COGNITIVE ASSESSMENT OF DEVIANT SEXUAL INTEREST USING THE EMOTIONAL STROOP AND FMRI

By

SHELLEY PRICE

A thesis submitted to the College of Life and Environmental Sciences of the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

School of Psychology
College of Life and Environmental Sciences
The University of Birmingham
December 2010

Abstract

This thesis examines differences in information-processing of sexual material using the emotional Stroop task to determine whether it is a reliable tool to be used in the assessment of deviant sexual interest for sexual abusers. Data was collected from several samples: adolescent sexual abusers receiving treatment from g-map services (Sale, Cheshire); adolescent offenders involved with a Youth Offending Team organisation in the West Midlands; adolescent non-offenders recruited from the community; adult sexual abusers and violent offenders who attend treatment with West Midlands Probation or who are detained at HMP Grendon (Oxfordshire, UK); and adult non-offenders in the West Midlands. Chapter one reviews the Stroop task and provides guidance to researchers for developing Stroop studies. Chapter two provides an overview of the methodologies used throughout the thesis and reports on the development of new word stimuli to measure sexual interest specific to sexual abusers. In Chapter three, adult sexual abusers demonstrated Stroop bias towards general sexual word content and towards words reflecting sexual interest more specific to sexual abusers. Chapter four found that the variant of the emotional Stroop task used in this thesis was not a suitable tool to measure sexual interest in adolescent samples. Chapter five compared the datasets of adult and adolescent samples and found that offender groups exhibit a general slowing effect overall on the emotional Stroop, and that differences in Stroop bias were evident between the different age cohorts. Chapters six/seven identified specific areas of the brains of sexual abusers that are disinhibited when processing emotional and sexual word content during the completion of an emotional Stroop task. The thesis concludes with a discussion of the overall findings, the limitations of the research using this paradigm, the implications of the findings, and suggestions for future research.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis supervisor Tony Beech for his wonderful guidance, knowledge, support that he so generously offered, and the freedom that he entrusted in me to conduct my work. When I began this venture I could only have imagined the amount of kindness and opportunity that were presented throughout the duration of the PhD, and for that I am grateful.

Thank you to my advisor Ian Mitchell for his support, lovely chats and enthusiasm for the research project, and thanks to my advisor Glyn Humphreys for providing me with the opportunity to conduct the fMRI research. In addition, special thanks are due to Harriet Allen for her brilliance and guidance provided so that I could complete the fMRI experiment and analyse the complex data.

Special thanks are also due to the staff at g-map services in Sale, particularly Bobbie Print, for sponsoring my research, being so helpful and for providing me with access to the adolescent participants. Thank you to West Midlands and Cornwall Probation, the Youth Offending Team in Dudley, and HMP Grendon for granting me access to their clients and for the planning efforts made by staff to make this research possible. Also, great thanks are due to the National Organisation for the Treatment of Abusers for the generous grant and to the Association for the Treatment of Sexual Abusers for presenting me with the Graduate Research Award for the research within this thesis.

An endless amount of gratitude goes to my colleagues and friends at the Centre for Forensic and Criminological Psychology, Shannon Vettor, Leigh Harkins, Sue Hanson, Alasdair Goodwill, Louise Dixon, Stella Briggs and Jessica Woodhams for welcoming me with such kindness from the first day I arrived at the University and for the laughter, encouragement and generosity to help. In addition, enormous

thanks are due to the wonderful people that I have met throughout the PhD, colleagues and to my closest friends who have made my experiences over the past three years so brilliant, you have made Birmingham feel like home to me: Kirsty Mack, Katerina Kantartzis, and Robert Cryer.

I would also like to thank my colleagues and friends in Ottawa for helping to make this possible, having confidence in me and for always being there to come home to: Karl Hanson, Guy Bourgon, Kevin Nunes, and Holly Manuel.

To Jackie Kachuik. Well, what can I say? You have been my rock! For the countless and endless phone calls, wonderful visits, and your ability to be so far away but seem right by my side throughout this process I am eternally indebted to you! You are my confidant, my soul mate and a sister to me.

Finally, I would like to thank my family for being so supportive no matter how far across the ocean I found myself. Mom, Pops, James, Howie, Marilyn, Emily and Scott, your patience and understanding throughout the years has been so important to me and I love you all very much. I miss you all the time, but as soon as I am home I feel as though I never left. To my parents, who without their confidence in me and belief in what I could do, this would not have been possible. There's nothing quite like home. And now you can finally say that I am finished school!

P.S. Life is what happens while you're busy making other plans.

TABLE OF CONTENTS

Introduction **1-6**

Statement of Authorship **7**

Chapter 1: The Promises and Perils of the Emotional Stroop Task: A General Review and Considerations for Use with Forensic Samples

Chapter Rationale **8**

Price, S. A., Beech, A. R., Mitchell, I., & Humphreys, G. W. (in press). The Promises and Perils of the Emotional Stroop Task: A General Review and Considerations for Use with Forensic Samples, *Journal of Sexual Aggression* **10-35**

Chapter 2: General Methodology and the Development of an Emotional Stroop Word Stimulus Set

Chapter Rationale **36**

Chapter 2: General Methodology and the Development of an Emotional Stroop Word Stimulus Set **37-48**

Chapter 3: Measuring Deviant Sexual Interest in Adult Samples Using the Emotional Stroop Task

Chapter Rationale **49**

Chapter 3: Measuring Deviant Sexual Interest in Adult Samples Using the Emotional Stroop Task **50-75**

Chapter 4: Measuring Deviant Sexual Interest in Adolescent Samples Using the Emotional Stroop Task

Chapter Rationale **76**

Chapter 4: Measuring Deviant Sexual Interest in Adolescent Samples Using the Emotional Stroop Task **77-100**

Chapter 5: Comparison of Measurements of Deviant Sexual Interest Between Adult and Adolescent Samples Using the Emotional Stroop Task

Chapter Rationale **101**

Chapter 5: Comparison of Measurements of Deviant Sexual Interest Between Adult and Adolescent Samples Using the Emotional Stroop Task **102-126**

Chapter 6: An Emotional Stroop Task Between Offenders and Controls and the Identification of Areas of Brain Activation Through the Use of Functional Magnetic Resonance Imaging (fMRI).

Chapter Rationale **127-128**

Price, S. A., Beech, A. R., Allen, H. A., & Mitchell, I. (submitted). An Emotional Stroop Task Between Offenders and Controls and the Identification of Areas of Brain Activation Through the Use of Functional Magnetic Resonance Imaging (fMRI). *Archives of Sexual Behavior.* **129-177**

Chapter 7: Further Development of fMRI Baseline Activation Patterns from an Emotional Stroop Task Paradigm for Use in work with Sexual Abusers

Chapter Rationale **178**

Further Development of fMRI Baseline Activation Patterns from an Emotional Stroop Task Paradigm for Use in work with Sexual Abusers **179-204**

Chapter 8: Discussion **205-214**

Appendices **215-220**

References **221-248**

LIST OF FIGURES

Figure	Title	Page
Chapter 3		
1	Emotional Stroop Bias Scores: Smith and Waterman Word Stimulus Set	**62**
2	Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman Stimuli	**64**
3	Emotional Stroop Bias Scores: Newly-Derived Word Stimulus Set	**67**
4	Means for the Main Effects of Type of Stroop Bias and Participant Group: Newly-Derived Word Stimuli	**68**
Chapter 4		
1	Mean Emotional Stroop Bias Scores: Smith and Waterman Stimulus Set	**90**
2	Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman Stimuli	**92**
3	Emotional Stroop Bias Scores for Newly-Derived Word Stimuli	**95**
4	Means for the Main Effects of Type of Stroop Bias and Participant Group: Newly-Derived Word Stimuli	**96**
Chapter 5		
1	Mean Age, BACS (CD) and BACS (EC): Adolescent and Adult Groups	**106**
2	Tests of Executive Function: Adolescent and Adult Groups	**107**

3	Negative Stroop Bias Scores by Participant Group	**111**
4	Sexual Stroop Bias Scores by Participant Group	**112**
5	Aggression Stroop Bias Scores by Participant Group	**113**
6	Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman Stimuli	**115**
7	EPD Stroop Bias Scores by Participant Group	**118**
8	SA Stroop Bias Scores by Participant Group	**119**
9	PD Stroop Bias Scores by Participant Group	**120**
10	Means for the Main Effects of Type of Stroop Bias and Participant Group: Newly-Derived Word Stimuli	**121**
Chapter 6		
1	Activation Areas for Contrasts of Interest for Control Subjects	**174**
2	Significant Clusters for Contrast Sexual - Neutral Words: Controls > Sex offenders	**175**
3	Significant Clusters for Contrasts of Interest: Controls > Sex offenders	**177**
Chapter 7		
1A	Uncorrected Areas of Activation for Non-Offending Controls: Incongruent – Congruent Contrast of Interest	**191**
1B	Significant Clusters for Contrast Incongruent – Congruent: Non-Offending Controls (threshold Z > 1.8, p=.05)	**192**
2A	Overlapping Areas of Activation for Contrasts of Interest in the Smith and Waterman Word Stimulus Set: Non-Offending Controls	**193**

2B	Significant Areas of Activation for Sexual – Neutral Contrast of Interest: Non-Offending Controls	**193**
2C	Uncorrected Areas of Activation for Non-Offending Controls: Emotion > Non-Emotion, Smith and Waterman Word Categories	**194**
2D	Significant Clusters for Contrast Emotional – Non-Emotional: Non-Offending Controls (threshold Z > 2.3, p=.05)	**195**
3A	Uncorrected Areas of Activation for Non-Offending Controls: Emotion > Non-Emotion word activation, Newly-Derived Word Stimuli	**196**
3B	Areas of activation for Contrast Emotion > Non-Emotion: Smith and Waterman Stimulus Set (blue) and Newly-Derived Word Stimuli (red)	**197**
4A	Significant Clusters for Contrast Incongruent – Congruent: Non-Offending Controls > Exhibitionist (threshold Z > 1.8, p=.05)	**199**
4B	Significant Clusters for Contrast SA – MSA (blue) and PD – MPD (red): Non-Offending Controls > Exhibitionist	**200**

