessment. In *International Perspectives on Violence Risk Assessment*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199386291.001.0001

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., Doyle, M., Folino, J. O., Godoy-Cervera, V., Grann, M., Ho, R. M. Y., Large, M. M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., de Ruiter, C., Seewald, K., & Otto, R. K. (2014). International Perspectives on the Practical Application of Violence Risk Assessment: A Global Survey of 44 Countries. *International Journal of Forensic Mental Health*, *13*(3), 193–206. https://doi.org/10.1080/14999013.2014.922141

Singh, J. P., Desmarais, S. L., & Van Dorn, R. A. (2013). Measurement of Predictive Validity in Violence Risk Assessment Studies: A Second-Order Systematic Review. *Behavioral Sciences & the Law*, *31*(1), 55–73. https://doi.org/10.1002/bsl.2053

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*(3), 499–513. https://doi.org/10.1016/j.cpr.2010.11.009

Singh, J. P., Grann, M., & Fazel, S. (2013). Authorship Bias in Violence Risk Assessment? A Systematic Review and Meta-Analysis. *PLoS ONE*, *8*(9), e72484. https://doi.org/10.1371/journal.pone.0072484

Skeem, J. L., & Monahan, J. (2014). Risk Redux: The Resurgence of Risk Assessment in Criminal Sanctioning. *Federal Sentencing Reporter*, *26*(3), 158–166.

https://doi.org/0.1525/fsr.2014.26.3.158

Skulmoski, G., Hartman, F., & Krahn, J. (2007). The Delphi Method for Graduate Research. *Journal of Information Technology Education: Research*, *6*, 001–021. https://doi.org/10.28945/199

Slaughter, B., Fann, J. R., & Ehde, D. (2003). Brain Injury Traumatic brain injury in a county jail population: prevalence, neuropsychological functioning and psychiatric disorders. *Taylor & Francis*, *17*(9), 731–741. https://doi.org/10.1080/0269905031000088649

Smith, S. T., Edens, J. F., & McDermott, B. E. (2013). Fearless Dominance and Self-Centered Impulsivity Interact to Predict Predatory Aggression among Forensic Psychiatric Inpatients. *International Journal of Forensic Mental Health*, *12*(1), 33–41. https://doi.org/10.1080/14999013.2012.760186

Snowden, R. J., Gray, N. S., Taylor, J., & MacCulloch, M. J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. *Psychological Medicine*, *37*(11), 1539–1549. https://doi.org/10.1017/S0033291707000876

Soyka, M., Graz, C., Bottlender, R., Dirschedl, P., & Schoech, H. (2007). Clinical correlates of later violence and criminal offences in schizophrenia. *Schizophrenia Research*, *94*(1–3), 89–98. https://doi.org/10.1016/j.schres.2007.03.027

Spikman, J. M., Deelman, B. G., & Van Zomeren, A. H. (2000). Executive functioning, attention and frontal lesions in patients with chronic CHI. *Journal of Clinical and Experimental Neuropsychology*, *22*(3), 325–338. https://doi.org/10.1076/1380-3395(200006)22:3;1-V;FT325

Sprong, M., Schothorst, P., Vos, E., Hox, J., & Van Engeland, H. (2007). Theory of mind in schizophrenia: Meta-analysis. *British Journal of Psychiatry*, *191*(JULY), 5–13. https://doi.org/10.1192/bjp.bp.107.035899

Sreenivasan, S., Kirkish, P., Garrick, T., Weinberger, L. E., & Phenix, A. (2000). Actuarial risk assessment models: A review of critical issues related to violence and sex-offender recidivism assessments. *Journal of the American Academy of Psychiatry and the Law*, *28*(4), 438–448. http://jaapl.org/

Stadtland, C., Hollweg, M., Kleindienst, N., Dietl, J., Reich, U., & Nedopil, N. (2005). Risk assessment and prediction of violent and sexual recidivism in sex offenders: Long-term predictive validity of four risk assessment instruments. *Journal of Forensic Psychiatry and*

Psychology, *16*(1), 92–108. https://doi.org/10.1080/1478994042000270247

Stein, D. J. (2000). The neurobiology of evil: Psychiatric perspectives on perpetrators. *Ethnicity and Health*, *5*(3–4), 303–315. https://doi.org/10.1080/713667458

Stephan, R. A., Alhassoon, O. M., Allen, K. E., Wollman, S. C., Hall, M., Thomas, W. J., Gamboa, J. M., Kimmel, C., Stern, M., Sari, C., Dalenberg, C. J., Sorg, S. F., & Grant, I. (2017). Meta-analyses of clinical neuropsychological tests of executive dysfunction and impulsivity in alcohol use disorder. *The American Journal of Drug and Alcohol Abuse*, *43*(1), 24–43. https://doi.org/10.1080/00952990.2016.1206113

Stokes, F. (1997). Using the Delphi Technique in Planning a Research Project on the Occupational Therapist's Role in Enabling People to make Vocational Choices following Illness or Injury. *British Journal of Occupational Therapy*, *60*(6), 263–267. https://doi.org/10.1177/030802269706000607

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Sturmey, P. (2004). Cognitive therapy with people with intellectual disabilities: a selective review and critique. *Clinical Psychology & Psychotherapy*, *11*(4), 222–232. https://doi.org/10.1002/cpp.409

Sturmey, P., & McMurran, M. (2011). Forensic Case Formulation. In *Forensic Case Formulation*. John Wiley and Sons. https://doi.org/10.1002/9781119977018

Sutherland, R., Sindicich, N., Barrett, E., Whittaker, E., Peacock, A., Hickey, S., & Burns, L. (2015). Motivations, substance use and other correlates amongst property and violent offenders who regularly inject drugs. *Addictive Behaviors*, *45*, 207–213. https://doi.org/10.1016/j.addbeh.2015.01.034

Tarter, R. E., Hegedus, A. M., Alterman, A. I., & Katz-Garris, L. (1983). Cognitive capacities of juvenile violent, nonviolent, and sexual offenders. *Journal of Nervous and Mental Disease*, *171*(9), 564–567. https://doi.org/10.1097/00005053-198309000-00007

Taylor, Kreutzer, J. S., Demm, S. R., & Meade, M. A. (2003). Traumatic brain injury and substance abuse: A review and analysis of the literature. *Neuropsychological Rehabilitation*, *13*(1–2), 165–188. https://doi.org/10.1080/09602010244000336

Teicher, M. H., & Samson, J. A. (2016). Annual Research Review: Enduring neurobiological effects of childhood abuse and neglect. *Journal of Child Psychology and Psychiatry and*

Allied Disciplines, *57*(3), 241–266. https://doi.org/10.1111/jcpp.12507

Teichner, G., & Golden, C. J. (2000). The relationship of neuropsychological impairment to conduct disorder in adolescence: A conceptual review. *Aggression and Violent Behavior*, *5*(6), 509–528. https://doi.org/10.1016/S1359-1789(98)00035-4

Tengström, A. (2001). Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia. *Nordic Journal of Psychiatry*, *55*(4), 243–249. https://doi.org/10.1080/080394801681019093

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., Thabane, M., Giangregorio, L., & Goldsmith, C. H. (2010). A tutorial on pilot studies: The what, why and how. *BMC Medical Research Methodology*, *10*. https://doi.org/10.1186/1471-2288-10-1

The NHS Confederation. (2012). *Defining mental health services*. www.nhsconfed.org

Tiihonen, J., Isohanni, M., Räsänen, P., Koiranen, M., & Moring, J. (1997). Specific major mental disorders and criminality: A 26-year prospective study of the 1966 Northern Finland birth cohort. *American Journal of Psychiatry*, *154*(6), 840–845. https://doi.org/10.1176/ajp.154.6.840

Tonnaer, F., Cima, M., & Arntz, A. (2016). Executive (DYS) functioning and impulsivity as possible vulnerability factors for aggression in forensic patients. *Journal of Nervous and Mental Disease*, *204*(4), 280–286. https://doi.org/10.1097/NMD.0000000000000485

Tooze, J. A., Grunwald, G. K., & Jones, R. H. (2002). Analysis of repeated measures data with clumping at zero. In *Statistical Methods in Medical Research* (Vol. 11, Issue 4, pp. 341–355). Sage PublicationsSage CA: Thousand Oaks, CA. https://doi.org/10.1191/0962280202sm291ra

Trisha, C., Golnoush, A., Jan-Marie, K., Torres, I. J., & Yatham, L. N. (2018). Cognitive functioning in first episode bipolar I disorder patients with and without history of psychosis. *Journal of Affective Disorders*, *227*, 109–116. https://doi.org/10.1016/j.jad.2017.10.003

Ttofi, M. M., Farrington, D. P., Piquero, A. R., Lösel, F., DeLisi, M., & Murray, J. (2016). Intelligence as a protective factor against offending: A meta-analytic review of prospective longitudinal studies. *Journal of Criminal Justice*, *45*, 4–18. https://doi.org/10.1016/j.jcrimjus.2016.02.003

Turkstra, L., Jones, D., & Toler, H. L. (2003). Brain injury and violent crime. *Brain Injury : [BI]*, *17*(1), 39–47. https://doi.org/10.1080/0269905021000010122

Ullman, D. (1989). *Neuropsychological correlates of assaultive behavior among incarcerated males - ProQuest* [The University of Nebraska-Lincoln]. https://search-proquest-com.ezproxy.is.ed.ac.uk/docview/303566911/?pq-origsite=primo

Ullrich, S., Keers, R., & Coid, J. (2014). Delusions, anger, and serious violence: New findings from the macarthur violence risk assessment study. *Schizophrenia Bulletin*, *40*(5), 1174–1181. https://doi.org/10.1093/schbul/sbt126

Umbach, R., Leonard, N. R., Luciana, M., Ling, S., & Laitner, C. (2019). The Iowa Gambling Task in Violent and Nonviolent Incarcerated Male Adolescents. *Criminal Justice and Behavior*, *46*(11), 1611–1629. https://doi.org/10.1177/0093854819847707

Umbrasas, K. V. (2018). Low intellectual ability does not predict violent crime in a military forensic sample. *Intelligence*, *71*, 41–45. https://doi.org/10.1016/j.intell.2018.10.003

Underwood, M. D., Bakalian, M. J., Escobar, T., Kassir, S., Mann, J. J., & Arango, V. (2019). Early-Life Adversity, but Not Suicide, Is Associated With Less Prefrontal Cortex Gray Matter in Adulthood. *International Journal of Neuropsychopharmacology*, *22*(5), 349–357. https://doi.org/10.1093/ijnp/pyz013

van den Dries, L., Juffer, F., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2010). Infants' Physical and Cognitive Development After International Adoption From Foster Care or Institutions in China. *Journal of Developmental & Behavioral Pediatrics*, *31*(2), 144–150. https://doi.org/10.1097/DBP.0b013e3181cdaa3a

Van Der Gronde, T., Kempes, M., Van El, C., Rinne, T., & Pieters, T. (2014). Neurobiological correlates in forensic assessment: A systematic review. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0110672

van Goozen, S. H. M., & Fairchild, G. (2006). Neuroendocrine and neurotransmitter correlates in children with antisocial behavior. *Hormones and Behavior*, *50*(4), 647–654. https://doi.org/10.1016/j.yhbeh.2006.06.021

Van Holst, R. J., Janke Van Holst, R., & Schilt, T. (2011). Drug-Related Decrease in Neuropsychological Functions of Abstinent Drug Users. *Current Drug Abuse Reviews*, *4*, 42–56. https://doi.org/10.2174/1874473711104010042

van Teijlingen, E. R., & Hundley, V. (2001). The importance of pilot studies. *Social Research*

Update, *35*. https://aura.abdn.ac.uk/handle/2164/157

Vanderhoff, H., Jeglic, E. L., & Donovick, P. J. (2011). Neuropsychological assessment in prisons: ethical and practical challenges. *Journal of Correctional Health Care : The Official Journal of the National Commission on Correctional Health Care*, *17*(1), 51–60. https://doi.org/10.1177/1078345810385914

Verdejo-García, A., Bechara, A., Recknor, E. C., & Pérez-García, M. (2006). Executive dysfunction in substance dependent individuals during drug use and abstinence: An examination of the behavioral, cognitive and emotional correlates of addiction. *Journal of the International Neuropsychological Society*, *12*(3), 405–415. https://doi.org/10.1017/S1355617706060486

Verdejo-García, A., López-Torrecillas, F., Giménez, C. O., & Pérez-García, M. (2004). Clinical implications and methodological challenges in the study of the neuropsychological correlates of cannabis, stimulant, and opioid abuse. In *Neuropsychology Review* (Vol. 14, Issue 1, pp. 1–41). Springer. https://doi.org/10.1023/B:NERV.0000026647.71528.83

Viechtbauer, W. (2010). Metafor: meta-analysis package for R. *Journal of Statisical Software*, *36*(3), 1–48. https://www.jstatsoft.org/

Viechtbauer, W., & Cheung, M. (2010). Outlier and influence diagnostics for meta-analysis. *Wiley Online Library*, *1*(2), 112–125. https://onlinelibrary.wiley.com/

Virkkunen, M., Kallio, E., Rawlings, R., Tokola, R., Poland, R. E., Guidotti, A., Nemeroff, C., Bissette, G., Kalogeras, K., Karonen, S. L., & Linnoila, M. (1994). Personality Profiles and State Aggressiveness in Finnish Alcoholic, Violent Offenders, Fire Setters, and Healthy Volunteers. *Archives of General Psychiatry*, *51*(1), 28–33. https://doi.org/10.1001/archpsyc.1994.03950010028004

Vojt, G. (2014). *The Implementation of Violence Risk Assessments into Forensic Psychiatric Care in Scotland* [The University of Edinburgh]. http://hdl.handle.net/1842/10034

Volavka, J. (1999). The neurobiology of violence: An update. *Journal of Neuropsychiatry and Clinical Neurosciences*, *11*(3), 307–314. https://doi.org/10.1176/jnp.11.3.307

Volavka, J., & Citrome, L. (2011). Pathways to aggression in schizophrenia affect results of treatment. *Schizophrenia Bulletin*, *37*(5), 921–929. https://doi.org/10.1093/schbul/sbr041

Walls, T. A., Fairlie, A. M., & Wood, M. D. (2009). Parents do matter: A longitudinal two-part mixed model of early college alcohol participation and intensity. *Journal of Studies on*

Alcohol and Drugs, *70*(6), 908–918. https://doi.org/10.15288/jsad.2009.70.908

Walsh, A., Wells, J., Gann, S. M., Walsh, A., Wells, J., & Gann, S. M. (2020). Mentally Ill and Mentally Deficient Offenders. In *Correctional Assessment, Casework, and Counseling* (pp. 353–373). Springer International Publishing. https://doi.org/10.1007/978-3-030-55226-8_17