LIST OF TABLES

Table	Title	Page
Chapter 1		
1	Key References: The Emotional Stroop Task	**17**
Chapter 3		
1	Violations to the Assumption of Normally Distributed Data for the Newly-Derived Word Stimulus Set	**57**
2	Mean Age, Scores on Study Measures and Number of Stroop Errors (SD) by Participant Group	**58**
3	Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Smith and Waterman (2004) Word Stimulus Set	**61**
4	Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Newly-Derived Word Stimulus Set	**66**
Chapter 4		
1	Violations to the Assumption of Normally Distributed Data for the Smith and Waterman Stimulus Set	**85**
2	Violations to the Assumption of Normally Distributed Data for the Smith and Waterman Stimulus Set	**86**
3	Mean Age, Scores on Measures and Number of Stroop Errors (SD) per Adolescent Group	**87**
4	Mean RT and emotional Stroop Bias Scores in Milliseconds (SD) for Smith and Waterman (2004) Word Stimulus Set	**89**
5	Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Newly-Derived Word Stimulus Set	**93**

Chapter 5		
1	Significant Differences from Non-Parametric Post-Hoc Comparisons of Mean RT's	**110**
2	Significant Differences from Non-Parametric Post-Hoc Comparisons of Mean RT's: Newly-Derived Word Stimuli	**117**
Chapter 6		
1	Scale Responses and Stroop Errors per Group with (Standard Deviations)	**167**
2	Mean Response Times (RT) and Bias Scores with (Standard Deviations) for the Emotional Stroop Task	**168**
3	Average Age, Scale Responses and Stroop Errors Per Group with (Standard Deviations) for the fMRI Emotional Stroop Task	**169**
4	Mean Response Times (RT) and Bias Scores with (Standard Deviations) for the fMRI Emotional Stroop Task	**170**
5	Areas of Activation for Control Subjects per Word Category	**171**
6	Areas of Significant Clusters of Experimental Target Word Contrasts: Controls	**172**
7	Areas of Significant Clusters of Experimental Target Word Contrasts for Controls > Sexual Offenders	**173**
Chapter 7		
1	Mean RT's and emotional Stroop bias scores for controls and case study	**190**

2	Areas of Significant Activation for the Emotional > Non-Emotional Contrast of Interest between the Newly-Derived Word Stimuli and Smith and Waterman Stimuli.	**198**

INTRODUCTION

Sexual crimes against children have a profound and detrimental effect on individuals and communities, and these types of crimes are a reality in the lives of many. In 2008/2009, there were 40,787 serious sexual offences (rape, sexual assaults and sexual activity with children) recorded by police in the United Kingdom (Home Office, July 2009). Research has shown us that deviant sexual interest is the best known predictor of sexual recidivism as measured by penile plethysmography (Hanson & Bussière, 1998). Not only is this factor considered important to incorporate into actuarial risk measures used for the assessment of sexual offenders (Craig, Browne & Beech, 2008; Hanson, Harris, Scott & Helmus, 2007; Hanson & Thornton, 2003; Prentky & Righthand, 2003; Wong, Olver, Nicholaichuk & Gordon, 2004), but it has also been identified as a factor that is important to include in sexual abuser treatment programmes as a target in treatment (Beech, Oliver, Fisher & Beckett, 2006; Marshall, Anderson & Fernandez, 1999; Yates, Goguen, Nicholaichuk, Williams & Long, 2000). Therefore, much of the research presently being conducted is focussed on being able to more accurately measure deviant sexual interests for the purposes of reducing recidivism, and in hopes of targeting treatment more accurately and specific to this factor (Thornton & Laws, 2009).

The nature of an individual's sexual interests is used to distinguish between repetitive sexual abusers, non-offenders and lower risk offenders. This is thought to be a determinant of an offender's degree of risk, or persistence in sexual offending (Thornton & Laws, 2009). However, the measurement of deviant sexual interests is a difficult task because these interests are implicit in nature. Actuarial risk measures have incorporated specific items that are used to measure deviant sexual interests. For example, the Static-2002 (Hanson & Thornton, 2003) considers three items to score

deviant sexual interest: (1) any sentencing occasions for non-contact sex offences; (2) any male victims; and (3) having two or more victims less than age 12, one unrelated. The Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II; Prentky & Righthand, 2003) also incorporates these items to assess adolescent samples of sexual abusers. The items, however, are found under the heading of 'sexual drive/preoccupation'. Importantly, these items have been found to be associated with level of risk and persistence in sexually abusive behaviours (Beckett, Beech, Fisher & Fordham, 1994; Beech, 1998; Hanson & Bussière, 1998; Långström, 2002; Rasmussen, 1999; Worling, 2002). However, a difficulty in using these variables to measure a concept such as deviant sexual interest is that they are static in nature and do not provide us with a clear indication of what the sexual abuser 'thinks' or 'believes' is attractive, sexually stimulating or sexually arousing. In effect, these variables would seem to merely touch the surface of what it is we are interested in measuring with regards to deviant sexual interests, failing to tap into the cognitive and affective components of the construct.

Generally, sexual interests are defined as "ideas, beliefs or preferences concerning what individuals find attractive, sexually stimulating or significant for sexual arousal to occur…sexual interests concern thoughts, ideas, beliefs, feelings or perceptions " (Price & Hanson, 2007, pp. 203-204). Sexual interests are considered 'deviant' when willing consent is lacking (Price & Hanson, 2007). It would seem then that we require a method of accessing sexual interest information at a more profound and implicit level (i.e., at an information-processing level), which would in turn be of critical value in the identification of treatment needs and the assessment of risk of sexual abusers (Thornton & Laws, 2009).

One measure that has been used in attempts to tap into the psychological nature of deviant sexual interests with offenders is the emotional Stroop task (Price & Hanson, 2007; Smith & Waterman, 2004). This task uses differences in reaction times (RT) to emotional words and neutral or control words as a measure of interference in information-processing (Larsen, Mercer & Balota, 2006; Wentura, Rothermund, & Bak, 2000). The emotional Stroop task is described as an information-processing measure because the response latencies experienced by individuals during the presentation of emotional stimuli are considered to be a function of the processing of the stimuli (Smith, 2009). Theory surrounding the emotional Stroop task maintains that the interference in information-processing caused by the emotional content reflects an individual's implicit attitudes, emotions, and motivations (Dalgleish, 2005; Klein, 1964), and consists of both cognitive and affective components (Smith, 2009).

Therefore, it would seem practical to apply the emotional Stroop task to measure an implicit construct such as deviant sexual interest. Research using the emotional Stroop task with offender samples has attempted to create sets of word stimuli that could be characterized as deviant in the sexual interest domain, and more specifically that could be able to differentiate between subgroups of sexual abusers (Price & Hanson, 2007). Efforts to do so are made in hopes of applying this tool as an alternative to traditional assessment methods for use with offender samples allowing researchers to evaluate the implicit attitudes of offenders without having to rely on their word.

In addition to the body of research examining emotional Stroop interference effects, developments in cognitive neuroscience have employed functional magnetic resonance imaging (fMRI) to identify brain structures that are responsible for the processing of emotionally salient stimuli (Ersche et al., 2010; Malhi, et al., 2005;

Shin, et al., 2001; van den Heuvel, et al., 2005). By attempting to piece together the cognitive and affective mechanisms involved in the processing of sexual and aggressive information during an emotional Stroop task, we could become better prepared to answer the question of whether we can use an emotional Stroop paradigm to measure sexual interest adequately.

Therefore, the overall aim of the thesis is to examine differences in information-processing of sexual material between groups of offenders and non-offending controls, and to determine whether the emotional Stroop task is a reliable tool to be used in the assessment of deviant sexual interest.

Specifically, the thesis will:

- Introduce the emotional Stroop task as a possible alternative to traditional methods of assessment;
- Outline the development of an empirically derived set of emotional Stroop word stimuli that are intended to reflect a facet of deviant sexual interests;
- Explore the utility of the emotional Stroop task with several offender groups varying in age and offence type;
- Examine the brain structures that are involved in the processing of deviant sexual information during the completion of emotional Stroop tasks through the use of functional magnetic resonance imaging (fMRI).

Structure of the Thesis

This thesis examines the measurement of deviant sexual interest using the emotional Stroop task and fMRI. Chapter 1 presents a review of emotional Stroop literature in general, as well as the work that has been completed using offender samples. Chapter 1 also outlines factors that researchers should take into account when developing a Stroop experiment that are known to have an effect on Stroop results, and provides suggestions of how to minimize potential confounding variables to elicit purer results.

Chapter 2 presents the general methodology used for the emotional Stroop experiments that comprise the thesis in Chapters 3, 4, and 5. Chapter 2 also describes the empirical development of a new set of word stimuli intended to reflect a facet of deviant sexual interest.

Chapter 3 reports on the differences in emotional Stroop response times (RT) and Stroop bias scores between three groups of adults: adult sexual abusers, adult violent offenders, and adult non-offending controls. Participant characteristics and experimental issues are considered in this chapter along with the utility of the emotional Stroop with adult samples.

Chapter 4 explores the utility of the emotional Stroop task with adolescents and potential participant characteristics that may influence Stroop results with this sample. Factors such as vocabulary, level of executive functioning, and evidence of learning disabilities are examined and discussed.

Chapter 5 presents a comparison between adolescent and adult participants in emotional Stroop interference effects, and highlights significant differences that are evident between the age groups on the measures of executive function.

Chapter 6 establishes baseline brain activation patterns in non-offending control subjects to offence-related word stimuli during the presentation of an emotional Stroop task using fMRI. Activation patterns from the control subjects are then compared to a small group of adult sexual abusers.

Chapter 7 presents a case study of an exhibitionist and compares the brain activation patterns of the case study to baseline activations established in Chapter 6. In particular, Chapters 6 and 7 attempt to create an understanding of the cognitive processes involved in the processing of offence-related information and how we can use this understanding to guide work with sexual abusers.

Chapter 8 concludes the thesis by providing a general discussion of the results and considers the implications of the findings and directions for future research using the Stroop task and fMRI.

Ethical Approval

The British Psychological Society code of ethical practice was adhered to in the design of the research projects. Ethical approval was gained from The School of Psychology Ethics Committee at the University of Birmingham, West Midlands Probation, Grendon Research Advisory Group (GRAG), the Birmingham University Imaging Centre (BUIC), Dudley Youth Offending Team and G-MAP Services Ltd.

STATEMENT OF AUTHORSHIP

Chapter 1 contains material that has been accepted for publication[1]. Chapter 6 was submitted as paper to the Association for the Treatment of Sexual Abusers (ATSA) and was awarded the Graduate Research Award and has been submitted for publication[2]. The case study content from Chapter 7 is a manuscript in preparation for publication. The authorship on each article indicates collaborative working. To clarify, I am the senior author and my supervisors Anthony Beech, Ian Mitchell and Glyn Humphreys are also named as authors. Additionally, Harriet Allen has been listed as an author on the paper for publication, and is a co-author on the manuscript in preparation because of her added contributions to the work conducted.

[1] Price, S.A., Beech, A.R., Mitchell, I., & Humphreys, G.W. (in press). The promises and perils of the emotional Stroop task: A general review and considerations for use with forensic samples. *Journal of Sexual Aggression*

[2] Price, S.A., Beech, A.R., Allen, H.A., & Mitchell, I. (submitted). An emotional Stroop task between offenders and controls and the identification of areas of brain activation through the use of functional magnetic resonance imaging (fMRI). *Archives of Sexual Behavior*.

CHAPTER 1

The Promises and Perils of the Emotional Stroop Task: A General Review and Considerations for Use with Forensic Samples

The aim of this chapter is to introduce the emotional Stroop task as an alternative measure to be used in the assessment of deviant sexual interest, and to encourage researchers to conduct future research using this paradigm. Factors that may influence Stroop results are considered as well as experimental and participant characteristics that should be taken into account when conducting experiments using the emotional Stroop task. Previous research findings using offender samples are reviewed, and suggestions for future research are suggested.

The following chapter was accepted for publication in the Journal of Sexual Aggression (in press). The format of the paper has been altered to achieve consistency with other chapters in this thesis.

Abstract

Assessing the implicit attitudes and beliefs of offenders has proven difficult over decades of work and research with offender samples. The purpose of this article is to outline the research related to the use of the emotional Stroop task and to encourage researchers and practitioners to use it as a measure to assess the implicit attitudes of offenders within their assessments. We consider the factors that may influence Stroop results along with the research considerations that need to be taken into account when using an emotional Stroop paradigm. Currently, a significant limitation of the emotional Stroop task is a lack of standard word stimulus sets that are able to distinguish between offender types. This limitation could seriously impede further development of this tool. Suggestions for future research are proposed.

The Promises and Perils of the Emotional Stroop Task: A General Review and Considerations for use with Forensic Samples

1. Introduction

The purpose of this paper is to argue that the Stroop task (Stroop, 1935) can be used as an alternative to traditional assessment methods for use with offender samples and to encourage researchers to build upon current research regarding implicit assessment measures in forensic research. This review is intended to inform researchers of the potential that the emotional Stroop task has in differentiating offenders from non-offenders, as well as differentiating between offender types at an information-processing level. The prospective applications of the emotional Stroop task and relevant advantages within the assessment, treatment and maintenance of offenders are discussed. Hopefully this review will also encourage researchers and clinicians to be aware of the factors that may influence Stroop results and to take these factors into consideration when planning studies using this tool. Judgments concerning these factors need to guide future experiments so that the experiments are administered and reported in a consistent manner.

2. The Stroop task

The traditional colour-word Stroop task was used by Stroop (1935) to examine interference between conflicting sources of information – colour word meaning and colour names (Stroop, 1935). McKenna and Sharma (1995) use the term 'intrusive cognitions' to refer to interference in information-processing produced by an irrelevant stimulus. In the colour-naming Stroop task the 'irrelevant stimulus' is the semantic meaning of a colour word.

In his first experiment, Stroop (1935) presented subjects with colour words printed in five ink colours (red, blue, green, brown and purple) and had subjects read the colour words aloud (a simple reading task). In the *experimental condition* the colour of the ink did not match the colour of the word presented (incongruent condition). In the *control condition* all colour words were printed in black ink. No significant differences were found between reading colour words presented in black ink versus reading colour words presented in incongruent colours. Stroop (1935) therefore concluded that ink colour had no effect on reading time (Stroop, 1935; MacLeod, 1991).

In a second experiment, Stroop (1935) presented the same list of colour words in incongruent ink colours but changed the task by asking participants to name the colour of the ink that the word was presented in. The control condition in this experiment was the presentation of solid colour squares. Significant differences were reported where subjects averaged a 47 second increase in the total time to respond between the two conditions (Stroop, 1935).

From these results, Stroop (1935) concluded that participants habitually found it easier to read a word than to name the colour that a word was presented in, which was indicated through the increase in response latencies on incongruent trials in the colour naming version of the study (Stroop, 1935). This difference in response latencies, called the 'Stroop effect', is said to provide a measure of interference in information-processing.

2.1 Interference theories

There are different accounts of how interference in information-processing occurs. The Relative Speed of Processing view (Cattell, 1886; Fraisse, 1969; Klein, 1964;

Morton & Chambers, 1973; Posner & Snyder, 1975) and the Automaticity view (Cattell, 1886; Laberge & Samuels, 1974; Logan, 1980; Posner & Snyder, 1975; Shiffrin & Schneider, 1977) propose that information-processing occurs in a sequential fashion.

The relative speed of processing view suggests that Stroop interference is a result of colour-naming taking longer to process than word reading. The theory assumes that: (1) the word response arrives at the response stage of information-processing first because subjects are faster to read words than they are to colour-name; and (2) it is the faster dimension (word reading) that interferes with the slower dimension (colour-naming). Cohen, Dunbar, and McClelland (1990) argue that if these assumptions hold true, then it should be possible to make colour information conflict with word reading by adjusting the presentation of the stimuli. However, researchers claim that relative speed of processing is not an adequate explanation for Stroop interference, because interference in word reading has not been observed when colour stimuli are presented prior to word stimuli (Cohen, Dunbar, & McClelland, 1990; Dunbar & MacLeod, 1984; Glaser & Glaser, 1982; MacLeod, 1991; MacLeod & Dunbar, 1988).

In the automaticity view it is assumed that schemas, developed through frequent use and stored in memory, set in motion a chain of unintentional cognitive-behavioural processes that are effortless and difficult to interrupt (Cox, Fadardi, & Pothos, 2006). Cohen, Dunbar, and McClelland (1990) suggest that the more often a processing pathway has been activated (i.e., word reading), the more easily it will be accessed for encoding a well-learned dimension in the presence of novel stimuli (ink colours). The authors believe that this holds particularly true when a better learned

dimension must be ignored, for a less well-learned dimension to be responded to (Cohen, Dunbar, & McClelland, 1990).

Although there have been no rigorous arguments against automaticity theories to date, MacLeod (1991) suggests that the Stroop effect is best explained using a Parallel Distributed Processing (PDP) ('connectionist') model (Cohen, Dunbar, & McClelland, 1990). This model states that parallel pathways are activated upon the presentation of the stimulus. In the case of the Stroop task, the parallel pathways are a colour information pathway and a word information pathway. In addition to the separate word and colour input units, two task demand (attention) units are present at the input level and are responsible for distributing attention to either the word or colour processing pathway. The task demand (naming colours of ink that stimuli are presented in) controls the degree to which the separate pathways are processed towards a response unit. Once one of the pathways (colour or word) crosses its response threshold a response is produced. Stroop interference is explained by this model as two simultaneously active pathways producing conflicting activation at their intersections – interactions that can occur at any point in the process, at multiple locations. Once a response is produced the conflict at the pathway intersection is no longer present, and the pathways are rendered inactive until the next stimulus presentation (Cohen, Dunbar, & McClelland, 1990; MacLeod, 1991; MacLeod & MacDonald, 2000).

2.2 The emotional Stroop task

In applied research, researchers have used the emotional Stroop task, which has extended Stroop stimuli to include emotion-inducing words, and/or disorder-related words. In the emotional Stroop task the inclusion of these diverse word stimuli allows

researchers to examine interference in other routes of information-processing above interference caused by colour-word stimuli (Dalgleish 1995; Williams, Mathews, & MacLeod, 1996). In this type of task congruent and incongruent trials do not exist because the stimuli presented are not colour words (Larsen, Mercer & Balota, 2006). Instead, control conditions are used in the emotional Stroop task and are generally in the form of neutral word stimuli (i.e., the word "desk") (MacLeod & MacDonald, 2000).

Participants administered the emotional Stroop task are still asked to name the ink colour of the stimuli presented, however, the content of the stimuli now includes words with emotional content (as opposed to colour words) and control words (with no emotional significance) (Larsen, Mercer, & Balota, 2006). Since the task is to name the ink colour of printed word stimuli, in order to have successful completion of the task participants must filter word name information, because it interferes with ink colour information (Larsen, Mercer & Balota, 2006; Mutter, Naylor & Patterson, 2005).

The reaction times (RT) to the words with emotional connotation are compared to the reaction times to the neutral or control words, and the difference (bias score) is said to be the result of interference caused by the emotional content of the word (Larsen, Mercer & Balota, 2006; Wentura, Rothermund, & Bak, 2000).