Walters, G. D. (2003). Predicting Institutional Adjustment and Recidivism With the Psychopathy Checklist Factor Scores: A Meta-Analysis. *Law and Human Behavior*, *27*(5), 541–558. https://doi.org/10.1023/A:1025490207678

Ward, T., & Beech, A. R. (2015). Dynamic risk factors: a theoretical dead-end? *Psychology, Crime and Law, 21*(2), 100–113. https://doi.org/10.1080/1068316X.2014.917854

Ward, T., Polaschek, D. L. L., & Beech, A. R. (2005). Theories of Sexual Offending. In T. Ward, D. L. L. Polaschek, & A. R. Beech (Eds.), *Theories of Sexual Offending*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470713648

Warren, J. I., South, S. C., Burnette, M. L., Rogers, A., Friend, R., Bale, R., & Van Patten, I. (2005). Understanding the risk factors for violence and criminality in women: The concurrent validity of the PCL-R and HCR-20. *International Journal of Law and Psychiatry*, *28*(3), 269–289. https://doi.org/10.1016/j.ijlp.2003.09.012

Webster, C. D., Douglas, K., Eaves, D., & Hart, S. D. (1997). Assessing risk of violence to others. In *Impulsivity: Theory, assessment, and treatment* (pp. 251–277). https://www.guilford.com/

Webster, C. D., Nicholls, T. L., Martin, M. Lou, Desmarais, S. L., & Brink, J. (2006). Short-Term Assessment of Risk and Treatability (START): The case for a new structured professional judgment scheme. *Behavioral Sciences and the Law, 24*(6), 747–766. https://doi.org/10.1002/bsl.737

Wechsler D. (2011). The Wechsler Abbreviated Scale of Intelligence (2nd ed). *Pearson*. https://www.pearsonclinical.co.uk/

Weil, Z. M., Corrigan, J. D., & Karelina, K. (2016). Alcohol abuse after traumatic brain injury: Experimental and clinical evidence. In *Neuroscience and Biobehavioral Reviews* (Vol. 62, pp. 89–99). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2016.01.005

Weiss, E. M., Kohler, C. G., Nolan, K. A., Czobor, P., Volavka, J., Platt, M. M., Brensinger, C., Loughead, J., Delazer, M., Gur, R. E., & Gur, R. C. (2006). The relationship between history of violent and criminal behavior and recognition of facial expression of emotions in men with schizophrenia and schizoaffective disorder. *Aggressive Behavior*, *32*(3), 187–194. https://doi.org/10.1002/ab.20120

Wells, G. ., Shea, B., O'connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2015). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis*. Ottawa Hospital Research Institute. https://ci.nii.ac.jp/naid/10020590649/

Welty, G. (1972). Problems Of Selecting Experts For Delphi Exercises. *Academy of Management Journal*, *15*(1), 121–124. https://doi.org/10.5465/254805

White, A. J., Meares, S., & Batchelor, J. (2014). The role of cognition in fitness to stand trial: a systematic review. *The Journal of Forensic Psychiatry & Psychology*, *25*(1), 77–99. https://doi.org/10.1080/14789949.2013.868916

Williams, H., Chitsabesan, P., Fazel, S., McMillan, T., Hughes, N., Parsonage, M., & Tonks, J. (2018). Traumatic brain injury: a potential cause of violent crime? In *The Lancet Psychiatry* (Vol. 5, Issue 10, pp. 836–844). Elsevier Ltd. https://doi.org/10.1016/S2215-0366(18)30062-2

Williams, H., Cordan, G., Mewse, A. J., Tonks, J., & Burgess, C. N. W. (2010). Self-reported traumatic brain injury in male young offenders: a risk factor for re-offending, poor mental health and violence? *Neuropsychological Rehabilitation*, *20*(February 2015), 801–812. https://doi.org/10.1080/09602011.2010.519613

Williams, H., Mewse, A. J., Tonks, J., Mills, S., Burgess, C. N. W., & Cordan, G. (2010). Traumatic brain injury in a prison population: Prevalence and risk for re-offending. *Brain Injury*, *24*(10), 1184–1188. https://doi.org/10.3109/02699052.2010.495697

Willmott, C., Ponsford, J., Hocking, C., & Schönberger, M. (2009). Factors Contributing to Attentional Impairments After Traumatic Brain Injury. *Neuropsychology*, *23*(4), 424–432. https://doi.org/10.1037/a0015058

Willner, P. (2005). The effectiveness of psychotherapeutic interventions for people with learning disabilities: A critical overview. *Journal of Intellectual Disability Research*, *49*(1), 73–85. https://doi.org/10.1111/j.1365-2788.2005.00633.x

Wilson, D. (n.d.). *Practical Meta-Analysis Effect Size Calculator [Online calculator]*. Retrieved August 10, 2020, from https://campbellcollaboration.org/research-resources/effect-size-calculator.html

Wilson, L. C., & Scarpa, A. (2012). Criminal Behavior: The need for an integrative approach that incorporates biological influences. *Journal of Contemporary Criminal Justice*, *28*(3), 366–381. https://doi.org/10.1177/1043986212450232

Witt, K., van Dorn, R., & Fazel, S. (2013). Risk Factors for Violence in Psychosis: Systematic Review and Meta-Regression Analysis of 110 Studies. *PLoS ONE*, *8*(2). https://doi.org/10.1371/journal.pone.0055942

Wolff, N., & Shi, J. (2012). Childhood and Adult Trauma Experiences of Incarcerated Persons and Their Relationship to Adult Behavioral Health Problems and Treatment. *International Journal of Environmental Research and Public Health*, *9*(5), 1908–1926. https://doi.org/10.3390/ijerph9051908

Woods, P., & Almvik, R. (2002). The brøset violence checklist (BVC). *Acta Psychiatrica Scandinavica*, *106*(SUPPL.S412), 103–105. https://doi.org/10.1034/j.1600-0447.106.s412.22.x

Woodworth, M., & Porter, S. (2002). In cold blood: Characteristics of criminal homicides as a function of psychopathy. *Journal of Abnormal Psychology*, *111*(3), 436–445. https://doi.org/10.1037/0021-843X.111.3.436

World Health Organization. (2019). *Substance abuse*. https://www.who.int/topics/substance_abuse/en/

Wortzel, H. S., Brenner, L. A., & Arciniegas, D. B. (2013). Traumatic Brain Injury and Chronic Traumatic Encephalopathy: A Forensic Neuropsychiatric Perspective. *Behavioral Sciences & the Law*, *31*(6), 721–738. https://doi.org/10.1002/bsl.2079

Wright, J. H. (2006). Cognitive Behavior Therapy: Basic Principles and Recent Advances. *FOCUS*, *4*(2), 173–178. https://doi.org/10.1176/foc.4.2.173

Wu, C. H., Tsai, T. H., Su, Y. F., Zhang, Z. H., Liu, W., Wu, M. K., Chang, C. H., Kuo, K. L., Lu, Y. Y., & Lin, C. L. (2016). Traumatic Brain Injury and Substance Related Disorder: A 10-Year Nationwide Cohort Study in Taiwan. *Neural Plasticity*, *2016*. https://doi.org/10.1155/2016/8030676

Wykes, T., Huddy, V., Cellard, C., McGurk, S. R., & Czobor, P. (2011). A meta-analysis of cognitive remediation for schizophrenia: Methodology and effect sizes. *American Journal of Psychiatry*, *168*(5), 472–485. https://doi.org/10.1176/appi.ajp.2010.10060855

Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*(5), 740–767. https://doi.org/10.1037/a0020473

Yeomans, R. R. (1996). Neuropsychological profiles of repeat offenders: Violent versus nonviolent girls. [Northwestern University]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering*. http://www.proquest.com

Zanelli, J., Reichenberg, A., Morgan, K., Fearon, P., Kravariti, E., Dazzan, P., Morgan, C., Zanelli, C., Demjaha, A., Jones, P. B., Doody, G. A., Kapur, S., & Murray, R. M. (2010). Specific and generalized neuropsychological deficits: A comparison of patients with various first-episode psychosis presentations. *American Journal of Psychiatry*, *167*(1), 78–85. https://doi.org/10.1176/appi.ajp.2009.09010118

Zhou, J., Witt, K., Zhang, Y., Chen, C., Qiu, C., Cao, L., & Wang, X. (2014). Anxiety, depression, impulsivity and substance misuse in violent and non-violent adolescent boys in detention in China. *Psychiatry Research*, *216*(3), 379–384. https://doi.org/10.1016/j.psychres.2014.01.024

Assessing the Cognitive Contributors to Violence: A Pilot and Feasibility Study Protocol

Sarah Janes[1*], Suzanne O'Rourke[1,2], Matthias Schwannauer[1]

[1]Department of Clinical Psychology, School of Health in Social Science, The University of Edinburgh, Medical School (Doorway 6), Teviot Place, EH8 9AG
[2]NHS State Hospitals Board for Scotland

ABSTRACT

Background: In recent years there has been considerable progress in the development, validation and use of violence risk assessments (VRA). Their predictive ability however remains modest and, due to the repetitive use of certain risk factors, collectively, they appear to have hit an allegorical 'glass ceiling'. Further limiting VRA is the use of self-report, collateral information, and file reviews to assess risk-related factors, rather than validated performance measures. In parallel, findings from neuropsychology and neurobiology have highlighted brain regions associated with violent behaviour. Thus, it is hypothesised that VRA may benefit from the integration of behaviourally measured neuropsychological risk factors.

Methods/Design: The study follows a feasibility and pilot design with a prospective, observational approach. It aims to investigate the feasibility of using a neuropsychological battery to aid in the identification of violence risk in an inpatient and community setting, and to pilot a neuropsychological battery examining risk factors for violence. The primary outcomes of interest are violent incidents or offences recorded during the 6-month follow-up periods.

Discussion: It is our hope that the results of this study will contribute to the development of a structured tool to aid in the identification and assessment of cognitive impairments shown to be predictive of violence risk.

Keywords: Forensic, violence, violence risk assessment, neuropsychology, forensic neuropsychology, risk factors, mentally disordered offenders, violent offenders

1. Background

1.1 Introduction to the literature

The development and validation of violence risk assessments (VRA) have made considerable progress in recent years transforming from 'prediction' only, into 'prediction and management' tools, with more focus on the individual needs of offenders. Subsequently, converging findings from neuropsychological and neurobiological research have identified brain regions associated with violent behaviour and have highlighted a relationship between neurocognitive impairments and violence although findings have been inhibited by methodological limitations and poorly defined outcomes. Whilst there are VRA that encompass

[*] Correspondence to Sarah Janes, Department of Clinical Psychology, School of Health in Social Science, The University of Edinburgh, Medical School (Doorway 6), Teviot Place, EH8 9AG. Email: S1461835@sms.ed.ac.uk

http://dx.doi.org/10.7565/ssp.2021.v4.5213

neuropsychologically informed risk factors, such as impulsivity and lack of insight, they are often assessed using collateral information and personal files, rather than validated performance measures. Moreover, meta-analytic findings have revealed that widely used VRA have low predictive validities, indicating that aside from the specific context in which the assessment will be used, there is no risk measure significantly *better* than another (Campbell et al., 2007; Desmarais et al., 2016; Yang et al., 2010). Further, Monahan and Skeem (2014) postulated that due to existing risk assessments using essentially the same factors and relying on self-report measures, only differing on how the factors are analysed, risk assessments have reached a natural limit, or a 'glass ceiling'. Thus, the addition of cognitive abilities to existing VRA, measured with validated neuropsychological tools, may have the ability to break the 'glass ceiling' and improve predictive accuracy, while subsequently identifying cognitive strengths and weaknesses of the individuals, and informing rehabilitation needs (Haarsma et al., 2020).

1.2 Violence risk assessments

Existing VRA range from unstructured to structured (Heilbrun, 2009; Skeem & Monahan, 2011), however, due to the inherent limitations of using unstructured clinical judgement alone (Skeem & Monahan, 2011), actuarial and structured professional judgement (SPJ) tools are recommended (Singh et al., 2014). There are several VRA with strong conceptual and empirical support for both violent and general offending in adults, youth, and psychiatric inpatients that fit within the actuarial and SPJ categories. For example, the Violence Risk Appraisal Guide (VRAG) (Harris et al., 1993), a well-validated actuarial measure, is designed to measure violent offending, the Historical, Clinical, Risk Scale (HCR-20) (Webster et al., 1997) is a risk-needs measure of violent offending, and is based on SPJ, and an actuarial risk-needs measure of general offending, as opposed to violent offending, is the Level of Service/Case Management Inventory (LS/CMI) (Andrews et al., 2006). Whilst these VRA are undoubtedly useful for the evaluation and prediction of risk, their predictive validity is rarely found to be more than 'modest' (Haque & Webster, 2013). A widely used method for investigating predictive validity is Area Under the Curve (AUC), which can span from 0 to 1, where .5 would be no better than chance that a violent offender would score high on a risk assessment over a non-violent offender, and 1 being near perfect accuracy (Singh, 2013). In 2011, Singh and colleagues investigated the predictive validity of nine risk assessments utilising AUC scores across 68 studies and found that the HCR-20, was only the fifth best predictor of violence with a pooled effect size of .70, with the highest measure being the Sexual Violence Risk Assessment-20 (Boer et al., 1997), revealing an AUC score of .78 (Haque & Webster, 2013; Singh et al., 2011). Within the same study, the VRAG resulted in AUC of .74, and the LSI-R was one of the lowest with an AUC of .67. Although an AUC score of .75 is interpreted as a large effect size, it appears that these assessments may benefit from more precise measures and specific domains.

1.3 Neuropsychology and violent behaviour

The cognitive abilities which have received the most attention in offending literature are executive functions (EFs), described as a constellation of higher level skills that aid in the control, regulation, and co-ordination of other cognitive abilities and behaviors, are controlled by the frontal lobes, and include working memory, poor inhibition, planning, response monitoring, and cognitive flexibility (Hoaken et al., 2007). Meta-analytic data has revealed a difference of 0.44 (Ogilvie et al., 2011) and 0.62 standard deviations (Morgan & Lilienfeld, 2000) between antisocial groups compared to non-antisocial controls on measures of EFs. While this group of cognitive abilities are highly interrelated, the failure to disaggregate them in research may inhibit the identification of specific neurocognitive mechanisms related to specific types of violence and offenders (Cruz et al., 2020). Thus, many have been investigated independently allowing for more specific examination of their relationship with violent behaviour. For instance, a 2017 study on inhibition, the ability to stop a mental process or action with or without trying to (MacLeod, 2007), found that violent prisoners performed significantly worse on the stop-signal task relative to non-violent prisoners with a partial correlation r = .20 (Meijers et al., 2017), and similarly, Kennedy and colleagues found that a measure of inhibition significantly differentiated violent and non-violent (d = 0.38) juvenile offenders referred for a court assessment (Kennedy et al., 2011). Likewise, Ross and Hoaken (2011) reported that inhibition in recidivist prisoners was more impaired relative to individuals who were in prison for the first time (partial η^2 = 0.07), and that response monitoring, defined as "evaluating the consequences of behaviour and making adjustments to optimise outcomes" (Thakkar et al., 2008, p. 2464), was also more impaired in recidivist prisoners (partial η^2= 0.05), though response monitoring has been less researched. However, both response monitoring and inhibition are core aspects of EFs, largely supported by the prefrontal cortex, which has been evidenced as one of the most significant brain structures to be compromised in violent and antisocial populations (Davidson et al., 2000; Raine & Yang, 2006), warranting a further investigation into the individual relationship between response monitoring and violence, especially prospectively. Attentional processes have also been implicated in violent behaviour, as they can affect learning, memory, and processing speed among many other cognitive abilities (Spreen & Strauss, 1998). Both correlational studies of forensic patients (Abidin et al., 2013) and between group studies of prisoners (Bryant et al., 1984) have found associations between attention deficits and violence, whereby violent offenders performed more poorly than non-violent offenders with an effects size of d= 0.83, and correlations with violent outcomes with a correlation of r= -.22. The relationship between attention and violence has largely been examined in inpatients, thus investigations of this construct in wider offending populations is warranted.

The diversity of EFs has been supported by Miyake and colleagues' EFs model (2000) which proposed that individuals can be impaired on a single executive domain and may not always have a general executive dysfunction. They supported this by demonstrating that three EFs (e.g., inhibition, shifting, and updating) individually relate to neuropsychological measures of frontal lobe functioning and IQ (Friedman et al., 2006). Given that individual EFs may be impaired while others are preserved, it is possible that the association between individual EFs and offending may differ between various subtypes of offenders and violence. Moreover, intelligence is correlated with EFs and neuroimaging studies have evidenced that intelligence and EFs depend on shared neural functioning (Barbey et al., 2012). Intelligence is a composite construct that in part comprises individual EFs (e.g. working memory) and the remainder is dependent on intact EFs for optimal performance (e.g. attention). Thus, intelligence may be a covariate in the relationship between EFs and violence, and given the significant overlap between intelligence in the form of full scale IQ and EFs, components of intelligence, namely, crystallized or fluid intelligence should be measured and controlled for in studies where it is appropriate.

Research has also highlighted impulsivity (Abidin et al., 2013; Coid et al., 2015; De Vogel & De Ruiter, 2006; Edwards et al., 2003; Howard et al., 2014; Zhou et al., 2014), risk taking (Lodewijks et al., 2008; Umbach et al., 2019), social cognition including affect recognition and empathy (Bock & Hosser, 2014; Brugman et al., 2016; Lodewijks et al., 2008; O'Reilly et al., 2015), and a lack of insight, specifically in mentally ill offenders (Alia-Klein et al., 2007; Bjørkly, 2006), as important neuropsychological abilities related to violent behaviour.

1.4 Neuropsychology and risk assessment

Based on the presented evidence, it can be hypothesized that validly measured neuropsychological domains have a place in VRA. Moreover, it is postulated that cognitive factors may add incremental validity. To our knowledge only one study, a doctoral thesis, specifically examined this. In 2015, LaDuke examined the potential for individual measures of EFs, such as attention, impulsivity, and verbal fluency, to add incremental validity to the LS/CMI (LaDuke, 2015). Findings revealed that, in a sample of prisoners (n= 95) recruited from a programme for treatment and re-entry services, only two of the measures (e.g., disinhibition and cognitive flexibility) predicted the outcome 'program failure', operationalized as returning to prison due to serious violations of programme rules. In addition, these measures not only significantly predicted program failure, but also demonstrated incremental validity over and above the LS/CMI independent of substance use. Moreover, trends toward significance were seen in another measure of disinhibition and measures of attention (LaDuke, 2015). Whilst this study implemented a rigorous prospective design, they were unable to measure violence as their outcome due to a low base rate of incidents.

Following the LaDuke (2015) study, to our knowledge, only one group of researchers have developed a risk assessment tool composed of cognitive factors. Haarsma and colleagues (2020) developed a mobile risk assessment to overcome the limitations of current VRA. They tested the tool on 730 probationers with an outcome of 'any new arrest', and findings revealed an AUC of .60 for cognitive performance alone, which is lower than existing risk assessments, however, when age, gender, and crime level were added to the model, the AUC score increased to .70, which is in line with existing assessments. While these findings are undoubtedly encouraging, limitations were detected. First, there is little explanation for how the tool domains were identified, aside from a literature review; second, the domains were seemingly not piloted before the development of the tool to ensure they individually and cumulatively predict re-offending; and third, with 400 instruments in use to assess, manage, and monitor violence risk (Singh et al., 2014), the development of new, stand-alone risk assessments is likely unnecessary, and a measure which complements an existing risk assessment may be more useful. Further, as evidenced in this study, the combination of cognitive risk factors and risk variables which are already comprised on risk measures, indicates the potential for additive value.

In sum, researchers have begun directly investigating the ways in which neurocognition can enhance the appraisal of risk, however, more research is necessary to overcome methodological limitations, and to evaluate the cumulative value of cognitive risk factors. Nonetheless, although the relationship between neurocognition and violence is largely characterized by small effect sizes, findings from Haarsma et al. (2020) are encouraging, though a gap remains for a tool which predicts violent outcomes. Moreover, VRA require the addition of new items to increase predictive accuracy and to break through the 'glass ceiling'. The current study aims to address some of these limitations by using rigorous methodology to identify neurocognitive domains and measures, administering validated cognitive measures, and by including inpatient and community offenders to increase generalizability.

1.5 Aims

The current study seeks to, (a) investigate the feasibility of using a neuropsychological battery to aid in the identification of violence risk in an inpatient and community setting, (b) pilot a neuropsychological battery of measures examining risk factors for violence identified through a meta-analysis and an international Delphi study (manuscripts in preparation), (c) identify neuropsychological measures that may improve the predictive accuracy of existing VRA, (d) identify cognitive impairments, measured using a valid neuropsychological tool, that explain the variance in violence risk in two groups of violent offenders, and (e) to identify cognitive abilities which predict inpatient vs community violence.

This study encompasses two distinct samples: forensic psychiatric inpatients (FPI) and community violent offenders (CVO). Currently, recruitment and baseline data collection are complete, with a total of n= 63 participants (e.g., FPI n= 32; CVO n= 31) and analyses have yet to commence.

1.6 Power calculation

A power calculation was conducted using G*Power 3.1 (Faul et al., 2007). To detect a medium effect size with 80% power and alpha set at 0.05 in a sample of n=32, a maximum of two predictors can be examined at one time, and a maximum of one predictor at time can be examined in the CVO group (n=31) to test primary and secondary hypotheses; consequently, the effects will not be cumulative. Covariates, including, crystallized intelligence, age, symptom severity (FPI only), lifetime history of TBI, and severity of TBI will be separately controlled for. Hypotheses will be examined in each group separately.

As the current study examined eight cognitive abilities, the following analysis plan has been developed a priori based on literature, meta-analyses, and a Delphi study to identify key cognitive domains which will be examined in our primary and secondary hypotheses. Where two predictors are allowed, attention and response inhibition will be examined, and attention only, where one predictor is allowed.

1.7 Primary hypotheses

Performance on measures of ***attention*** and ***response inhibition*** will:

1. Be significantly poorer in FPI relative to CVO.
2. Explain the variance in violent incidents/offences prospectively, over, and above traditional VRA.
3. Explain the variance in violent incidents (FPI) or violent offences (CVO) prospectively, where increased impairments will be associated with increased violence and violence severity.

1.8 Secondary hypotheses

Performance on measures of ***attention*** and ***response inhibition*** will explain the variance in perceived risk (FPI), antisocial behaviour (CVO), impulsivity, aggression, and violence severity, retrospectively and prospectively, where poorer performance will demonstrate an increase in these outcomes (Table 4), and will increase odds of perpetrating reactive violence relative to instrumental violence.

2. Methods/Design

2.1 Design

This is a feasibility study examining the viability of implementing a neuropsychological battery in forensic services to aid in the identification and assessment of violence risk, and a pilot study investigating the predictive and incremental ability of neuropsychological measures on risk-related outcomes. It follows a prospective, observational design with a retrospective component. All participants were tested at baseline, and outcomes were collected for 6-month prospective and retrospective follow-ups.

2.2 Participants

FPI were recruited from forensic mental health services, part of NHS Scotland, including high, medium, and low secure inpatient settings. CVO were recruited from Criminal Justice Social Work in Edinburgh, Scotland. Inclusion and exclusion criteria can be found in Table 1.

Table 1. Inclusion and exclusion criteria.

Group	Inclusion	Exclusion
FPI	• Male • 18-60 years old • Completed admission case review • Not being considered for transfer • Can give informed consent • English as their first language • Diagnosis of a psychotic illness	• Patients who have intellectual disabilities • Non-violent/non-contact sex offences or other offences without the presence of a violent offence • Disabilities which may impede their ability to engage in the assessment process (e.g., significant hearing, sight, and motor impairments)
CVO	• Male • 18-60 years old • Can give informed consent • English as their first language • Be under licence and therefore in the care of criminal justice social work for at least 6 months after testing • Has been convicted of a violent offence.	• Participants who have intellectual disabilities • No conviction of a violent offence • Disabilities which may impede their ability to engage in the assessment process

Note. CVO received £20.00 as reimbursement for their time spent participating in the study.

2.3 Data Collection and outcomes

A case-note review checklist was utilised to record data for the purpose of describing the sample. Historical and risk-related data, and a record of potential contributors to the aetiology of cognitive impairments were recorded from participant files where they were available (Table 2).

Table 2. Data collected from participant files.

Historical	Risk-Related	Potential Contributors [a]
•Age, ethnicity, education level/attainment •Date of current admission/ became service user •Source of current admission/services •Legislation for current detention (FPI only) •Current conviction status (if applicable) •Primary and secondary diagnosis (if applicable) •Year of first diagnosis with a psychotic illness (FPI only)	•Reason(s) for current admission/services •Offence(s) or alleged offence(s) leading to current admission/ services •Violent incidences during current admission/ services (if applicable) •Total number of previous convictions •Most serious previous offence •Previous types of offences	•Birth trauma •Abnormal infant development •Childhood history of physical or sexual abuse or neglect •Diagnosis/history of drug or alcohol dependence/misuse •Abnormal infant development •Neurological injuries or neurologically relevant diagnoses

Note. [a] This information was only available for a subset of CVO.

Additional data collected included:

Symptom Severity and Positive Symptoms (FPI Only*).* The Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987) will be used to measure symptom severity and positive symptoms. The PANSS is a routinely administered measure in some of the forensic psychiatry sites from which FPIs were recruited. When a PANSS was not available, the researcher completed the measure at the time of baseline testing. Symptom severity is operationalised as PANSS total score, and positive symptoms is operationalised as PANSS positive symptoms score.

Violence and Offending Proneness. The retrieval of routinely administered assessments (where available) included, the Historical, Clinical, Risk-20 Scale (HCR-20), Version 3 (Douglas et al., 2014) as a measure of violence proneness, and the Level of Service Inventory-Revised: Screening Version (LSI-R:SV) (Andrews & Bonta, 1995), and the Level of Service/Case Management Inventory (Andrews et al., 2006), as measures of offending proneness. Measure composite scores will be utilised for analyses.

Lifetime History and Severity of TBI. *The Ohio State University Traumatic Brain Injury Identification* (OSU-TBI) (Corrigan & Bogner, 2007) is a standardized procedure for learning about a person's lifetime history of traumatic brain injury. To measure lifetime history of head injury, number of TBIs with loss of consciousness was employed, and to measure TBI severity, the worst injury recorded was rated on a 5-point Likert scale ranging from, no history of TBI or very minor head injury to severe TBI.

Neuropsychological Battery. Development of the battery followed the National Institute of Mental Health's MATRICS initiative's development of the MATRICS Consensus Cognitive Battery (MCCB), which sought to identify not only domains identified by the research literature, but also those that have yet to be the subject of sufficient research, through expert consensus (Nuechterlein et al., 2006). The following measures were chosen based on a Delphi study and meta-analysis both examining cognitive predictors of violence. Operational definitions of independent variables (IV) are listed in Table 3.

Table 3. Operational definitions of independent variables.

IV	Measure	Operational Definition
Attention	Conners' Continuous Performance Test -Version 3 (Conners, 1995)	Number of omission errors
Response Inhibition	Stroop Colour and Word Test (Golden, 1978; Golden & Freshwater, 1978)	Interference score
Response Monitoring	Modified Wisconsin Card Sorting Test (Schretlen, 2010)	Number of perseverative errors
Risk Taking	Iowa Gambling Task-Version 2 (Bechara et al., 1994)	Proportion of cards chosen from disadvantageous card piles
Reasoning	Wechsler Abbreviated Scale of Intelligence, 2nd Edition (Wechsler, 2011)	Perceptual Reasoning Index (PRI)
Affect Recognition (negative and positive)	The Awareness of Social Inference Test (McDonald et al., 2003)	Number of correctly identified emotions
Cognitive Empathy	The Awareness of Social Inference Test (McDonald et al., 2003)	Number of correctly identified feelings
Lack of Insight	Positive and Negative Syndrome Scale (Kay et al., 1987)	Lack of Insight scale score
Crystallized Intelligence	Wechsler Abbreviated Scale of Intelligence, 2nd Edition (Wechsler, 2011)	Verbal Comprehension Index (VCI)

Note. Attention and response inhibition will be used to examine the primary and secondary hypotheses; the remaining variables will be subject to exploratory analyses. Variables are listed in the order in which they will be explored. Crystallized intelligence will serve as a covariate, and lack of insight will only be examined in the FPI group.

Primary and secondary outcomes and operational definitions can be found in Table 4.

Table 4. Primary outcomes and operational definitions.