2.3 Theory behind the emotional Stroop task

In both the traditional colour-word Stroop, and the emotional Stroop, participants will automatically read the irrelevant words presented, when the task is instead to name the colour of ink the word is presented in (McKenna & Sharma, 1995). However, Dalgleish (2005) and Klein (1964) have argued that the emotional Stroop task

produces additional interference in information-processing caused not only by the semantic meaning attached to the word, but also because of the emotional content that reflects an individual's implicit attitudes, emotions, and motivations.

Implicit attitudes are evaluations that come to mind quickly when confronted with an object eliciting emotion, and are considered somewhat automatic because they are more engrained in an individual's cognitions and require less processing than weaker attitudes (Petty, Fazio & Briñol, 2008). Implicit attitudes have also been called 'automatic affective reactions' (Gawronski & Bodenhousen, 2006, p. 696) because the level of processing required is considered to be far less than the level of processing required in expressing explicit attitudes. Noteworthy is that an individual's implicit attitudes do not necessarily correspond with their explicit attitudes and an individual may or may not actually be aware that their implicit attitudes do not reflect their expressed opinions (Petty & Briñol, 2006).

Williams, Mathews and MacLeod (1996) suggest that if we adopt a PDP model (Cohen, Dunbar & McClelland, 1990) to explain interference displayed during the emotional Stroop task, then stronger connections in the word pathway will exist with words that are considered to be emotionally salient to a particular subject. Concomitantly, participants will display higher RTs to the emotionally salient words.

The task demand (naming the colour of ink that stimuli are presented in) in the emotional Stroop remains the same and still controls the degree to which the colour information pathway and word information pathway are processed towards a response unit. However, it is proposed that there is now stronger activation to overcome in the word processing pathway, until one of the pathways crosses its threshold and a response is produced (Cohen, Dunbar & McClelland, 1990; MacLeod, 1991; MacLeod & MacDonald, 2000; Williams, Mathews & MacLeod, 1996).

The emotional Stroop task has been applied in victimization studies and across a wide range of disorders including: phobias and anxiety disorders; panic disorders; depression; addictions; and conduct disorder (see Table 1 for key references). More recently, research has been conducted using the emotional Stroop task to examine patterns in bias scores (Stroop interference) of violent and sexual offenders using sexual word stimuli (Price & Hanson, 2007; Smith & Waterman, 2004).

Table 1

Key References: The Emotional Stroop Task

Research Focus	Key References
Phobias and anxiety disorders	Holle, Neely & Heimberg (1997); Koven, Heller, Banich & Miller (2003); Mathews, Mogg, Kentish & Eysenck (1995); Mogg, Bradley, Williams & Mathews (1993)
Panic disorders	McNally, Amir, Louro, Lukash, Riemann & Calamari (1994); McNally, Riemann & Kim (1990); McNally, Riemann, Louro, Lukach & Kim (1992)
Depression	Hill & Knowles (1991); Klieger & Cordner (1990); Segal, Truchon, Horowitz, Gemar & Guirguis (1995)
Addictions	Banich et al. (2007); Cox, Fadardi & Pothos (2006); Gross, Jarvik, & Rosenblatt (1993); Kramer & Goldman (2003); Waters & Feyerabend (2000); Wertz & Sayette (2001)
Conduct disorder	Dolan & Rennie (2007)
Victimization	Cassiday, McNally & Zeitlan (1992); Dubner & Motta (1999); Foa, Feske, Murdock, Kozak & McCarthy (1991); McNally, Clancy, Schacter & Pitman (2000); McNally, Metzger, Lasko, Clancy & Pitman (1998)
Belief in a just world	Hafer (2000)
Psychopathy	Hiatt, Schmitt & Newman (2004)
Violent/Sexual Offenders	Ó Ciardha & Gormley (2009) (pictorial Stroop); Price & Hanson (2007); Smith & Waterman (2003, 2004)

A consistent finding with all variations of the emotional Stroop task is that people have difficulty ignoring the word meaning while naming the colour (Johnson & Hasher, 1987; Salo, Henik, & Robertson, 2001). It is then not surprising that when the word stimuli directly relates to the participant, higher response latencies result (Logan & Goetsch, 1993).

3. Factors Influencing the Stroop Effect

As with any other method of assessment, there are certain considerations and limitations that need to be addressed before entering into a research experiment that uses Stroop methodologies. The following section will highlight certain factors that may influence Stroop results that should be taken into consideration when planning an experiment using the emotional Stroop task with any population.

3.1 Experimental design

Before conducting an experiment using the emotional Stroop task it is important to consider elements of the experiment that may affect Stroop results. For example, Verhaeghen and Meersman (1998) tested variations in the type of baseline used, the number of stimuli presented and the way in which reaction times are measured in the Stroop task. They found that presenting coloured patches produced a larger Stroop effect than noncolour words (i.e., neutral words), and coloured symbols, in the baseline conditions. This seems to be the case because coloured patches tend to be more distracting and have more of an effect (MacLeod, 1991). Verhaeghen and Meersman (1998) also found that presenting more than one stimulus at a time produced a larger Stroop effect than the sequential presentation of items; and that using the printed stimuli and stopwatch method of measuring reaction times produced

a larger Stroop effect than computerized testing and recording. An advantage of manually recording reaction time is that the participant is not required to learn any response mappings; however, it can increase error variance into the response time measure (Mutter, Naylor, & Patterson, 2005).

The advantages of using a single-trial computerized version of the Stroop task (instead of printed stimuli and a stopwatch) are that: (1) it allows for a measure of per-item reaction times as opposed to a summation of reaction times across a large stimulus set; (2) it allows for the presentation of different word and trial types to be presented randomly; (3) errors for single words can be omitted when calculating reaction time measures; and (4) displaying the stimulus centrally fixated (as opposed to having all stimuli on a card in sequence) does not require the subject to generate eye movements or to move attention down columns of stimuli, an advantage that is especially important when presenting Stroop tasks in fMRI studies (Salo, Henik, & Roberston, 2001).

Importantly, participants are asked not to correct errors while completing the Stroop task, because this impedes in the calculation of response latencies, and errors are simply removed from the data. Research in cognitive psychology, and cognitive neuroscience, has demonstrated that participants tend to shift their response behaviours following conflict trials (e.g., incongruent trials in a Stroop task) and errors over the course of an experiment (Holmes & Pizzagalli, 2007). Botvinick, Nystrom, Fissell, Carter, and Cohen (1999) note that incompatible responses trigger control processes that result in adjustments in performance. Adjustments in performance post-conflict (i.e., post-incongruent trial) are exhibited by participants through decreased error rates and slowed reaction times following the particular conflict trial - a phenomenon that has been coined the Gratton Effect (Gratton, Coles,

& Donchin, 1992). The same pattern of behavioural adjustment is found following an error, whereby subjects tend to display patterns of increased accuracy and slowed reaction times - a post-error adjustment known as the Rabbitt/Laming Effect (Laming, 1979; Rabbitt, 1966). Post-error slowing has also been observed and identified as a potential confounding variable. Although error trials are typically excluded from analysis, it is suggested that post-error trials may not be comparable to other correct responses and should be considered for exclusion in the analysis of response latencies as well (Egner & Hirsch, 2005; Kerns et al., 2004). In practice, we would suggest that researchers consider the possibility of post-error and post-conflict effects, and to run analyses both with and without these effects to explore whether they have a significant effect on the findings of the study.

3.2 Stimulus set considerations

Since the word stimuli presented in the emotional Stroop task are rarely the same (i.e., there are no standard sets of word stimuli used to examine sexual preoccupation, depression, etc…), it is crucial that efforts are made to minimize potential confounds that naturally accompany the word stimuli in this task. Therefore, it is essential that stimuli are matched on all lexical features that could influence response latencies, such as word length and word frequency (Burt, 2002; Larsen, Mercer, & Balota, 2006). Specifically, research in the area of word recognition speed has shown that infrequently used words take longer to recognize than frequently used words. Concomitantly, participants are slower to colour-name low frequency words compared to high frequency words (Larsen, Mercer, & Balota, 2006). Additionally, the length in letters of word stimuli is thought to influence recognition speed in that,

as might be expected, longer words take more time to process than shorter words (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004).

Larsen, Mercer, and Balota (2006) performed a literature review on Stroop research to examine whether emotional Stroop studies conformed to the assumption that word stimuli should be matched for word length and frequency. Interestingly, out of 72 empirical articles that used the emotional Stroop task, only 32 of those articles provided the lists of words that were used in the research. The authors found that both negative and positive words used in these studies were significantly longer in length than the neutral words; and that negative words and disorder-specific words used in the literature were significantly rarer than the neutral words. Thus, when Stroop interference effects are calculated by subtracting neutral words from experimental words, if the experimental words are much longer and rarer than the neutral stimuli, one cannot disregard that the effects may be due to differences in lexical characteristics between stimulus categories. Word frequency and word length are both lexical features that contribute to both speed of word recognition and latencies in colour-naming. Accordingly, these characteristics should ideally be controlled for across all word stimulus categories (Balota, et al., 2004; Larsen, Mercer, & Balota, 2006).

The ability to interpret the emotional Stroop task depends upon the selection of neutral stimuli because they provide the baseline by which interference is measured. When conducting an emotional Stroop task the selection and matching of neutral and control words can change performance on the task, as can the way in which they are presented, which could explain some of the current inconsistencies in the clinical literature (Salo, Henik, & Roberston, 2001). For example, Salo, Henik, and Robertson (2001) compared response times of neutral, congruent and incongruent

stimuli using three different presentation conditions: (1) blocked format (all neutral stimuli presented on a screen at the same time); (2) blocked single-trial format; and (3) single-trial randomized format. The authors found that RTs to neutral stimuli were faster when they were presented in a blocked format, than they were when presented in blocked single-trial and single-trial randomized formats. The same RT changes between presentation types were not observed for the congruent and incongruent conditions. Consequently, the authors suggest that these findings make it difficult to interpret interference reported in the clinical literature because there is inconsistent use of neutral stimuli (Salo, Henik, & Roberston, 2001).