	Outcome	Operationalisation
Primary Outcomes	Violent Incidents	Institution-recorded incidents of violence will be recorded and triangulated with patient notes to ensure consistency. *"An intentional act of physical aggression against another individual that is likely to cause physical injury"* (Meloy, 2006, p. 539), *where the individual is the clear aggressor or instigator*. Mean number of incidents over 6 months.
	Violent Offences & Charges [a]	Violent offences/ charges operationalised by the same definition as above. Mean number of offences/charges over 6 months.
Secondary Outcomes	Level of Observation (perceived risk)	Daily mean level of observation over 6-months (183 days) prospectively and retrospectively (FPI only).
	Days on Enhanced Levels (perceived risk)	Mean number of days on enhanced levels prospectively and retrospectively over 6 months (FPI only).
	Length of Admission (perceived risk)	Time in months that a FPI has been in their secure environment.
	Breaches of License/Conditions	Mean number of breaches of license/conditions prospectively and retrospectively over 6 months. Breaches will not be limited to violence and aggression only (CVO only).
	Antisocial Behaviour	Mean number of charges or convictions of non-violent/non-aggressive offences which occurred during the follow-up period in the CVO sample, over 6 months.
	Aggression	Mean number of non-contact aggressive incidents/offences in both samples (e.g., verbal aggression, damage to property, threats, racial aggression, carrying a weapon, intimidating behaviours) over 6 months.
	Impulsivity [b]	Total score of BIS-11 (Patton et al., 1995)
	Severity of violence & Aggression	Mean severity of all violent and aggressive incidents and offences combined over 6 months. Codes range from 0 (completely non-violent) to 4 (severe violence) (Gunn & Robertson, 1976; Robertson et al., 1987).
	Reactive vs Instrumental Violence	Median score of all violent incidents and offences over 6 months. Codes range from 1= Purely reactive to 4= Purely Instrumental (Woodworth & Porter, 2002).

Note. BIS= Barratt Impulsivity Scale. [a] When reporting results, a clear distinction will be made between charges and convictions. [b] As impulsivity is already a well-established risk factor for violence, it will serve as a proxy outcome measure in the current study.

3. Statistical Methods

3.1 Unity of samples

The distinct difference between the FPI and the CVO sample is the presence of major mental illness in FPI and inpatient vs community violence. Thus, all FPI will be analysed together first, and then between group tests will look for any significant differences on measures and demographic information between the FPI and CVO sample. If there are no significant differences between groups, all participants will be analysed in one model, using a dummy variable to indicate the two groups, though, due to the differences between outcome measures (inpatient vs community), the sample will only be combined to examine self-reported impulsivity. To examine hypotheses on the entire sample (n=63), and to detect a medium effect with 80% power, six variables can be examined in one regression.

3.2 Data analysis

Statistical analyses will be conducted using R: A language and environment for statistical computing (R Core Team, 2013).

3.3 Feasibility outcomes

Descriptive statistics will be reported for feasibility outcomes. The predictive accuracy of cognitive measures will be assessed utilising Receiver Operating Characteristic (ROC) analysis, and area under the curve (AUC). For this analysis, any participant who is violent during the follow-up period will be coded as 1, with 0 indicating the absence of violence.

3.4 Feasibility criteria

A future, larger study will be considered if the current study meets the following criteria:

1. Mean tolerability for a single measure must not fall below 3.5. A score of 3 or below indicates that the battery is unpleasant for participants. Scores ranging from 7-4.0 indicate extremely pleasant to neutral. If the mean tolerability rating for a single measure falls below 3.5, the reasons for this will be examined, and its replacement with a new measure for the same construct will be considered.
2. The mean recruitment rate for all sites combined must not fall below 30%. In the case that less than 30% of participants are recruited from all those eligible, inclusion criteria and recruitment procedures will be investigated. To keep the sample that is already recruited, only changes to the inclusion criteria which do not threaten the validity of the study will be considered for review. If this is not seen as a possible solution, the study will be deemed infeasible.
3. If the completion rate for the core measures (WASI-II, CPT-3, Stroop, MWCST, IGT-2, TASIT) does not reach 70%, reasons for non-completion will be reviewed and changes will be made to ensure the battery is feasible for these populations to complete.
4. If variables specified in our hypotheses do not show sufficient sensitivity, specificity, predictive accuracy or add incremental validity to existing risk measures, the measures used for this study will be re-considered using literature and further meta-analyses.

3.5 Primary and secondary outcomes

Patient characteristics will be reported using descriptive statistics. Both offending groups will be compared on baseline performance and participants who were violent will be compared to those who were not during the follow-up period, using t-tests or the non-parametric equivalent.

Linear regressions will be utilized to examine primary and secondary outcomes, with the exception of reactive and instrumental violence, in which ordinal logistic regressions will be used. If there are a sufficient number of recorded incidents or offences, the mean number of violent incidents/offences per month of the follow-up period will be used as a continuous dependent variable. Should there be a low base rate of incidents, those who were violent during follow-up will be coded as one, and those who were not will be coded as zero, and logistic regressions will be used. If the majority of the primary outcomes are not naturally binary, a two-part model for semicontinuous data will be considered. Two separate linear regressions will be run to examine the variance in violent outcomes explained by risk assessments and individual predictors. The proportion of explained variance (R^2) from each model will be used to discuss and compare the models.

4. Discussion

Existing research presents convincing evidence that there is indeed a relationship between cognitive functioning and violent behaviour, however methodological inconsistencies remain. Moreover, the accuracy of existing risk assessments continues to remain at a moderate level. The current protocol aims to identify cognitive domains related to violent and risk-related outcomes, and to examine their utility against existing risk assessments in two forensic samples.

4.1 Strengths

There are several noteworthy strengths in the current study. First, the neuropsychological battery was developed based on rigorous meta-analytic reviews, and an international Delphi study, and care was taken to consider the abilities of the targeted populations; second, all neurocognitive measures are well-validated, clinical tools increasing the reliability and validity of results, and allowing for the calculation of the proportion of participants who have clinically significant neurocognitive impairments relative to norms, facilitating use in clinical practice; third, individuals were recruited from high, medium, and low secure inpatient settings, in addition to violent offenders living in the community, increasing the generalisability of results; fourth, a coding tool to examine instrumental and reactive violence has been included to investigate factors which may be specific to types of violence; fifth, validated measures were included to measure key co-variates including TBI and symptom severity; and sixth, all variables have been clearly operationalised.

4.2 Limitations

This study is not without its limitations. Like other observational studies in this field, there is a risk that the base rate of observational violence will be low, as a result of FPI being in intensive treatment and stable environments, and low level offences that may not be recorded as charges or convictions in CVO, are common within these populations, which may impede our ability to examine our primary hypotheses. To counter this, we have included secondary outcome measures, which are indicative of perceived violence risk, and it is hoped that these will serve as a suitable criterion variable in the case of low adverse incidences in both FPI and CVO samples. Also, this is a field which has struggled to obtain larger samples, given that antisocial personality disorder and other psychiatric illnesses are common among these populations, thus this has been designed as a feasibility and pilot study to inform a larger trial. Conducting pilot studies can answer questions about feasibility, recruitment potential, and patient acceptability, as well as, sourcing a justification for larger, future studies, and can provide best practice guidelines for logistical issues for a larger study. Moreover, conducting a pilot in advance of a primary study can increase the potential for a successful future study (Thabane et al., 2010). Accordingly, to remain transparent, we have pre-emptively chosen variables which were supported by the literature and expert opinion to test the primary and secondary hypotheses of the current study. Finally, the findings will not be generalizable to other sexes, as only men were recruited.

5. Conclusion

It is our hope that the results of this study and, if indicated, a larger future study, will contribute to the development of a structured professional tool to aid in the identification and assessment of cognitive impairments shown to be predictive of violence risk, our knowledge of the predictive utility of measures, and to determine whether the addition of the assessment battery increases the predictive accuracy of existing measures of risk and/or positively contributes to the formulation of offending risk.

Declarations

Conflict of interests: The authors declare that they have no conflict of interest.

Funding Sources: This study is part of a Ph.D. funded by the Principal's Career Development Scholarship and the Edinburgh Global Research Scholarship from the University of Edinburgh. A portion of the study was funded by the State Hospital Research Committee in Lanarkshire, UK.

Ethics: This study obtained ethical approval by the West of Scotland NHS Ethics Committee (Reference: 17/WS/0189).

References

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry*, *13*(1), 197. https://doi.org/10.1186/1471-244X-13-197

Alia-Klein, N., O'Rourke, T. M., Goldstein, R. Z., & Malaspina, D. (2007). Insight into illness and adherence to psychotropic medications are separately associated with violence severity in a forensic sample. *Aggressive Behavior*, *33*(1), 86–96. https://doi.org/10.1002/ab.20170

Andrews, D. ., & Bonta, J. (1995). *The level of supervision inventory-revised* (Vol. 106, pp. 19–52). Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Journals.Sagepub.Com*, *17*(1), 7. https://doi.org/10.1177/0011128705281756

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain, 135*(4), 1154–1164. https://doi.org/10.1093/brain/aws021

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50*(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bjørkly, S. (2006). Empirical evidence of a relationship between insight and risk of violence in the mentally ill - A review of the literature. *Aggression and Violent Behavior, 11*(4), 414–423. https://doi.org/10.1016/j.avb.2006.01.006

Bock, E. M., & Hosser, D. (2014). Empathy as a predictor of recidivism among young adult offenders. *Psychology, Crime and Law, 20*(2), 101–115. https://doi.org/10.1080/1068316X.2012.749472

Boer, D., Hart, S., Kropp, P., & Webster, C. (1997). Manual for the sexual violence risk-20. Professional guidelines for assessing risk of sexual violence. The Mental Health. *Law & Policy Institute.*

Brugman, S., Lobbestael, J., von Borries, A. K. L., Bulten, B. E. H., Cima, M., Schuhmann, T., Dambacher, F., Sack, A. T., & Arntz, A. (2016). Cognitive predictors of violent incidents in forensic psychiatric inpatients. *Psychiatry Research, 237*, 229–237. https://doi.org/10.1016/j.psychres.2016.01.035

Bryant, E. T., Scott, M. L., Golden, C. J., & Tori, C. D. (1984). Neuropsychological deficits, learning disability, and violent behavior. *Journal of Consulting and Clinical Psychology, 52*(2), 323–324.

Campbell, M. A., French, S. A., & Gendreau, P. (2007). *Assessing the utility of risk assessment tools and personality measures in the prediction of violent revidivism for adult offenders.* Public Safety Canada. https://www.securitepublique.gc.ca/cnt/rsrcs/pblctns/rsk-ssmnt-tls/rsk-ssmnt-tls-eng.pdf

Coid, J. W., Kallis, C., Doyle, M., Shaw, J., & Ullrich, S. (2015). Identifying causal risk factors for violence among discharged patients. *PLoS ONE, 10*(11). https://doi.org/10.1371/journal.pone.0142493

Conners, C. (1995). *Conners' continuous performance test.* Multi-Health Systems.

Corrigan, J. D., & Bogner, J. (2007). Screening and Identification of TBI. *Journal of Head Trauma Rehabilitation, 22*(6), 315–317. https://doi.org/10.1097/01.HTR.0000300226.67748.3e

Cruz, A. R., de Castro-Rodrigues, A., & Barbosa, F. (2020). Executive dysfunction, violence and aggression. *Aggression and Violent Behavior, 51*, 101380. https://doi.org/10.1016/j.avb.2020.101380

Davidson, R. J., Putnam, K. M., & Larson, C. L. (2000). Dysfunction in the neural circuitry of emotion regulation - A possible prelude to violence. Science, *289*(5479), 591–594. American Association for the Advancement of Science. https://doi.org/10.1126/science.289.5479.591

De Vogel, V., & De Ruiter, C. (2006). Structured professional judgment of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime and Law, 12*(3), 321–336. https://doi.org/10.1080/10683160600569029

Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in US correctional settings. *Psycnet.Apa.Org.* https://doi.org/10.1037/ser0000075

Douglas, K. S., Hart, S. D., Webster, C. D., Belfrage, H., Guy, L. S., & Wilson, C. M. (2014). Historical-Clinical-Risk Management-20, Version 3 (HCR-20 V3): Development and Overview. *International Journal of Forensic Mental Health, 13*(2), 93–108. https://doi.org/10.1080/14999013.2014.906519

Edwards, D. W., Scott, C. L., Yarvis, R. M., Paizis, C. L., & Panizzon, M. S. (2003). Impulsiveness, impulsive aggression, personality disorder, and spousal violence. *Violence and Victims, 18*(1), 3–14. https://doi.org/10.1891/vivi.2003.18.1.3

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Golden, C. (1978). *A manual for the clinical and experimental use of the Stroop color and word test*. https://nsuworks.nova.edu/cps_facbooks/47/

Golden, C., & Freshwater, S. (1978). *Stroop color and word test*. http://v-psyche.com/doc/Clinical Test/Stroop Color and Word Test.doc

Gunn, J., & Robertson, G. (1976). Drawing a Criminal Profile. *British Journal of Criminology, 16*. https://heinonline.org/HOL/Page?handle=hein.journals/bjcrim16&id=168&div=21&collection=journals

Haarsma, G., Davenport, S., White, D. C., Ormachea, P. A., Sheena, E., & Eagleman, D. M. (2020). Assessing risk among correctional community probation populations: Predicting reoffense with mobile neurocognitive assessment software. *Frontiers in Psychology, 10*, 2926. https://doi.org/10.3389/fpsyg.2019.02926

Haque, Q., & Webster, C. D. (2013). Structured professional judgement and sequential redirections. *Criminal Behaviour and Mental Health, 23*(4), 241–251. https://doi.org/10.1002/cbm.1886

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*(4), 315–335. https://doi.org/10.1177/0093854893020004001

Heilbrun, K. (2009). *Evaluation for risk of violence in adults*. Oxford University Press, Inc.

Hoaken, P. N. S., Allaby, D. B., & Earle, J. (2007). Executive cognitive functioning and the recognition of facial expressions of emotion in incarcerated violent offenders, non-violent offenders, and controls. *Aggressive Behavior, 33*(5), 412–421. https://doi.org/10.1002/ab.20194

Howard, R. C., Khalifa, N., & Duggan, C. (2014). Antisocial personality disorder comorbid with borderline pathology and psychopathy is associated with severe violence in a forensic sample. *Journal of Forensic Psychiatry and Psychology, 25*(6), 658–672. https://doi.org/10.1080/14789949.2014.943797

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Kennedy, T. D., Burnett, K. F., & Edmonds, W. A. (2011). Intellectual, behavioral, and personality correlates of violent vs. non-violent juvenile offenders. *Aggressive Behavior, 37*(4), 315–325. https://doi.org/10.1002/ab.20393

LaDuke. (2015). *Can neuropsychology inform violence risk assessment?* [Drexel University]. https://idea.library.drexel.edu/islandora/object/idea%3A6618/datastream/OBJ/download/Can_neuropsychology_inform_violence_risk_assessment_.pdf

Lodewijks, H. P. B., Doreleijers, T. A. H., de Ruiter, C., & Borum, R. (2008). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry*, *31*(3), 263–271. https://doi.org/10.1016/j.ijlp.2008.04.009
MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. S. Gorfein & C. M. Macleod (Eds.), *Inhibition in cognition* (pp. 3–23). American Psychological Association. https://doi.org/10.1134/s1019331609030137
McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. The Journal of head trauma rehabilitation, 18(3), 219-238. https:doi.org/10.1097/00001199-200305000-00001
Meijers, J., Harte, J. M., Meynen, G., & Cuijpers, P. (2017). Differences in executive functioning between violent and non-violent offenders. *Psychological Medicine*, *47*(10), 1784–1793. https://doi.org/10.1017/S0033291717000241
Meloy, J. R. (2006). Empirical basis and Forensic aplication of affective and predatory Violence. *Australian & New Zealand Journal of Psychiatry*, *40*(6–7), 539–547. https://doi.org/10.1080/j.1440-1614.2006.01837.x
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "'frontal lobe'" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734
Monahan, J., & Skeem, J. L. (2014). Risk redux: The resurgance of risk assessment in criminal sanctioning. *Federal Sanctioning Reporter*, *26*(3), 158–166. https://doi.org/10.1525/fsr.2014.26.3.158
Morgan, & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, *20*(1), 113–136. https://doi.org/10.1016/S0272-7358(98)00096-8
Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., Essock, S., Fenton, W. S., Frese, F. J., Gold, J. M., Goldberg, T., Heaton, R. K., Keefe, R. S., Kraemer, H., Mesholam-Gately, R., Seidman, L. J., Stover, E., Weinberger, D. R., Young, A. S., … Marder, S. R. (2006). The MATRICS Consensus Cognitive Battery, part 1: Test selection, reliability, and validity. American Journal of Psychiatry, 165(2), 203–213. https://doi.org/10.1176/appi.ajp.2007.07010042
O'Reilly, K., Donohoe, G., Coyle, C., O'Sullivan, D., Rowe, A., Losty, M., McDonagh, T., McGuinness, L., Ennis, Y., Watts, E., Brennan, L., Owens, E., Davoren, M., Mullaney, R., Abidin, Z., & Kennedy, H. G. (2015). Prospective cohort study of the relationship between neuro-cognition, social cognition and violence in forensic patients with schizophrenia and schizoaffective disorder. *BMC Psychiatry*, *15*(1). https://doi.org/10.1186/s12888-015-0548-0
Ogilvie, J. M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. *Criminology*, *49*(4), 1063–1107. https://doi.org/10.1111/j.1745-9125.2011.00252.x
Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*, 768–774. http://www.impulsivity.org/measurement/bis11
R Core Team. (2013). *R: A language and environment for statistical computing*.
Raine, A., & Yang, Y. (2006). Neural foundations to moral reasoning and antisocial behavior. *Social Cognitive and Affective Neuroscience*, *1*(3), 203–213. https://doi.org/10.1093/scan/nsl033
Robertson, G., Taylor, P. J., & Gunn, J. C. (1987). Does violence have cognitive correlates? *British Journal of Psychiatry*, *151*(JULY), 63–68. https://doi.org/10.1192/bjp.151.1.63

Ross, E. H., & Hoaken, P. N. S. (2011). Executive cognitive functioning abilities of male first time and return Canadian federal inmates. *Canadian Journal of Criminology and Criminal Justice*, *53*(4), 377–403. https://doi.org/10.3138/cjccj.53.4.377

Schretlen, D. (2010). *Modified Wisconsin Card Sorting Test®: M-WCST; Professional Manual*. PAR.

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences and the Law*, *31*(1), 8–22. https://doi.org/10.1002/bsl.2052

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., Doyle, M., Folino, J. O., Godoy-Cervera, V., Grann, M., Ho, R. M. Y., Large, M. M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., de Ruiter, C., Seewald, K., & Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 vountries. *International Journal of Forensic Mental Health*, *13*(3), 193–206. https://doi.org/10.1080/14999013.2014.922141

Singh, J. P., Serper, M., Reinharth, J., & Fazel, S. (2011). Structured assessment of violence risk in schizophrenia and other psychiatric disorders: A systematic review of the validity, reliability, and item content of 10 available instruments. *Schizophrenia Bulletin*, *37*(5), 899–912. https://doi.org/10.1093/schbul/sbr093

Skeem, J. L., & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, *20*(1), 38–42. https://doi.org/10.1177/0963721410397271

Spreen, O., & Strauss, E. (1998). *A Compendium of Neuropsychological Tests* (2nd ed.). Oxford University Press.

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., Thabane, M., Giangregorio, L., & Goldsmith, C. H. (2010). A tutorial on pilot studies: The what, why and how. *BMC Medical Research Methodology*, *10*. https://doi.org/10.1186/1471-2288-10-1

Thakkar, K. N., Polli, F. E., Joseph, R. M., Tuch, D. S., Hadjikhani, N., Barton, J. J. S., & Manoach, D. S. (2008). Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD). *Brain*, *131*(9), 2464–2478. https://doi.org/10.1093/brain/awn099

Umbach, R., Leonard, N. R., Luciana, M., Ling, S., & Laitner, C. (2019). The Iowa gambling task in violent and nonviolent incarcerated male adolescents. *Criminal Justice and Behavior*, *46*(11), 1611–1629. https://doi.org/10.1177/0093854819847707

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). HCR–20: assessing risk for violence. In *Mental Health Law and Policy Institute*.

Wechsler, D. (2011). WASI-II: Wechsler abbreviated scale of intelligence. *PsychCorp*.

Woodworth, M., & Porter, S. (2002). In cold blood: Characteristics of criminal homicides as a function of psychopathy. *Journal of Abnormal Psychology*, *111*(3), 436–445. https://doi.org/10.1037/0021-843X.111.3.436

Yang, M., Wong, S. C. P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*(5), 740–767. https://doi.org/10.1037/a0020473

Zhou, J., Witt, K., Zhang, Y., Chen, C., Qiu, C., Cao, L., & Wang, X. (2014). Anxiety, depression, impulsivity and substance misuse in violent and non-violent adolescent boys in detention in China. *Psychiatry Research*, *216*(3), 379–384. https://doi.org/10.1016/j.psychres.2014.01.024

Appendix B: Table of Measures Used in Included Studies in Systematic Review

Cognitive domains and measures with corresponding study

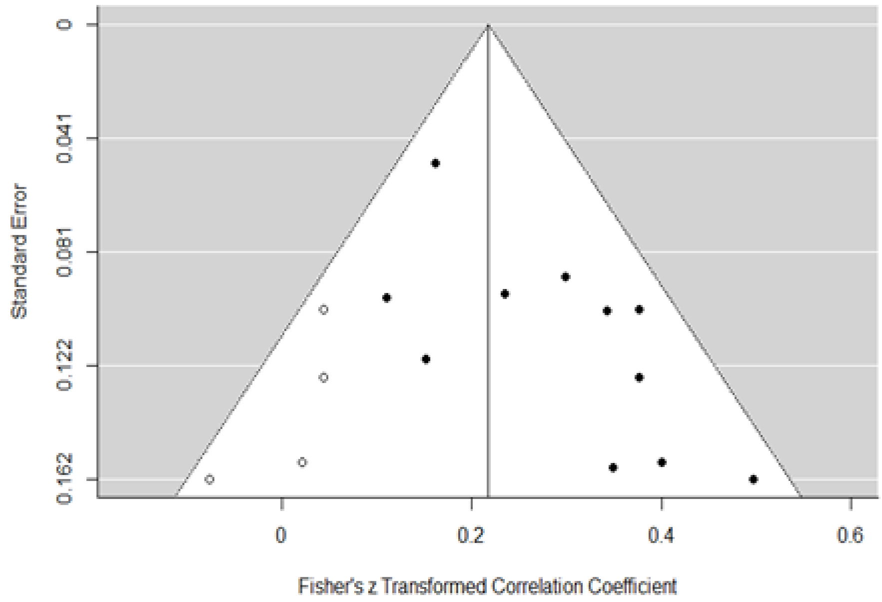
Domain	Measure	Study
Intelligence	Wechsler Intelligence Scales	Beggs & Grace, 2008; Brimigion et al., 2014; Busch et al., 1990; Cornell & Wilson, 1992; Dejong et al., 1992; Feichtinger, 2007; Fullam & Dolan, 2008; Gretton, 1998; Hays et al., 1978; Howard et al., 2014; Meijers et al., 2017; Nazmie et al., 2013; Tarter et al., 1983; Umbrasas, 2019
	Test of Non-Verbal Intelligence	Greenfield et al., 2007
	Vienna Matrices Test	Bock et al., 2014
	Raven Standard Progressive Matrices	Kuin et al., 2017; Rimmer, 1998
Verbal Comprehension	Wechsler Intelligence Scales	Brimigion et al., 2014; Busch et al., 1990; Fullam & Dolan, 2008; Tarter et al., 1983; Umbrasas, 2018
	Luria Nebraska Neuropsychological Battery	Bryant et al., 1984
	Token Test	Ullman, 1988
	Peabody Picture Vocabulary	Kennedy et al., 2011
Memory	Luria Nebraska Neuropsychological Battery	Bryant et al., 1984
	Memory Scale- Neuropsychological Impairment Score	Rimmer, 1998
	Wechsler Memory Scales	Tarter et al., 1983
	California Verbal Learning Test (long delay recall)	Feichtinger, 2007
Working Memory	Luria Nebraska Neuropsychological Battery	Bryant et al., 1984
	Wechsler Intelligence Scales	Brimigion et al., 2014; Busch et al., 1990; Tarter et al., 1983
	MATRICS Cognitive Consensus Battery	O’Reilly et al., 2015
	Cambridge Automated Neuropsychological Test Battery- Spatial Working Memory Task	Meijers et al., 2017
Processing Speed	Wechsler Adult Intelligence Scales	Brimigion et al., 2014; Busch et al., 1990; Tarter et al., 1983
	MATRICS Cognitive Consensus Battery	O’Reilly et al., 2015
	Trail Making Test B/A	Kuin et al., 2017
Reasoning	Wechsler Intelligence Scales	Brimigion et al., 2014; Busch et al., 1990; Fullam & Dolan, 2008; Tarter et al., 1983; Umbrasas, 2018
	MATRICS Cognitive Consensus Battery	O’Reilly et al., 2015

Response Inhibition	Colour-Word interference	Feichtinger, 2007
	Stop Probability Tasks	Fullam & Dolan, 2008
	BACS inhibition subtest	Kennedy et al., 2011
	Stop Signal Task- CANTAB	Meijers et al., 2017
Impulsivity	BIS-11	Duwors, 1998; Edwards et al., 2003; McDermott et al., 2008; McKee, 2004; Zhou et al., 2014
	Impulsiveness Scale- Short Form	Chan et al., 2012
	HCR-20	Abidin et al., 2013; Belfrage et al., 2000; Coid et al., 2015; De Vogel & De Ruiter, 2011
	UPPS Impulsivity Scale	Howard et al., 2014
	Psychopathic Personality Inventory	Smith et al., 2013
	BIS-11/Balloon Analogue Task/17 Venturesomeness Scale	Tonnaer et al., 2016
	Structured Assessment of Violence Risk in Youth	Lodewijks et al., 2008
VIQ	Wechsler Intelligence Scales	Busch et al., 1990; Cornell & Wilson, 1992; Feichtinger, 2007; Gretton, 1998; Hays et al., 1978; Nazmie et al., 2013; Tarter et al., 1983
	Wide Range Achievement Test-3	Umbach et al., 2019
Performance IQ	Wechsler Adult Intelligence Scales	Cornell & Wilson, 1992; Feichtinger, 2007; Gretton, 1999; Hays et al., 1978; Tarter et al., 1983
Attention	Luria Nebraska Neuropsychological Battery	Ullman, 1988; Bryant et al., 1984
	MATRICS Cognitive Consensus Battery	O'Reilly et al., 2015
	Neuropsychological Impairment Scale	Rimmer, 1998
	Positive and Negative Syndrome Scale	Abidin et al., 2013
	Test of Attentional Performance	Tonnaer et al., 2016
	Trail Making Test- A	Kuin et al., 2017; Nazmie et al., 2013
	Detroit Test of Learning Aptitude	Tarter et al., 1983
	Choice-Reaction Time Task	Meijer et al., 2017
Insight	HCR-20	Belfrage et al., 2000; De Vogel & De Ruiter, 2011; Abidin et al., 2013; Coid et al., 2015
	Positive and Negative Syndrome Scale	Abidin et al., 2013
	Scale to Assess Unawareness in Mental Disorder in Schizophrenia	Alia-Klein et al., 2013
Empathy	Empathy Skills Questionnaire, Empathy Measure-Adult Version	Moulden, 2009
	Interpersonal Reactivity Index	Nigel et al., 2018, Goldstein et al., 2001
	Structured Assessment of Violence Risk in Youth	Lodewijks et al., 2008
	Interpersonal Reactivity Index	Goldstein et al., 2001

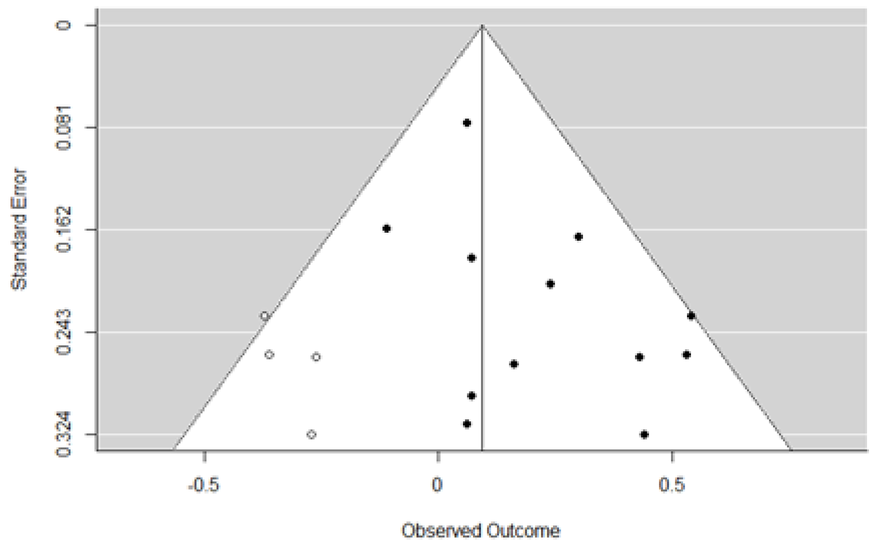
Cognitive Flexibility	Trail Making Test B	Nazmie et al., 2013; Rimmer, 1998
	Wisconsin Card Sorting Test	Ullman, 1988
	Test of Attentional Performance	Tonnaer et al., 2016
	DKEFS trails and sorting	Feichtinger, 2007
	CANTAB	Fullam & Dolan, 2008; Meijers et al., 2017
Motor Skills	Luria Nebraska Neuropsychological Battery	Bryant et al., 1984
	Bender Visual Motor Gestalt Test	Busch et al., 1990
	Finger Tapping, Purdue Pegboard, Star Tracing	Tarter et al., 1983
Planning	Stockings of Cambridge	Fullam & Dolan, 2008; Meijers et al., 2017
	Picture Arrangement	Busch et al., 1990; Tarter et al., 1983
	Means End Problem Solving Procedure	Rimmer, 1998
	DKEFS Tower Test	Feichtinger, 2007
	Porteus Maze	Greenfield et al., 2007

Appendix C: Trim-and-Fill Funnel Plots

Trim-and-Fill funnel plot for Impulsivity



Trim-and-Fill funnel plot for Intelligence



Note. Open circles on the left represent missing studies.