Larsen, Mercer, and Balota (2006) suggest that anything that influences word recognition speed could also influence colour-naming speed, which would in parallel influence how Stroop interference is interpreted. The authors note that Stroop interference contains two components that influence speed of colour-naming: (1) RT difference between emotion-inducing and control words, and (2) lexical difference between words with emotional connotation and control words. If the latter are controlled for, it is possible that smaller or larger emotional Stroop effects would result because the additional interference caused by lexical characteristics would be removed from all word categories, concomitantly resulting in a purer reflection of the actual emotional Stroop interference (Larsen, Mercer, & Balota, 2006).

3.3 Participant characteristics

Certain participant characteristics have been said to affect the way in which participants are able to successfully complete the Stroop task. These include: age, reading skill and gender differences.

3.3.1 Age

Age would appear to affect one's ability to successfully complete the Stroop task (Houx, Jolles, & Vreeling, 1993; Kieley & Hartley, 1997; Spieler, Balota, & Faust, 1996; West & Bell, 1997). Research has suggested that greater Stroop interference is experienced by adults over the age of 60, and that they occasionally experience greater facilitation, indicating that older adults are less able to ignore the irrelevant word dimension (Mutter, Naylor, & Patterson, 2005). Comalli, Wapner, and Werner (1962) have found that Stroop interference remains constant across middle adulthood and begins to increase in the 65 – 80 year-old age group. However, Johnson, Bouchard, Segal, Keyes, and Samuels (2003) report increased ability on the Stroop task until the age of 60 when this phenomenon begins to reverse.

Since the Stroop task is a test of inhibitory control, it is often assumed that increased Stroop interference exhibited in older adults represents a deficiency in these inhibitory mechanisms rendering the Stroop task age sensitive. However, Verhaeghen and Meersman (1998), argue that it is simply a general slowing effect that accounts for an increase in Stroop interference. The authors tested two groups using the Stroop task against several moderator variables: baseline condition (i.e., patches of colour, coloured noncolour words, coloured symbols); presentation of stimuli (one stimulus at a time vs. more than one at a time); and type of reaction time recording (printed stimuli recorded with a stopwatch versus computerized presentation and logging of reaction time). Verhaeghen and Meersman (1998) did not find any significant effect of age, thus concluding that the interference effect in the Stroop task is not age sensitive, and that a general slowing effect is responsible for any perceived effects of age in Stroop interference.

It is important to note that many of the studies that have been conducted to date examining age-sensitivity of the Stroop task on interference effects (Kieley & Hartley, 1997; Spieler, Balota, & Faust, 1996; West & Bell, 1997) have used samples that represent extreme age ranges (participants in their 20's compared to participants aged 65 and older), rendering it difficult to gauge when perceived declines in successful Stroop task completion have their onset.

3.3.2 Reading skill

Reading skill has displayed a high correlation with the Stroop task (Johnson et al., 2003; MacLeod, 1991) whereby skilled readers tend to have more difficulty with the Stroop task when word meaning and print colour are incongruent (i.e., word red written in blue ink) (Mutter, Naylor, & Patterson, 2005). In very young children (lacking basic reading skills), colour-naming speed exceeds word-reading speed, and there is little interference effect (Schiller, 1966).

3.3.3 Gender differences

Gender differences on the Stroop task have also been studied and, in general, women appear slightly faster at naming colours and word scores (Johnson et al., 2003; MacLeod, 1991). However, gender differences have not been found for word reading or for colour-word Stroop interference (Jorgenson, Davies, Opella, & Angerstein, 1980, 1981; MacLeod, 1991). Offender research involving the emotional Stroop task has not yet examined potential gender differences using these specialized word stimuli because all samples used to date have used all male participants (Hiatt, Schmitt, & Newman, 2004; Price & Hanson, 2007; Smith & Waterman, 2004). However, Smith and Waterman (2005) examined the presence of gender differences in information-

processing bias for aggression-themed words in a sample of undergraduate students and found that male subjects were significantly slower to colour-name words relating to overt aggression (i.e., physical aggression), as opposed to covert aggression (i.e., malicious gossip). The women were slower to colour-name words relating to covert aggression however this finding was not significant. Stimulus salience (i.e., previous experience) was thought to be a plausible explanation for these differences in significant findings when the self-reported levels of aggression on behalf of the men were considered (Smith & Waterman, 2005).

4. The emotional Stroop task using offender samples

Implicit measures may be useful for tapping into an individual's automatic tendencies (Petty & Briñol, 2006). The emotional Stroop task provides one example of this type of measure and the task has been widely used in psychological research to assess implicit attitudes and associations; however the use of the Stroop task in the forensic field is quite small. The research that is available in this area seems to indicate similar results to those reported in other branches of psychological research that have used this paradigm. In other words, when the word stimuli directly influence or relate to the participant, higher response latencies result.

4.1 Stroop effects in psychopaths

Information-processing abnormalities, specifically, a perceived immunity to unattended contextual information, has been observed in psychopathic individuals (Forth & Hare, 1989; Hare, 1978; Harpur & Hare, 1990; Kosson, 1996, 1998). These results have led to the assumption that psychopaths will demonstrate less, rather than more, Stroop interference. Hiatt, Schmitt and Newman (2004) examined this

assumption exploring Stroop interference in a group of primary psychopaths. The authors presented three separate Stroop variations to subjects: (a) a basic colour-word Stroop task; (b) a picture-word Stroop task; and (c) a spatially separated colour-word Stroop task. It was found that psychopathic participants were just as affected as controls by the presence of incongruent contextual information in the traditional Stroop colour-word test. However, in the picture-word and spatially separated colour-word versions of the Stroop task, psychopathic participants displayed less interference than control subjects. Therefore, subjects with psychopathy demonstrated reduced Stroop interference only when a slight spatial separation between relevant and irrelevant stimulus dimensions was present. Based on the results from these experiments Hiatt et al., (2004) concluded that the critical determinant of the attentional performance of psychopathic individuals was dependent on the spatial relationship of the stimuli (i.e., whether stimuli overlap, as in the colour-word version, or spatially separated, as in the picture-word Stroop task). Hiatt et al. subsequently concluded that their results confirm the hypothesis that psychopathic individuals are unable to attend to incongruent contextual information when it is spatially separated.

4.2 Stroop effects in sexual and violent offenders

Smith and Waterman (2003) examined information-processing bias for aggression-themed words and self-reported scores of aggression between violent offenders (n=33), nonviolent offenders (n=17) and undergraduates (n=30). The sample of violent offenders displayed significantly more interference for aggressive and negatively-themed words than the undergraduates. Additionally, increased Stroop interference was observed with higher anger scores on the Aggression Questionnaire (BPAQ; Buss & Perry, 1992); a result that was reflected in the undergraduate sample

as well. The authors suggest that information-processing biases for aggressive material are present in individuals who are prone to engage in aggressive behaviours.

Further work conducted by Smith and Waterman (2004) examined Stroop interference effects of sexual and aggressive word stimuli for violent and sexual offender populations. They tested 10 sexual offenders, 10 violent offenders, 10 non-violent offenders and 13 undergraduates using an emotional Stroop task that consisted of the random presentation of neutral words, aggression-related words, positive and negative affect words, colour words and sexual words all matched for word length and word frequency. It was found that sexual offenders displayed longer latencies to sexually-themed words, while violent offenders took longer to name the colour of aggression-themed words. Additionally, the more aggressive rapists displayed longer latencies to the aggression-themed words, compared with men who were less overtly violent (i.e., those charged with indecent assault). However, for the sexual material alone, different types of sexual offenders (i.e., child molesters versus rapists) could not be distinguished between. Smith and Waterman (2004) concluded that the way in which information is perceived by different types of offenders activates contrasting cognitive associations, and that these associations are reflected in the Stroop latencies.

Price and Hanson (2007) replicated Smith and Waterman's (2004) study using a sample of 15 child abusers, 15 rapists, 15 violent non-sexual offenders, 15 non-violent non-sexual offenders and 15 community (non-offending) participants. The general direction of the Stroop findings was the same between the two studies. For example, Price and Hanson (2007) also found a significant difference between the groups with rapists being significantly slower to name the colour of sexually-themed words compared to the community sample. Rapists also tended to be slower (a non-

significant effect) than the child abusers, violent offenders, and non-violent non-sexual offenders to name the colour of sexually-themed words.

Both groups of sexual offenders in this study had significantly larger mean RT's to the neutral words. Additionally, the child abusers significantly differed from the community and non-violent non-sexual offender groups when naming the colour words, and were slower than the rapists and violent offenders, but this difference was non-significant. Overall, it appears that the sexual offenders in this study were slower to colour-name and this could have been due to a confounding variable that was not included in the analysis of the study (i.e., level of executive function or reading skill).

In the same study, Price and Hanson (2007) developed a new word stimulus set consisting of child abuse words, rape-themed words, intimacy-themed words, and neutral words. These newly developed word categories were derived because the sexual words presented in Smith and Waterman's (2004) study were a mixture of both rape-related and child abuse-related words. It was the intention of the authors to explore whether different types of sexual offenders could be differentiated between if presented with separate word lists relating to their specific type of sexual offending behaviours[3]. No significant differences were found between offender types using the new words.