Appendix D: Concepts and Search Terms for Systematic Review

Concepts and Search terms

Cognitive deficits	“Cognitive deficit*”, Neuropsych*, “neuropsych* deficit*”, “cognit* psychia*”, “brain injury”, “head injury*”, “cognit* defect*”, neurocognitive, “neurocognitive deficits”, TBI, “traumatic brain injury”, intel*, memor*, “executive funct*”, “executive cognitive funct*”, atten*, emotion*, “facial affect recog*”, impuls*, “social cognition”, “theory of mind”, neurophysiological, comprehension, intention, learning, perception, insight, think*, volition, “cognitive inhibition”, empathy, “affective empathy”, inhibit*
Violence	Violen*, hostil*, aggress*, homicid*, murder*, rape, “child abuse”, abuse, paedophil*, pedophil*, “sex* violence”, maltreatment
Offenders	“violen* offend*”, offend*, prison*, delinquent, inmate, convict, felon, “sex offend*”, rapist
Mental Illness	“ment* disorder*” or “ment* ill*” or sychiatry* or “forensic* sychiatry*” or “state hospital”, secure, Schizophreni*, “antisocial personality disorder”, APD, psychopathy, ADHD, psychosis, psychotic, “schizoaffective disorder”, “bipolar disorder”, “first episode psychosis”
Risk factors	risk, contributor*, predict*, recidiv*, relapse, “re-offend*”

Appendix E: Example of Search for Systematic Review

Example of search from PsycInfo

1. cognitive impairment/ or brain damage/ or cognitive ability/ or dysexecutive syndrome/ or intellectual development disorder/ or memory disorders/ or thought disturbances/
2. "cognitive deficit".mp.
3. neurocognition/ or cognitive processes/ or neuropsychology/
4. "neuropsych* deficit*".mp.
5. head injuries/ or brain concussion/ or brain damage/ or traumatic brain injury/
6. executive function/ or cognitive control/ or set shifting/ or task switching/
7. attention/
8. learning/ or memory/
9. learning disorders/
10. emotions/ or emotion recognition/
11. exp social cognition/
12. "facial affect recognition".mp.
13. impulsiveness/ or cognitive style/ or attention deficit disorder/ or attention deficit disorder with hyperactivity/ or behavioral disinhibition/ or impulse control disorders/
14. impuls*.mp.
15. "theory of mind"/ or cognitive development/ or comprehension/ or mentalization/ or mind/ or social perception/
16. exp intention/
17. neurophysiology/ or neurosciences/
18. insight/ or personality processes/ or intuition/ or "perceptiveness (personality)"/
19. exp THINKING/
20. exp VOLITION/
21. "cognitive inhibition".mp.
22. exp EMPATHY/
23. "affective empathy".mp.
24. "cognitive empathy".mp.
25. inhibit*.mp.
26. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
27. violen*.mp. or exp Aggressive Behavior/
28. homicide/ or behavior disorders/ or violent crime/ or filicide/ or mass murder/ or serial homicide/ or infanticide/
29. exp HOSTILITY/
30. violence/ or antisocial behavior/ or conflict/ or domestic violence/ or intimate partner violence/ or patient violence/ or school violence/ or violent crime/ or coercion/ or dangerousness/ or physical abuse/
31. rape/ or sexual abuse/
32. child abuse/ or child neglect/ or emotional abuse/ or pedophilia/ or physical abuse/ or sexual abuse/ or verbal abuse/
33. (pedophil* or paedophil*).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
34. abuse.mp.
35. "sex violence".mp.

36. maltreatment.mp.

37. 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36

38. criminals/ or perpetrators/ or female criminals/ or male criminals/ or mentally ill offenders/ or juvenile delinquency/ or prisoners/

39. offend*.mp.

40. (inmate or convict).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

41. exp Sex Offenses/

42. "sex offender".mp.

43. rapist.mp.

44. 38 or 39 or 40 or 41 or 42 or 43

45. mental disorders/ or personality disorders/ or exp psychosis/ or exp schizoaffective disorder/

46. "mental* ill*".mp.

47. forensic psychiatry/ or psychiatry/ or forensic psychology/

48. psychiatric hospitals/ or psychiatric units/ or sanatoriums/

49. "state hospital".mp.

50. "secure hospital".mp.

51. exp SCHIZOPHRENIA/

52. "first episode psychosis".mp.

53. antisocial personality disorder/ or autism spectrum disorders/ or psychopathy/

54. exp Bipolar Disorder/

55. 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54

56. exp Risk Factors/

57. contributor*.mp.

58. predict*.mp.

59. exp RECIDIVISM/

60. relapse.mp.

61. re-offend*.mp.

62. 56 or 57 or 58 or 59 or 60 or 61

63. 26 or 55

64. 37 and 44 and 62 and 63

Appendix F: Ethical Approval for Delphi Study

School of Health in Social Science
Clinical and Health Psychology

The University of Edinburgh
Medical School
Doorway 6, Teviot Place
Edinburgh EH8 9AG

Telephone 0131 651 3969
Fax 0131 650 3891
Email submitting.ethics@ed.ac.uk

Sarah Janes
PhD Student, Clinical Psychology
School of Health in Social Science
University of Edinburgh

01 July 2016

Dear Sarah,

Application for Level 1 Ethical Approval

Reference: CLIN292
Project Title: Assessing the Cognitive Contributors to Violence Risk
Academic Supervisor: Suzanne O'Rourke / Matthias Schwannauer

Thank you for submitting the above research project for review by the Department of Clinical and Health Psychology Ethics Research Panel. I can confirm that the submission has been independently reviewed and was approved on the 28th June 2016.

Should there be any change to the research protocol it is important that you alert us to this as this may necessitate further review.

Yours sincerely,

Kirsty Gardner
Administrative Secretary, Clinical Psychology

Appendix G: REC Approval for Pilot and Feasibility Study

WoSRES
West of Scotland Research Ethics Service

NHS
Greater Glasgow and Clyde

Mrs. Sarah Janes
PhD Student, Clinical Psychology
University of Edinburgh
School of Health and Social Science
Medical School
Doorway 6,TeviotPlace
Edinburgh EH89AG

West of Scotland REC 3
West of Scotland Research Ethics Service
West Glasgow Ambulatory Care Hospital
(former Royal Hospital for Sick Children Yorkhill)
Dalnair Street
Glasgow G3 8SJ
www.nhsggc.org.uk

Date 29th September 2017
Your Ref
Our Ref
Direct line 0141 232 1805
E-mail WOSREC3@ggc.scot.nhs.uk

Dear Mrs Janes

Study title: **Cognitive Contributors to Risk of Harm to Others: A Feasibility Study**
REC reference: **17/WS/0189**
IRAS project ID: **231852**

Thank you for your submitting your response to the conditions letter. I can confirm the REC has received the documents listed below and that these comply with the approval conditions detailed in our letter dated 01 September 2017

Documents received

The documents received were as follows:

Document	*Version*	*Date*
Participant consent form [Consent Form]	2.0	25 September 2017
Participant information sheet (PIS) [PIS V2.0]	2.0	25 September 2017
Research protocol or project proposal [Proposal V2.1]	2.1	25 September 2017

Approved documents

The final list of approved documentation for the study is therefore as follows:

Document	*Version*	*Date*
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [University of Edinburgh Professional Indemnity]	1	
GP/consultant information sheets or letters [RMO letter/information sheet]	1.0	01 August 2017
IRAS Application Form [IRAS_Form_04082017]		04 August 2017
Letter from funder [Letter from Funder]	1	21 July 2017
Other [University of Edinburgh Clinical Trials Liability]		
Other [EL Certificate]		

Other [PL Confirmation]		
Other [GCP Training Certificate]		21 July 2017
Other [ACL Training Certificate]		21 July 2017
Other [Eligibility Criteria Checklist]	1.0	01 August 2017
Participant consent form [Consent Form]	2.0	25 September 2017
Participant information sheet (PIS) [PIS V2.0]	2.0	25 September 2017
Referee's report or other scientific critique report [Scientific Critique from State Hospital]	1	15 May 2017
Research protocol or project proposal [Proposal V2.1]	2.1	25 September 2017
Summary CV for Chief Investigator (CI) [CV]		01 August 2017
Summary CV for supervisor (student research) [CV]	1	01 August 2017

You should ensure that the sponsor has a copy of the final documentation for the study. It is the sponsor's responsibility to ensure that the documentation is made available to R&D offices at all participating sites.

17/WS/0189	**Please quote this number on all correspondence**

Yours sincerely

Liz Jamieson

Liz Jamieson
REC Manager

Copy to: Ms Charlotte Smith, University of Edinburgh
Mr Jamie Pitcairn, The State Hospital

Appendix H: Management Approval for Pilot and Feasibility Study

The State Hospital

Carstairs
Lanark ML11 8RP
Telephone 01555 840293
Fax 01555 840024
E-mail info@tsh.nhs.uk
http://www.tsh.scot.nhs.uk

NHS SCOTLAND

Ms Sarah Janes
Edinburgh University

Sent via e-mail

Date 15 September 2017
Your Ref
Our Ref DA/jm

Enquiries to Jacqueline McDade
Direct Line 01555 842013
E-mail jmcdade@nhs.net

Dear Ms Janes

Re: Cognitive Contributors to Risk of Harm to Others: A Feasibility Study

Having considered the views of the Research Committee and noted that you have a letter of favourable opinion from the West of Scotland Research Ethics Service, I write to give you Managerial Approval to proceed with your project. This is subject to you fulfilling the requirements of the State Hospital Research Committee.

May I take this opportunity to wish you every success in your endeavour.

Yours sincerely

DAlcock

DR DUNCAN ALCOCK
Associate Medical Director

cc Jamie Pitcairn, Research and Development Manager.
Professor Lindsay Thomson, Medical Director.

Appendix I: Letter of Access for NHS Glasgow for Pilot and Feasibility Study

NHS
Greater Glasgow and Clyde

Administrator: Mrs Elaine O'Neill
Telephone Number: 0141 232 1815
E-Mail: elaine.o'neill2@ggc.scot.nhs.uk
Website: www.nhsggc.org.uk/r&d

R&D Management Office
West Glasgow ACH
Dalnair Street
Glasgow G3 8SW

24 July 2018

Dr Emma Drysdale
West Glasgow ACH
Dalnair Street
Glasgow G3 8SJ

NHS GG&C Board Approval

Dear Dr E Drysdale,

Study Title:	**Cognitive Contributors to Risk of Harm to Others: A Feasibility Study**
Principal Investigator:	**Dr Emma Drysdale**
GG&C HB site	**Rowanbank Clinic**
Sponsor	**The University of Edinburgh**
R&D reference:	**GN18MH417**
REC reference:	**17/WS/0189**
Protocol no:	**V5.0; 05/07/18**

I am pleased to confirm that Greater Glasgow & Clyde Health Board is now able to grant **Approval** for the above study.

Conditions of Approval

1. **For Clinical Trials** as defined by the Medicines for Human Use Clinical Trial Regulations, 2004

   a. During the life span of the study GGHB requires the following information relating to this site
      i. Notification of any potential serious breaches.
      ii. Notification of any regulatory inspections.

It is your responsibility to ensure that all staff involved in the study at this site have the appropriate GCP training according to the GGHB GCP policy (www.nhsggc.org.uk/content/default.asp?page=s1411), evidence of such training to be filed in the site file.

2. **For all studies** the following information is required during their lifespan.
   a. Recruitment Numbers on a monthly basis
   b. Any change of staff named on the original SSI form
   c. Any amendments – Substantial or Non Substantial
   d. Notification of Trial/study end including final recruitment figures
   e. Final Report & Copies of Publications/Abstracts

Please add this approval to your study file as this letter may be subject to audit and monitoring.

Your personal information will be held on a secure national web-based NHS database.

I wish you every success with this research study

Yours sincerely,

Mrs Elaine O'Neill
Senior Research Administrator

Cc: Charlotte Smith (The University of Edinburgh)

Appendix J: Letter of Access for NHS Fife for Pilot and Feasibility Study

Research & Development

Research & Education Centre, Queen Margaret Hospital, Whitefield Road, Dunfermline, KY12 OSU

NHS Fife

Mrs Sarah Janes
171 Pleasance
EDINBURGH
EH8 9RU

7 February 2019
Our Ref 19-003 231852
17/WS/0189
Enquiries to Aileen Yell
Telephone 01383 623623 Ext 20940
E-mail aileenyell@nhs.net
Website www.nhsfife.org

Dear Mrs Janes

Letter of access for research
PROJECT TITLE : Cognitive contributors to risk of harm to others : a feasibility study

This letter confirms your right of access to conduct research through NHS Fife for the purpose and on the terms and conditions set out below. This right of access commences**7 February 2019** and ends on **31 August 2019** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. The information supplied about your role in research at NHS Fife has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to NHS Fife premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through NHS Fife, you will remain accountable to your employer University of Edinburgh, but you are required to follow the reasonable instructions of Mrs Moira Scott, Consultant Clinical & Forensic Psychologist, Lynebank Hospital, Dunfermline in this NHS organisation or those given on her behalf in relation to the terms of this right of access.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

You must act in accordance with NHS Fife policies and procedures, which are available to you upon request, and the UK Policy Framework for Health and Social Care Research.

You are required to co-operate with NHS Fife in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on NHS Fife premises. You must observe the same standards of care and propriety in dealing with patients, staff, visitors, equipment and premises as is expected of any other contract holder and you must act appropriately, responsibly and professionally at all times.

You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice (http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf) and the Data Protection Act 1998. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement. Please also ensure that while on the premises you wear your ID badge at all times, or are able to prove your identity if challenged. Please note that this NHS organisation accepts no responsibility for damage to or loss of personal property.

We may terminate your right to attend at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of this NHS organisation or if you are convicted of any criminal offence. As from 26 July 2010, your HEI employer may initiate your Independent Safeguarding Authority (IISA) registration (where applicable), and thereafter, will continue to monitor your ISA registration status via the on-line ISA service. Should you cease to be ISA-registered, this letter of access is immediately terminated. Your employer will immediately withdraw you from undertaking this or any other regulated activity. You MUST stop undertaking any regulated activity.

Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you. NHS Fife will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 2018. Any breach of the Data Protection Act 2018 may result in legal action against you and/or your substantive employer.

If your current role or involvement in research changes, or any of the information provided in your Research Passport changes, you must inform your employer through their normal procedures. You must also inform your nominated manager in this NHS organisation.

Yours sincerely

Dr Amanda Wood
Assistant R&D Director

cc: *Emily Gribbin, School Academic Administrator, University of Edinburgh (emily.gribbin@ed.ac.uk)*

Appendix K: Approval from CJSW R&D for Pilot and Feasibility Study

Yvonne Gannon <Yvonne.Gannon@edinburgh.gov.uk>

on behalf of

strategyandinsight <strategyandinsight@edinburgh.gov.uk>

Tue 7/31/2018 10:01 AM

To:

- JANES Sarah

Hello Sarah

I am pleased to confirm that your research request has been approved. Please contact Stuart Milnes (stuart.milnes@edinburgh.gov.uk) to discuss further.

Please don't hesitate to get in touch if you have any queries.

Kind regards

Yvonne

Yvonne Gannon | Senior Policy & Insight Officer (Insight Unit) | Strategy & Insight Division | Chief Executive | The City of Edinburgh Council | Business Centre 2.1 | Waverley Court | 4 East Market Street | Edinburgh | EH8 8BG | Tel: 0131 553 8334 | Email: yvonne.gannon@edinburgh.gcsx.gov.uk

Appendix L: Substantial Amendment to Pay Participants

WoSRES
West of Scotland Research Ethics Service

NHS
Greater Glasgow and Clyde

Mrs Sarah Janes
Clinical Psychology
University of Edinburgh
The University of Edinburgh
School of Health in Social Science
Medical School - Doorway 6
Teviot Place
Edinburgh
EH8 9AG

West of Scotland REC 3
Research Ethics
Clinical Research and Development
Dykebar Hospital
Grahamston Road
Paisley PA2 7DE

Date 11 July 2019
Direct line 0141 314 0211
E-mail WoSREC3@ggc.scot.nhs.uk

Dear Mrs Janes

Study title:	**Cognitive Contributors to Risk of Harm to Others: A Feasibility Study**
REC reference:	**17/WS/0189**
Amendment number:	**REC Ref AM06**
Amendment date:	**09 July 2019**
IRAS project ID:	**231852**

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion
The members of the Committee taking part in the review gave a **favourable ethical opinion** of the amendment on the basis described in the notice of amendment form and supporting documentation.

The sub-committee asked by email on 09 July 2019 that you amend the PIS to clarify that participants would be paid £10 per session up to a maximum of £20.

You responded by email on 11 July 2019 with an amended PIS.

The sub-committee were satisfied with this response.

Approved documents
The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Notice of Substantial Amendment (non-CTIMP) [Notice of Substantial Amendment]	REC Ref AM06	09 July 2019
Other [Payment Receipt]	V 1.0	02 July 2019
Participant information sheet (PIS) [PIS]	V 3.0	02 July 2019
Research protocol or project proposal [Protocol]	V 9.0	02 July 2019

Membership of the Committee
The members of the Committee who took part in the review are listed on the attached sheet.

Appendix M: RMO Letter

Cognitive Contributors to Risk of Harm to Others
01 August 2017 v1.0
IRAS Project ID: 231852

RMO Information Sheet

Date:

Responsible Medical Officer

Address:

Dear,

Re: Cognitive Contributors to the Risk of Harm to Others

During the next few months we would like to offer all new admissions and continuing care patients the opportunity to take part in the study titled, '**Cognitive Contributors to the Risk of Harm to Others'. We are contacting you to provide you with information and seek your consent to approach the patients for whom you are clinically responsible and your opinion on whether they have the capacity to consent to research.** The study is being conducted by Sarah Janes, Clinical Psychology PhD student and has received a favourable opinion by the State Hospitals Research Committee and West of Scotland REC 3 [17/WS/0189].

The study seeks to identify if cognitive impairments identified as predictive of violence through meta-analytic review are also predictive of inpatient violence in a forensic mental health setting and whether their routine assessment would be acceptable and feasible for this population. For this purpose we propose to undertake an assessment battery comprising seven assessments over two to three meetings each of which would take approximately one hour, although these will be adjusted according to participant's preference and capacity. To provide a positive contribution to patient care the results will be summarised for both clinical teams and the patient participants at the time of assessment and the measures themselves archived in the patient's medical records at the end of the study. The battery will examine the following cognitive abilities; intelligence, attention, impulsivity, empathy, response monitoring, risk taking behaviours, inhibitory control and social cognition. After assessments are completed, a 6, 12 and 24 month follow-up utilising Datix records and recorded incidents on patient risk will be collected and analysed. The participant will only be tested at baseline, and not at the follow-

If you feel it is appropriate for the patients under your care to be approached, a member of the psychology team independent of the research, will ask them if they would like to meet with me to hear about the study. The project will be explained to them by a member of the psychology team using the attached form and seven days will be given before their consent is sought, to allow them to consider whether they wish to participate. They will, of course, be free to withdraw from the project at any time.

A copy of the participant information sheet is enclosed for your information. Should you have any questions regarding this study, please do not hesitate to contact me by email s1461835@sms.ed.ac.uk, or 0747-400-7737.

Yours sincerely,

Sarah Janes + Dr Suzanne O'Rourke

Clinical Psychology PhD Student   Consultant Forensic Clinical Neuropsychologist

School *of* Health in Social Science
The University of Edinburgh
Medical School
Doorway 6, Teviot Place
Edinburgh EH8 9AG

Appendix N: Participant Information Sheet

Participant Information Sheet

Cognitive Contributors to the Risk of Harm to Others

You are being invited to take part in a research study. Before you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish. Please ask a member of your care team to contact us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

The purpose of this feasibility study is to explore if certain cognitive difficulties (in areas such as attention, impulsiveness, empathy, insight, thinking, intelligence and recognising the emotions of others) predict risk of harm to others. The study is being conducted in part fulfilment of a PhD in Clinical Psychology at the University of Edinburgh.

Why have I been asked to take part?

Previous published studies have suggested that difficulties in how we think about and process information, control our behaviour and understand other people's emotions may help explain why people can sometimes harm others. We think that patients in the State Hospital may have some of these difficulties and knowing this might help improve risk assessments and patient care.

Do I have to take part?

No, it is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. Deciding not to take part or withdrawing from the study will not affect the healthcare that you receive, or your legal rights.

What will happen if I take part?

To find out if you are having these kinds of difficulties, we'd like to ask you to do some assessments with us. These can involve many different things such as word meanings, computer puzzles, spotting the missing bit of pictures, and reading words. We will spread the assessments over different days so that sessions are not too long. If you decide to take part you can ask for a break or change your mind at any time. If you decide to take part we will meet two or three times for about an hour each time.

To find out what other assessments you have completed, and some information about your background, we would also like to look at your case notes. We also hope to access your personal records for up to 24 months after some or all of the assessments have been completed to get information on your follow-up care and interactions with others during this time. This information will be kept strictly confidential and no identifiable information will be used. We cannot do this without your permission and so we will ask you if you are willing to sign a form to say that this is okay. We will do this even if you choose not to complete all of the testing sessions. Again, it is entirely up to you whether you want to let us look at your files and saying no will not affect your care in any way. You can decide to change your mind at any time during the 24 months if you decide that you do not want us to see your records.

What are the possible benefits of taking part?

In order to help you gain some personal benefit from taking part in this study we will ask you if you would like your team to be given a summary of your results and whether we can place your assessments in your medical records after the feasibility study has ended. This will reduce the chance of you needing to do the same tasks again soon and the results may help your team understand your needs.

We will provide you with a written summary of your own results for you to keep.

What are the possible disadvantages and risks of taking part?
It is not thought that there are many disadvantages; however, it is possible that you may feel tired or find some of the tasks difficult. If that is the case remember you can stop the session ask Sarah Janes for help or withdraw from the study at any time. It is also possible that we may identify a problem that you are having with your thinking, learning, controlling your actions, or even your attention. We will give your clinical team all of your assessments. We will do so straight away if there is anything in the results that are clinically important but otherwise will do so once the study has finished.

What happens when the study is finished?
At the end of the research:

- A copy of the assessment information will be put into your confidential medical file after the feasibility study is over, i.e. after the follow up period of 6 months unless there is a clear clinical reason.
- We will anonymise your results and put them into a computer file with no names so that we can describe the kinds of difficulties patients in the hospital have. No one will be able to tell you have taken part in the study or how you did from this file.
- We will give your clinical team a one page summary of your results to help them understand your difficulties and to go over your results with you if you are interested.

Will my taking part in the study be kept confidential?
All the information we collect during the course of the research will be kept confidential and there are strict laws which safeguard your privacy at every stage. In certain circumstances if we feel you or another person is at risk of harm, confidentiality will be breached, or if you tell the researcher about any previously undisclosed criminal activity or other related information, the researcher will inform your responsible medical officer or a member of your direct health care team, however, we will talk to you first to let you know that this is going to happen.

Study researchers will need access to your medical records to carry out this research.

To ensure that the study is being run correctly, we will ask your consent for responsible representatives from the Sponsor University of Edinburgh to access your medical records and data collected during the study, where it is relevant to you taking part in this research. The Sponsor is responsible for overall management of the study and providing insurance and indemnity.

What will happen to the results of the study?
The study will be written up as conference presentations, academic papers, and a PhD thesis. You will not be identifiable in any published results.

Who is organising the research and why?
This study is being organised/sponsored by the University of Edinburgh. The study is funded by the State Hospital.

Who has reviewed the study?

The study proposal has been reviewed by the State Hospital. All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee (REC). The State Hospital and the Research Ethics Committee have given their approval for this study to go ahead.

If you have any further questions about the study please contact ask a member of your clinical team to contact Sarah Janes on *s1461835@sms.ed.ac.uk* **or, if your matter is urgent, Suzanne O'Rourke on Suzanne.o'rourke@nhs.net.**

If you would like to discuss this study with someone independent of the study please contact:

Dr. Angus MacBeth

The University of Edinburgh
Tel: 0131 650 3893

Email: angus.macbeth@ed.ac.uk

If you wish to make a complaint about the study please contact:

The University of Edinburgh's Research Governance Team by email at: resgov@accord.scot

Thank you for taking the time to read this information sheet.

Yours Sincerely,

Sarah Janes + Dr Suzanne O'Rourke
Clinical Psychology PhD Student Consultant Forensic Clinical Neuropsychologist

Appendix O: Participant Consent Form

CONSENT FORM

Cognitive Contributors to the Risk of Harm to Others

Participant ID:

Sarah Janes

SCHOOL of HEALTH IN SOCIAL SCIENCE
The University of Edinburgh
Medical School
Doorway 6, Teviot Place
Edinburgh EH8 9AG

Please initial box

1. I confirm that I have read and understand the information sheet (Version 4.0, 20 November 2017) for the above study and have had the opportunity to consider the information and ask questions. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected. ☐

3. I give consent for any completed assessments to be used for the purposes of this study even if I choose to withdraw from the study or am no longer able to take part before all assessments are completed. ☐

4. I understand that relevant sections of my medical notes and assessment data collected during the study may be looked at by individuals from the University of Edinburgh where it is relevant to my taking part in this research. I give permission for these individuals to have access to my records. ☐

5. I give consent for my completed assessments to be put into my medical record and for the clinical team to know the results after the feasibility study has ended, i.e. after 6 months, unless there is a clear clinical reason for them to be shared earlier. ☐

6. I give consent for members of the research team from the current study to have access to my medical records in paper, electronic or database form for up to 24 months after some or all assessments have been completed. ☐

6. I agree to take part in the above study. ☐

Name of Participant	Date	Signature
Name of Person taking consent	Date	Signature

NOTE: Original (x1) to be retained in site file; Copy (x1) to be included in patient notes; Copy (x1) to be retained by the participant.