Price and Hanson suggested several explanations as to why the second experiment did not yield the anticipated results. First, they proposed that because the words were not empirically derived the stimuli may not have actually reflected the motives, thoughts or feelings experienced by sexual offenders. Second, Price and Hanson note that their sample of child molesters were mainly intra-familial offenders (80%), and were generally of low risk to reoffend (Hanson & Thornton, 2000),

[3] It is worth mentioning that the newly derived word stimuli for the second experiment in this study were not empirically derived through any formal assessment (i.e., using a panel of offenders), but instead were intuitively derived by the authors.

suggesting that perhaps the activation that needs to be overcome in the word processing pathway of intra-familial child abusers regarding the emotional salience of the sexually deviant child abuser stimuli may actually be weaker than that of extra-familial child abusers, and more reflective of the control sample.

There is converging evidence that the strength of affective interference would be less for intra-familial child abusers than it would be for a high-risk sexual abuser displaying stronger, more deviant attitudes and cognitions related to deviant sexual stimuli. Recent research using Rapid Serial Visual Presentation (RSVP), which uses pictures of children as stimuli, has found significant differences between extra-familial child abusers and controls, concomitantly yielding no significant differences between controls and intra-familial child abusers. This would suggest that the emotional salience of deviant sexual stimuli is stronger for extra-familial child abusers than for intra-familial child abusers (Flak, 2009).

It seems appropriate now to reconsider that a key limitation of the research thus far using the emotional Stroop paradigm is a significant lack of standardization surrounding the word lists used. Despite efforts made to minimize potential confounds surrounding the word stimuli in this task by matching words on lexical characteristics, efforts should be made to develop standard sets of word lists that will be able to consistently discriminate between different types of offenders.

Additional work conducted by Ó Ciardha and Gormley (2009) used a modified pictorial Stroop task on a sample of 24 sexual abusers who had committed sexual offences against children and 24 controls. Some of the images of the children and adults were taken from the Not Real People (NRP) image set (Laws & Gress, 2004, visit http://pacific-psych.com/products/nrp-not-real-people-visual-stimuli-set/). Additional images were created by the authors using a similar morphing technique

used in the creation of the NRP images. The backgrounds of all images were removed and were coloured in leaving the target colour (i.e., red, green blue and yellow) in all image trials located on the to-be-ignored stimulus (i.e., target image).

Ó Ciardha and Gormley (2009) found that overall the sexual abusers in the study did not significantly differ from non-offenders on the pictorial modified Stroop task. However, when subgroups of sexual abusers were compared to the control subjects, offenders with an admitted sexual interest in children did not significantly differ in their average reaction times to adults and children, while control subjects did. The authors note that their results should be interpreted with caution considering the small sample sizes within the subsets of the groups, and the heterogeneity of the offender sample.

5. The advancement of work with offender samples

Although Price and Hanson's (2007) second experiment did not yield the anticipated results, it does raise some interesting questions regarding how best to conduct emotional Stroop task experiments with offender samples that will hopefully yield significant differences between offender types and demonstrate the tool to be useful as an implicit measure of assessment in offender research. Currently, the use of the emotional Stroop task within offender research is far from robust. However, recent research using this paradigm with offenders (Hiatt, Schmitt, & Newman, 2004; Ó Ciardha & Gormley, 2009; Price & Hanson, 2007; Smith & Waterman, 2003, 2004) has revealed that the tool could hold promise in its application within the assessment and treatment of offenders.

In addition to being a minimally invasive and cost-effective measure, the Stroop task requires very little training to be able to administer. An important feature

of Stroop techniques is that they have the potential to reduce fakeability and social desirability concerns associated with self-report measures. Self-report measures of participants' interests and desires are intrinsically vulnerable to faking because they rely on the participant to communicate their preferences. These concerns are particularly pertinent with offender populations, because offenders may want to deny their deviant thoughts or behaviours (Abel, Jordan, Hand, Holland, & Phipps, 2001; Flak, Beech, & Fisher, 2006). The emotional Stroop task may circumvent these problems as it represents an implicit measure of interference in information-processing by using emotionally significant words to generate interference (Larsen, Mercer, & Balota, 2006; Wentura, Rothermund, & Bak, 2000). Consequently it is less likely to be subject to the limitations associated with explicit self-report measures (Flak, Beech, & Fisher, 2006).

If the emotional Stroop task is applied in offender assessments it could allow researchers to evaluate the implicit attitudes of offenders (without having to rely on their word) and could prove very useful in obtaining additional information required for more reliable offender assessments. This tool could also be used as a method for motivating offenders to be more truthful regarding the information that they disclose.

Importantly, flexibility around the choice of word stimuli used in these types of studies makes it possible to cater the word lists to the individual, and to examine an endless range of implicit thoughts and attitudes towards offending. This could become particularly valuable in identifying possible treatment targets.

Should researchers choose to use a pictorial Stroop task with offender samples, this methodology can have some advantages over the use of word stimuli related to the salience of information portrayed in the images used. For example, the age, gender, ethnicity, and pose and emotion portrayed in the image can all be

manipulated, a task that is more difficult to accomplish with the use of words. However, this manipulation of the target stimuli may not be appropriate in all cases (i.e., with offenders interested in sadistic acts), and the use of words may be a more appropriate choice (Ó Ciardha & Gormley, 2009b).

Smith and Waterman (2003) suggest that treatment efficacy may also be measured using this method. If applied in a research study, one would use the response latencies from the emotional Stroop task as being indicative of the strength of implicit attitudes. Presumably, post-treatment attitudes would differ from pre-treatment attitudes (i.e., they would become weaker).

Based on the fact that research involving the emotional Stroop task and offenders is so preliminary, one could only presume that the potential applications suggested in this paper are viable. However, this lack of research leaves plenty of room for the tool to be used in an attempt to answer a wide range of research questions regarding the most appropriate uses for this tool with offenders. Where the emotional Stroop task would be most beneficial and valuable to offender assessment, management or treatment remains a research question that needs to be answered. Additionally, future research might be able to determine whether the Stroop task could play a role in the prediction of recidivism or as a monitoring tool.

6. Discussion

As previously mentioned, few studies to date have used the emotional Stroop task to examine control processes in general and the control of responses to emotionally salient stimuli in particular. It is the intention of this article to encourage researchers in the forensic field to consider this paradigm for the assessment of implicit attitudes of different offender types. Its use with samples of violent and sexual offenders

suggests that the emotional Stroop technique holds promise, however the research needs to progress to a stage where the Stroop task is able to yield significant differences between offender types (Price & Hanson, 2007; Smith & Waterman, 2004).

Most recently, researchers have turned to the presentation of Stroop stimuli in MRI machines (Adleman et al., 2002; Banich et al., 2007; Price, Beech & Allen, 2010) to examine the additional component of brain patterns that occur when individuals are prevented from completing an automatic task (i.e., simply reading the word presented). Functional neuroimaging techniques are used to identify which brain regions are most active while completing specified tasks. Experiments using the traditional colour-word Stroop task show that the anterior cingulate cortex (ACC) seems to display the greatest activation (increased blood flow) during incongruent conditions. It is thought that the anterior cingulate plays a role in attentional control and the detection of conflict; however these assumptions have been under considerable debate (Carter et al., 1998; Leung, Skudlarski, Gatenby, Peterson, & Gore, 2000; MacLeod & MacDonald, 2000; Veen & Carter, 2005).

Importantly, it is also thought that the ACC also plays a role in affective processing by being involved in assessing the salience of emotional information and the regulation of emotional responses (Bush, Luu, & Posner, 2000; Whalen, et al., 1998). This is especially important information to consider when administering an emotional Stroop task and research in this area using fMRI techniques could contribute a great deal to the field, and towards informing researchers further about the information-processing patterns that occur while completing this task.

Like all other methods of offender assessment, the emotional Stroop task has its limitations and considerations that should be taken into account prior to advancing

into an experiment using this paradigm. It is the responsibility of researchers to be aware of the factors that influence Stroop results, and to adopt the most appropriate ways of administering the tool in order to get the most pure results reflective of the implicit attitudes of their samples, and to ensure that there is nothing inherently wrong in the way that they have implemented the task.

Price and Hanson (2007) suggest that in order to accomplish this, the first initiative should be to take a step back in the emotional Stroop research with offender samples by developing empirically-derived word stimuli that are based on the accounts of offenders. This method may then hold more promise in differentiating between offender types, as these accounts would be a reflection of specific types of implicit associations. Information on how Stroop stimuli are developed and matched with control stimuli could be a good start to solidifying the research available. The development of standard sets of Stroop stimulus sets catered to different types of offenders could aid in to the standard use of the Stroop task within forensic settings.

It is important that researchers carefully consider their word stimulus sets (including lexical characteristics) and the way in which reaction times are recorded as potential experimental confounds because these factors have been known to influence Stroop effects (Burt, 2002; Larsen, Mercer, & Balota, 2006; Salo, Henik, & Robertson, 2001). It is also important to consider and control for as many participant variables as possible, such as age and reading skill because participants may exhibit differential Stroop effects based on these factors (Kieley & Hartley, 1997; Mutter, Naylor, & Patterson, 2005; Schiller, 1966; Spieler, Balota, & Faust, 1996; West & Bell, 1997).

Despite the research available that provides evidence for the ability of the Stroop task to display interference in automatic information-processing, failure to

consider the factors that influence Stroop latencies has yielded mixed results. However, it may be that if these limitations and considerations are addressed properly, the result could be more concrete offender research using this paradigm, and a strong contribution to the research involving the implicit attitudes of offenders. The purpose of this paper is to encourage researchers to build upon current research regarding the potential uses of the emotional Stroop task within offender assessment. It is also meant to encourage researchers and clinicians to be aware of the factors that may influence Stroop interference effects to ensure that future experiments are administered and reported in a consistent manner.

CHAPTER 2

General Methodology and the Development of an Emotional Stroop Word Stimulus Set

The purpose of this chapter is to outline the general methodology used throughout the studies discussed in the chapters which follow. An additional aim of this chapter is to describe the development of a set of emotional Stroop word stimuli that were empirically derived to measure deviant sexual interest. The set of newly derived Stroop stimuli were developed with the intention of being able to yield significant differences between sexual offenders and non-sexual offenders, as well as between subtypes of sexual offenders. By concentrating on the development of appropriate word stimuli, the improvement of this measure may enable the Stroop task to be used as a possible alternative to traditional methods for assessing sexual interest preferences.