Appendix P: Case Note Review Checklist

Cognitive Contributors to Harm toward Others

Case Note Checklist

Demographics

1. **Date of current admission:**
2. **Source of current admission:**
3. **Legislation for current detention:**
4. **Current conviction status:**
5. **Diagnosis:**
   - **Primary:**
   - **Secondary:**
6. **Year of first diagnosis with a psychotic illness:**

Data related to Risk

1. **Reason(s) for current admission:**
2. **Offence(s) or alleged offence(s) leading to current admission:**
3. **Violent incidences during current admission (if applicable):**
4. **Total number of previous convictions:**
5. **The most serious previous offence:**
6. **Previous types of offences committed:**

Additional descriptive data of relevance given their potential contribution to the aetiology of cognitive impairments

1. **Birth trauma?:**
2. **Abnormal infant development?:**

3. **Childhood history of physical abuse?:**

4. **Childhood history of sexual abuse?:**

1. **Childhood history of neglect?:**

2. **Diagnosis of alcohol dependence?:**

3. **Diagnosis of drug dependence?:**

4. **Neurological injuries or neurologically relevant diagnoses:**

Routinely Administered Assessments

1. **HCR-20 (Version completed closest to time of testing):**

H: ____________________

C: ____________________

R: ____________________

Subscales:

Please tick scores for the following factors below:	0 No Evidence/ Low Probability	1 Possible or Sufficient Evidence/ Moderate Probability	2 Definite or Clear Evidence/ High Probability
H1. Previous Violence	☐	☐	☐
H2. Other Antisocial behaviour	☐	☐	☐
H3. Relationships	☐	☐	☐
H4. Employment problems	☐	☐	☐
H5. Substance use problems	☐	☐	☐
H6. Major mental illness	☐	☐	☐
H7. Personality disorder	☐	☐	☐

H8. Traumatic experiences	☐	☐	☐
H9. Violent attitudes	☐	☐	☐
H10. Treatment/SV response	☐	☐	☐
C1. Insight	☐	☐	☐
C2. Violent ideation or intent	☐	☐	☐
C3. Symptoms of Major Mental Illness	☐	☐	☐
C4. Instability	☐	☐	☐
C5. Treatment/SV Response	☐	☐	☐
R1. Prof. services and plans	☐	☐	☐
R2. Living situation	☐	☐	☐
R3. Personal support	☐	☐	☐
R4. Treatment/SV response	☐	☐	☐
R5. Stress or Coping	☐	☐	☐

- **Future Violence/Case Prioritization (Circle one):** High/Medium/Low
- **Serious Physical Harm (Circle one):** High/Medium/Low
- **Imminent Violence (Circle one):** High/Medium/Low

2. **PANSS Scores (if not collected by researcher):**
    - **Positive Scale:** ___________
    - **Negative Scale:** __________
    - **P-N Total:** _____________
    - **General:** ________________

		Positive Scale
Item #	**Domain**	
P1	**Delusions**	
P2	**Conceptual Disorganization**	
P3	**Hallucinatory Behaviours**	
P4	**Excitement**	
P5	**Grandiosity**	
P6	**Suspiciousness**	
P7	**Hostility**	
		Negative Scale
N1	**Blunted Affect**	
N2	**Emotional Withdrawal**	
N3	**Poor Rapport**	
N4	**Passive/Apathetic Social Withdrawal**	
N5	**Difficulty in Abstract Thinking**	
N6	**Lack of Spontaneity/Flow of Convo.**	
N7	**Stereotyped Thinking**	
		General Scale
G1	**Somatic Concern**	

G2	Anxiety	
G3	Guilt Feelings	
G4	Tension	
G5	Mannerisms/Posturing	
G6	Depression	
G7	Motor Retardation	
G8	Uncooperativeness	
G9	Unusual Thought Content	
G10	Disorientation	
G11	Poor Attention	
G12	Lack of Judgement and Insight	
G13	Disturbance of Volition	
G14	Poor Impulse Control	
G15	Preoccupation	
G16	Active Social Avoidance	

Appendix Q: Tolerability Measure

How did you find that measure?

7	6	5	4	3	2	1
Extremely Pleasant	Very Pleasant	Pleasant	Neutral	Unpleasant	Very Unpleasant	Extremely Unpleasant

Comments:

Participant ID:
Measure:
Date:

Appendix R: Correlation Matrix for Inpatient Sample- Prospective Outcomes

Correlation Matrix for Inpatients and Prospective Outcomes (n=32)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Head Injuries^																	
2. Head Injuries LOC^	.73**																
3. Severity TBI^	.82**	.71**															
4. HCR-20	.18	.20	.25														
5. Lack of Insight	.18	.07	-.02	.38*													
6. Reasoning	-.02	-.11	-.06	-.62**	-.49**												
7. Inattention	-.06	-.08	-.01	.25	.25	-.69**											
8. Response Inhibition	-.05	-.03	-.09	-.16	-.13	.23	-.25										
9. Response Monitor	-.20	-.19	-.07	-.25	-.27	.52**	-.37*	.17									
10. Risk Taking	-.22	-.00	-.13	-.30	-.19	.29	-.17	-.04	-.23								
11. Affect Recognition	.10	-.16	.13	-.20	-.34	.40*	-.48**	.12	.25	.07							
12. Cognitive Empathy	.02	-.15	-.14	-.39*	-.33	.53**	-.40*	.21	.32	.21	.56**						
13. Violence Severity	-.09	-.15	-.08	.24	.31	-.25	-.01	-.13	-.20	-.00	.22	.09					
14. Aggression	.08	-.11	.07	.21	.11	-.23	.15	-.13	-.31	-.01	.28	.12	.82**				
15. Level of Obs.	.05	-.04	-.09	.03	.07	.06	-.12	-.15	-.25	-.07	.00	.11	.38*	.27			
16. Days on Levels	.05	-.04	-.09	.01	.04	.08	-.10	-.14	-.25	-.06	.01	.13	.37**	.28	1.00**		
17.Violence Frequency	.02	.03	.23	.14	-.14	-.21	.35	-.20	-.24	.08	.18	-.13	.45**	.69**	.02	.02	
18. Length of Admiss.	-.14	.12	-.02	.02	-.08	-.07	-.16	-.01	-.09	.15	-.04	-.25	-.19	-.13	-.01	.00	-.16

Note. Items 13-17 are prospective outcomes; LOC= Loss of consciousness; TBI= traumatic brain injury; HCR-20= Historical, clinical, Risk 20; Obs.=Observations, Admiss= Admission, ^*n*=31

* indicates $p < .05$.

** indicates $p < .01$.

Appendix S: Correlation Matrix for Inpatient Sample- Retrospective Outcomes

Correlation Matrix for Inpatients and Retrospective Outcomes (n=32)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Head Injuries^																
2. Head Injuries LOC^	.73**															
3. Severity TBI^	.82**	.71**														
4. HCR-20	.18	.20	.25													
5. Lack of Insight	.18	.07	-.02	.38*												
6. Reasoning	-.02	-.11	-.06	-.62**	-.49**											
7. Inattention	-.06	-.08	-.01	.25	.25	-.69**										
8. Response Inhibition	-.05	-.03	-.09	-.16	-.13	.23	-.25									
9. Response Monitor	-.20	-.19	-.07	-.25	-.27	.52**	-.37*	.17								
10. Risk Taking	-.22	-.00	-.13	-.30	-.19	.29	-.17	-.04	-.23							
11. Affect Recognition	.10	-.16	.13	-.20	-.34	.40*	-.48**	.12	.25	.07						
12. Cognitive Empathy	.02	-.15	-.14	-.39*	-.33	.53**	-.40*	.21	.32	.21	.56**					
13. Violence Severity	.28	.32	.14	.58**	.44*	-.37*	.14	-.22	.38*	-.14	-.14	-.14				
14. Aggression	-.06	-.15	-.08	.29	.12	-.17	.24	-.11	-.29	-.10	-.17	.04	.56**			
15. Level of Obs.	.20	.29	.13	.49**	.21	-.26	.16	-.25	-.26	-.24	-.09	-.21	.65**	.57**		
16. Days on Levels	.19	.30	.10	.53**	.33	-.29	.18	-.32	-.25	-.21	-.14	-.26	.70**	.48**	.97**	
17.Violence Frequency	.09	.29	.07	.28	.05	-.23	.13	-.02	-.22	-.11	-.16	-.11	.54**	.66**	.72**	.58**

Note. Items 13-17 are retrospective outcomes; LOC= Loss of consciousness; TBI= traumatic brain injury; HCR-20= Historical, clinical, Risk 20; Obs.=Observations; ^*n*=31

* indicates $p < .05$.

** indicates $p < .01$.

Appendix T: Correlation Matrix for Community Sample- Prospective Outcomes

Correlation Matrix for Community Sample- Prospective Outcomes (n=31)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Head Injuries														
2. Head Injuries LOC	.30													
3. Severity TBI	.57	.77												
4. Reasoning	.33	-.07	-.04											
5. Inattention	.09	.06	.02	-.27										
6. Response Inhibition	.05	-.06	-.03	.25	-.14									
7. Response Monitor	.19	-.18	-.04	-.08	-.02	.02								
8. Risk Taking	.04	.35	.32	-.17	.24	-.16	.09							
9. Affect Recognition	.03	-.03	-.08	.10	-.14	.40*	.06	-.22						
10. Cognitive Empathy	.36*	.03	.07	.36*	-.44*	.07	-.00	-.01	.41*					
11. Violence Severity	-.03	.09	.08	.13	-.19	-.08	-.19	.15	-.25	.02				
12. Aggression	.08	-.11	.07	.21	.11	-.23	-.05	.04	.05	.25	.36			
13. Violence Frequency	.05	-.04	-.09	.03	.07	.06	-.04	.12	-.13	.06	.88**	.46*		
14. Antisocial	.05	-.04	-.09	.01	.04	.08	.27	-.23	-.10	.10	-.11	-.11	-.11	
15. Breaches	.02	.03	.23	.14	-.14	-.21	-.21	.14	-.30	-.09	.91**	-.05	.69**	-.08

Note. Items 11-15 are prospective outcomes.
* indicates $p < .05$.
** indicates $p < .01$.

Appendix U: Correlation Matrix for Community Sample- Retrospective Outcomes

Correlation Matrix for Community Sample- Retrospective Outcomes (n=31)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Head Injuries														
2. Head Injuries LOC	.30													
3. Severity TBI	.57	.77												
4. Reasoning	.33	-.07	-.04											
5. Inattention	.09	.06	.02	-.27										
6. Response Inhibition	.05	-.06	-.03	.25	-.14									
7. Response Monitor	.19	-.18	-.04	-.08	-.02	.02								
8. Risk Taking	.04	.35	.32	-.17	.24	-.16	.09							
9. Affect Recognition	.03	-.03	-.08	.10	-.14	.40*	.06	-.22						
10. Cognitive Empathy	.36*	.03	.07	.36*	-.44*	.07	-.00	-.01	.41*					
11. Violence Severity	-.09	-.15	-.08	.24	.31	-.25	-.12	-.18	.10	.23				
12. Aggression	.08	-.11	.07	.21	.11	-.23	.23	.28	-.17	-.02	.02			
13. Violence Frequency	.05	-.04	-.09	.03	.07	.06	-.22	-.25	.11	.28	.96**	-.22		
14. Antisocial	.05	-.04	-.09	.01	.04	.08	.38*	.30	-.20	-.10	-.25	.45*	-.35	
15.Breaches	.02	.03	.23	.14	-.14	-.21	-.00	-.23	-.05	-.15	-.29	-.13	-.25	.27

Note. Items 11-15 are retrospective outcomes.
* indicates $p < .05$.
** indicates $p < .01$.

Appendix V: Correlation Matrix for Entire Sample and Self-Reported Impulsivity

Correlation Matrix for Entire Sample (n=63)

Variable	1	2	3	4	5	6	7
1. Reasoning							
2. Inattention	-.48**						
3. Response Inhibition	.26*	-.20					
4. Response Monitoring	.25*	-.20	.17				
5. Risk Taking	.10	.01	-.10	.14			
6. Affect Recognition	.30*	-.34**	.25	.22	-.02		
7. Cognitive Empathy	.46**	-.38**	.25	.30*	.10	.54**	
8. Self-Reported Impulsivity	.01	.11	.09	-.08	-.13	.21	.28*

Note. Correlation matrix for core variables and self-reported impulsivity for entire sample
* indicates $p < .05$.
** indicates $p < .01$.

Appendix W: Residuals Graphs for Selected Regressions

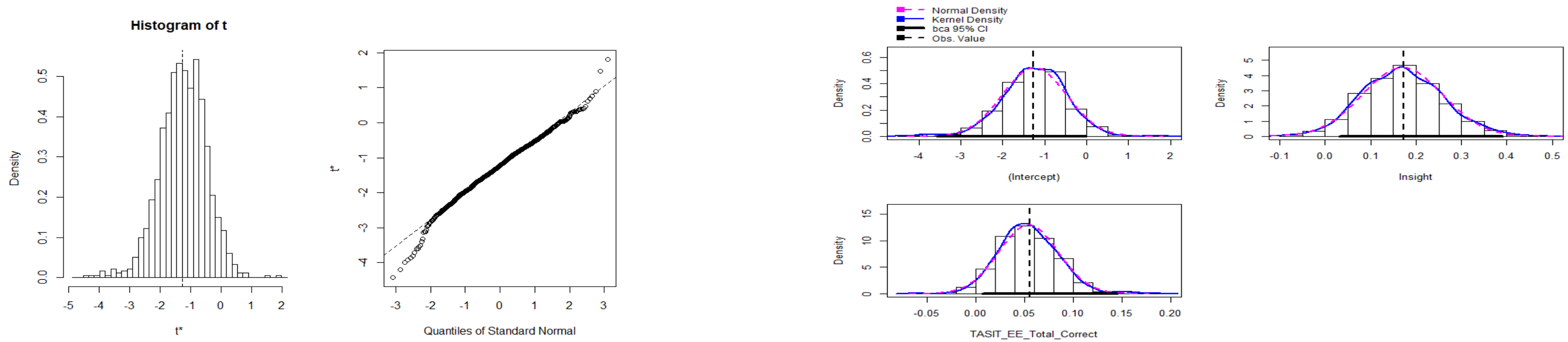


Figure 1. Bootstrapped residuals for affect recognition, and lack of insight with prospective violence severity as the criterion

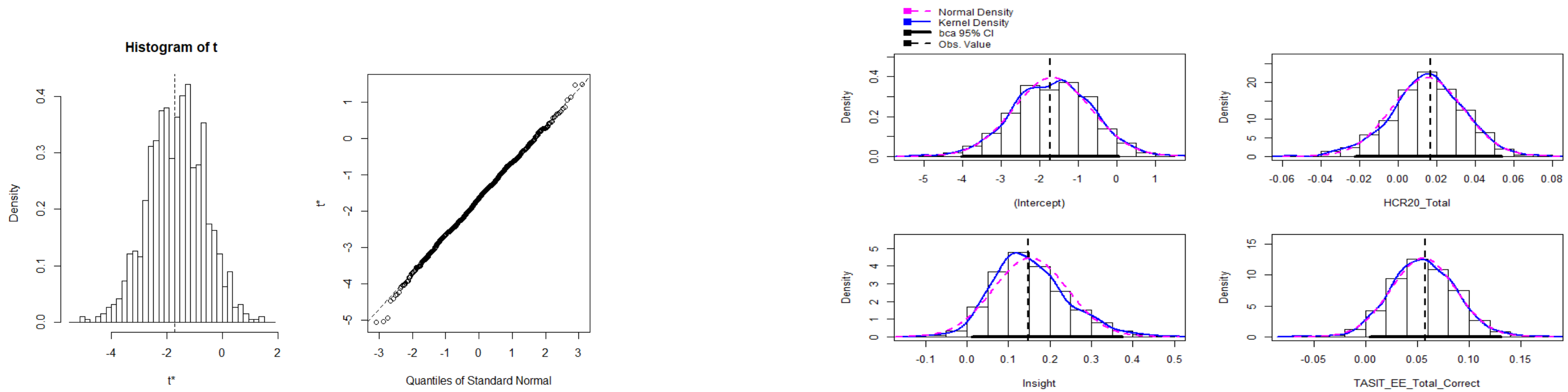


Figure 2. Bootstrapped residuals for HCR-20, affect recognition, and lack of insight with prospective violence severity as the criterion