Introduction

Previous research using the emotional Stroop task with offenders (Price & Hanson, 2007; Smith & Waterman, 2003, 2004) has focussed on examining Stroop effects in adult samples. Based on the research surrounding the potential effects that participant characteristics can have on Stroop results (as described in Chapter 1) it was an aim of this thesis to examine Stroop bias effects displayed by participants from different age groups, offending groups and levels of executive function. Participant characteristics that were considered when deciding on the apparatus and materials to be used in the experiments included age, reading skill and level of executive function. Chapter 1 also described a number of experimental variables that should be taken into consideration when conducting an emotional Stroop experiment. Experimental variables that were considered when planning the studies in this thesis included the type of baseline, number of stimuli presented at a time, method for recording reaction time, and lexical features of the word stimulus sets. Adherence to the consideration of participant and experimental variables across the studies was practiced in order to minimize the effects of confounding variables and to allow for a cross comparison of the studies that comprise this thesis.

The following description of the procedures, apparatus and materials were used for the studies outlined in Chapters 3, 4 and 5. The additional methods involved in the use of fMRI are detailed in Chapter 6.

A description of the methods and procedures is presented below as well as support for their use when considering the potential confounding effects that participant characteristics and experimental variables can have on Stroop effects. In

order to control for these effects the methods used throughout this thesis were kept consistent across the studies.

Method

Apparatus/Materials

Information concerning participants' age, presence of a learning disability, handedness, and whether they had an existing criminal record was collected. Information of offenders' index offence, prior offences, victim age, victim gender and relationship to the offender was collected in order to aid in the sub-categorization of offender groups.

Participants completed the British Picture Vocabulary Scale (BPVS-II; Dunn, Dunn, Whetton, & Burley, 1997); the Hayling and Brixton tests of executive function (Burgess & Shallice, 1997); and a shortened version of the Beliefs About Children Scale (BACS; Beckett; 1987). Scoring items on the BACS yields 2 subscale scores: cognitive distortions (CD) and emotional congruence with children (EC).

The BPVS was used to confirm that participants had an appropriate level of understanding of the words that were presented in the emotional Stroop task, and to ensure that the effects of reading skill deficits were minimized. For example, since little interference effects are experienced by individuals lacking basic reading skills (Schiller, 1966) it was of interest to test vocabulary to ensure that all participants were able to experience interference in information-processing that is caused by processing of word information. The BPVS is intended for use with subjects between the ages of 3:00 and 15:08 as a measure of verbal ability or verbal intelligence. Examiners are to calculate a raw score, then convert the raw score into a standardized score and

percentile rank listed in the Testbook provided. Most of the participants in the current studies were above the age limit where standardized scores and percentile ranks were provided. Raw scores were therefore calculated for each participant and analyses were conducted on these scores. The BPVS was used as a measure of vocabulary understanding in anticipation of the level of learning difficulties that may be prevalent in the youth offender samples. Extended use of this measure with the adult samples, despite the lack of provision of standardized scores, was due to the desire to maintain consistency across the studies, and to allow for comparison between the different age groups.

The Hayling Sentence Completion test is a measure of response initiation and response suppression. It consists of two sets of 15 sentences each missing the last word. In the first section the examiner reads each sentence aloud and the participant has to simply complete the sentences, yielding a simple measure of response initiation speed. The second part of the Hayling requires subjects to complete a sentence with a nonsense ending word (and suppress a sensible one), giving measures of response suppression ability and thinking time. Average inter-rater reliabilities of up to 96.0% have been found for final scoring of the Hayling test (Bielak, Mansueti, Strauss & Dixon, 2006).

The Brixton test is a visuospatial sequencing task that measures the ability to detect rules in sequences of stimuli. It takes between five and ten minutes to administer, and yields an easily understood scaled score between 1 and 10. The Brixton consists of a 56-page stimulus booklet, each page displaying two rows of five circles numbered between 1 and 10. One circle on each page is filled in with blue colour. It is the participant's task to identify where the blue dot would be on the following page each time, based on a pattern or rule governed in the previous page.

The outcome measure of this task is the total number of errors across 55 trials (Burgess & Shallice, 1997).

Computerized versions of the emotional Stroop task were presented randomly on a Toshiba laptop on a 12"x9" screen. Word stimuli were presented using version 2.0 of E-Prime software (Psychology Software Tools, Inc.). The colour-identification response latencies for each trial were detected and recorded by a five-button serial response box (Psychology Software Tools, Inc., model 200A) with four task-specific coloured buttons identified (green, red, blue and white). All studies used button-press response recordings because of the noisy environments the research was often carried out in (i.e., prisons, group treatment settings), and in order to maintain consistency across the experiments. Note that the fMRI study presented in Chapter 6 necessitated the use of button-press response recordings to reduce participants' movement within the scanner, and thus minimize movement induced artefacts in the data. Two emotional Stroop word stimulus sets were used in each experiment: (1) Smith and Waterman (2004, see Appendix A) word stimulus set; and (2) a set of newly derived word stimuli that were intended to reflect deviant sexual interests of sexual abusers.

Word Development

Price and Hanson (2007) derived a new set of word stimuli to be used to measure deviant sexual interests. This was necessary as it was assumed that the sexual words presented in Smith and Waterman's (2004) study were of a general sexual nature, and consequently not refined enough to distinguish between subtypes of sexual abusers. In testing their set of newly derived stimuli, Price and Hanson (2007) found no significant differences between the word types. The authors proposed that the stimuli they had developed may not have reflected the motives, thoughts or feelings

experienced by sexual offenders because the words were not empirically derived. Therefore, the aim of empirically deriving a new set of word stimuli was to divert away from general sexual terms used in previous research and to ensure that the words used as stimuli were a true reflection of deviant sexual interest preferences of sexual abusers.

Four groups of individuals aided in the development of a new word stimulus set for the purpose of this thesis: workers from the Lucy Faithfull Foundation (13), workers from G-MAP services (5), and two groups of sexual offenders (27). The newly-derived word stimuli were created from the responses of these groups to question 6a of the Relapse Prevention Questionnaire (RPQ; Beckett, Fisher, Mann & Thornton, 1996): How would you describe who would be most at risk from you? It was thought that by deriving the word stimuli from offenders' accounts of what they find sexually stimulating or attractive, a purer reflection of deviant sexual interests would result. Input was requested of workers because of their experience and daily interactions with sexual abusers.

A focus group was held with all staff members from the Family Programme at the Lucy Faithfull Foundation (LFF) in Alvechurch, United Kingdom in December 2007. The Lucy Faithfull Foundation is a child protection charity that aims to reduce a child's risk of being sexually abused. They provide specialist assessment of families where sexual abuse has occurred, or is believed to have occurred. The majority of their work is undertaken by the support of the Family Civil Courts. The workers at the LFF are all regarded as expert witnesses in the Family Courts, and most are specialists that assess adult perpetrators, alleged adult perpetrators and the partners of perpetrators. The focus group was facilitated by a worker at the LFF and the researcher.

The focus group consisted of 13 workers from across the United Kingdom. 193 word stimuli were derived throughout the session. In addition to the words derived from question 6a of the RPQ, workers provided the researcher with additional sexual action words which they felt were commonly used by the sexual offenders that they had contact with. Duplicate words were then removed from all of the responses resulting in a total of 113 word stimuli of how sexual abusers might describe victims at risk from them.

The same focus question was asked of 10 workers at G-MAP services, Sale, Cheshire in hard copy format. G-MAP is an independent outpatient organisation that offers a range of services to young people who display inappropriate sexual behaviours, their families, carers and the professionals that work with them. Five workers responded to the focus question and a total of 49 words were submitted, 40 of which were non-duplicates to the word stimuli derived from the focus group at LFF.

Additionally, the researcher visited the Selly Oak Probation office in Birmingham, UK to review the files of sexual offenders' responses to question 6a of the RPQ. Ten questionnaires were reviewed yielding nine word stimuli, two of which were non-duplicates (i.e., trusting, shy). Six of the ten questionnaires reviewed did not provide adequate responses. For example, three sexual abusers provided the answer "anyone/everyone is at risk from me", and three sexual abusers responded by providing age ranges of those at risk from them.

Finally, the UK Cornwall Probation office provided an additional 17 RPQ questionnaire responses from sexual abusers to question 6a which were reviewed by the researcher. This resulted in the extraction of 23 word stimuli, 9 of which were non-duplicates. Nine of the sexual abusers from this sample provided inadequate answers such as "nobody is at risk because I am treated", "myself", or provided age

ranges of potential victims. Unfortunately, many of the questionnaires reviewed were incomplete and responses from the offenders were not overly descriptive. Therefore a majority of the resulting words were derived from the front-line workers as opposed to the offenders themselves.

Once word stimuli were generated from these four sources they were arranged in alphabetical order and given to five PhD students at the University of Birmingham to be categorized into 4 separate categories: emotional/personality descriptors (EPD), sexual actions (SA), physical descriptors (PD) and aggressive/rape words (AR). Responses to this categorization task were then reviewed and words where 3 or less PhD students agreed on the word category were removed. This resulted in 75 word stimuli in total from the 4 different categories.

Further to this analysis, the aggressive/rape words were removed from the study because there were very few words belonging to this category, and the words did not necessarily reflect a rape theme, nor did they 'fit' with the other word stimuli. For example, in general, the words seemed to reflect a more aggressive word type (i.e., slag, slapper, bitch), and were likely put into the rape word category because of their overarching theme of aggression. Additionally, it was felt that the aggressive words might present as too sensitive in nature to be used with young non-offending adolescents. Control words were derived from the MRC Psycholinguistic Database, and all words were matched for word frequency, word length and word type (i.e., adjectives, verbs). The final version of the newly derived word stimulus set can be seen in Appendix B.

Analysis

Analysis was conducted on all available data, including variables such as age, BACS subscales, tests of executive function, vocabulary, number of errors made on the Stroop tasks, mean Stroop RT's and emotional Stroop bias scores, to examine whether the assumptions of parametric testing were met. Emotional Stroop bias scores were calculated by subtracting the mean RT of neutral or matched words from the mean scores of target words (i.e., words with emotional content). This resulted in five Stroop bias scores to be compared across the groups for the Smith and Waterman (2004) word stimulus set (i.e., positive, negative, color, aggression and sexual) and three Stroop bias scores for the newly derived set of stimuli (i.e., emotional/personality descriptors, sexual actions and physical descriptors). For the new word set, overall differences between the experimental words and matched words were also explored, resulting in an additional emotional Stroop bias score for the experimental words.

To test whether the data was normally distributed the Shapiro-Wilk test was used over the Kolmogorov-Smirnov test because it is more accurate and provides an exact significance (Field, 2005). Homogeneity of variance was tested using the Levene Test.

RT and Stroop bias scores were explored for outliers using histograms and boxplots. When outliers were identified the participant's full set of data was removed from the data set, if the removal from the data set could be justified. Once outliers were removed, if the data still did not meet the assumptions of parametric testing attempts were made to transform the mean RT's and Stroop bias scores using the log and inverse functions (Gress & Laws, 2009) in order to correct problems with the distribution of the data across word categories.

Difficulties with the assumptions of parametric tests are not uncommon when working with RT data. The data often tend to be skewed (Gress & Laws, 2009) and extreme values may be representative of the research question being examined (Miller, 1991). For example, if we expect that sexual abusers will take longer to process deviant sexual information because it is more salient to that particular group, one would not want to treat extreme values in the sexual word category for this group as outliers. Outlines of the violations to parametric testing are provided in Chapters 3, 4 and 5 to highlight the complicated nature of the response patterns using the Stroop task paradigm and differences in response patterns by different groups.

Once the exploration of the data was complete, all mean RT's and emotional Stroop bias scores were entered into separate one-way ANOVA's with participant group as the between-groups variable. Post-hoc analyses were conducted using a Bonferroni correction to control for the Type I error rate. However Gabriel's post-hoc procedure was also conducted because the sample sizes were unequal and this test has greater power (Field, 2005). Effect sizes for the one-way ANOVA's were calculated with the following equation:

$$\omega = \sqrt{\frac{SS_M - (df_M)MS_R}{SS_T + MS_R}}$$

Although ANOVA is considered to be robust to slight violations in the assumptions of parametric testing additional analyses were conducted on data that did not meet the assumptions of parametric testing. The non-parametric equivalent of the one-way ANOVA, the Kruskal-Wallis test, was used to explore whether significant differences were present between the groups. When it was necessary to conduct the Kruskal-Wallis test, Mann-Whitney tests were used as post-hoc procedures with a Bonferroni

correction to ensure that Type I error rates did not inflate past a .05 significance level (Field, 2005). Therefore, a critical value of .025 was used for the post-hoc analyses because two Mann-Whitney tests were run for each post-hoc comparison: sexual abusers compared to offending controls and sexual abusers compared to non-offending controls. Effect sizes for the post-hoc comparisons were calculated using the following equation:

$$r = \frac{z}{\sqrt{N}}$$

Because all of the groups completed the same Stroop tasks, it was possible to run a two-way mixed ANOVA with type of emotional Stroop bias score as the repeated measures variable and participant group as the between groups variable. Interaction effects were examined to determine whether the type of Stroop bias and participant group have a combined effect on the resulting Stroop bias scores. When the added assumption of sphericity was violated (i.e., the variance of the differences between the levels differed) the Greenhouse-Geisser correction was used to assess the significance of the F-statistics (Field, 2005).

The effect sizes for this analysis were conducted using the following equation:

$$r = \sqrt{\frac{F(1, df_R)}{F(1, df_R) + df_R}}$$

In addition to testing for significant differences in Stroop effects between the groups it was of interest to test whether covariates were present that might also have an

influence on the results, and to control for these variables within the analysis. Therefore, analysis of covariance (ANCOVA) was carried out on the dependent variables (i.e., mean RT's and Stroop bias scores) with the participant characteristic variables that have been shown to influence Stroop results (i.e., age, level of executive function, and level of vocabulary) set as covariates. ANCOVA's were also conducted on the BACS subscales of cognitive distortions and emotional congruence with children. The additional assumption of homogeneity of the regression slopes was run for each ANCOVA. Post-hoc comparisons for the ANCOVAs were carried out as suggested by Field (2005) by comparing the main effects with a Bonferroni correction, significance level of .025. It is important to note that although ANCOVAs were run for each dependent variable, when the data violated the assumptions of parametric testing and non-parametric measures were necessary, the results of the ANCOVA's were interpreted with caution because there is no available non-parametric equivalent of ANCOVA.

Procedures

Participants were provided with an information sheet that outlined the procedures of the experiment and were then asked to sign consent forms agreeing to take part in the study. Participants were required to learn colour-mappings of the response buttons prior to the presentation of target words by completing four practice blocks (25 words per block) of neutral word stimuli in order to memorize the colour response mappings on the serial response box. Participants then completed one of the two possible emotional Stroop tasks. All Stroop word stimuli were presented randomly and prior to each word a fixation "x" appeared at the centre of the screen for 500ms. Participants were asked to press the button corresponding to the colour of ink in which the word

was presented and to ignore the semantic meaning of the word. No opportunity to correct mistakes was provided because once a response was provided the next word stimulus was prompted. Following the first emotional Stroop task participants completed the BPVS. The second emotional Stroop task was then completed by participants, followed by the Hayling and Brixton tests of executive function, and finally a shortened version of the BACS questionnaire. The emotional Stroop tasks were counterbalanced in order to reduce the likelihood of order effects. The BACS was always completed at the end of the session due to the sensitive nature of the questions and to avoid priming effects that would be possible if the questionnaire were to be filled out prior to viewing the sexual word content involved in the emotional Stroop tasks.

Data Management and Confidentiality

Participant names and identifiers were not collected for the purpose of this study, therefore information collected was anonymous and participants were managed using a numbering system.

CHAPTER 3

Measuring Deviant Sexual Interest in Adult Samples Using the Emotional Stroop Task

The aim of this study was to examine differences in mean RT's and emotional Stroop bias scores between adult sexual abusers, adult offending controls and adult non-offending controls. Suggestions made in Chapter 1 regarding best methods of practice for conducting an emotional Stroop experiment with offenders were implemented and the utility of the emotional Stroop task with these samples was explored. The chapter presents similarities and differences in response patterns to the Smith and Waterman (2004) stimulus set between the participants in this study and participants from previous work. Additionally, mean RT's and Stroop biases were examined between the groups for the newly-derived word stimuli and the adaptation of the word stimuli is explored to determine whether these new words reflect sexual interests more specific to sexual abusers. The chapter concludes with the implications of the findings and provides suggestions for future research.

Introduction

For decades researchers and clinicians have been using penile plethysmography (PPG) to measure deviant sexual interests. However, there have been questions raised regarding its construct validity and PPG's ability to accurately measure sexual interest, as opposed to simply measuring sexual arousal (Laws, 2003, 2009; O'Donohue & Letourneau, 1992). It is widely accepted that clinicians and researchers tend to measure the penile response (i.e., arousal) because it is a tangible response to measure. However, it has also been noted that the ways in which we conduct PPG assessments puts us at risk of overlooking affective and cognitive components that are also relevant to sexual behaviour (Barlow, 1977; Laws, 2003, 2009; O'Donohue & Letourneau, 1992).

In theory, research has demonstrated that there are differences in the cognitive processes of sexual abusers related to their interests (Abel, Gore, Holland, Camp, Becker, and Rathner, 1989; Beech et al., 2008; Geer, Estupinan and Manguno-Mire, 2000; Nunes, Firestone, & Baldwin, 2007; Price & Hanson, 2007; Smith & Waterman, 2004; Ward, Hudson, Johnston, and Marshall, 1997). It is also argued that one should therefore expect sexual abusers to possess cognitive and affective associations specific to their experiences and interests, and that they would exhibit biases towards salient stimuli that are relevant to these experiences and interests (McNally, 1998; Smith, 2009). It is with these ideas in mind, and questions surrounding the construct validity of PPG assessment, that work on the development of information-processing paradigms to measure sexual interest have originated from.

The emotional Stroop task is a task that is considered to be an information-processing measure of sexual interest. The response latencies experienced by an

individual following the presentation of salient word stimuli is believed to be a function of the processing of the stimuli presented, and a reflection of the strength of representations associated with the words, both cognitively and affectively (Smith, 2009).

Emotional Stroop studies that have been conducted to date examining sexual and aggressive interests of offender populations have shown that sexual offenders display longer response latencies to sexual word stimuli compared to control subjects (Price & Hanson, 2007; Smith & Waterman, 2004). Similarly, violent offenders display a response bias towards aggression-themed word stimuli over control subjects (Smith & Waterman 2003, 2004). These studies confirm the assumptions that sexual abusers would exhibit biases towards salient stimuli that are relevant to their experiences and interests.

The measurement of sexual interests informs us in the treatment of sexual abusers (Beech, Oliver, Fisher & Beckett, 2006; Marshall, Anderson & Fernandez, 1999; Yates, Goguen, Nicholaichuk, Williams & Long, 2000); in risk assessments (Craig, Browne & Beech, 2008; Hanson, Harris, Scott & Helmus, 2007; Hanson & Thornton, 2003; Prentky & Righthand, 2003; Wong, Olver, Nicholaichuk & Gordon, 2004); and as a predictor of sexual recidivism (Hanson & Bussière, 1998). In practice, sexual interests also aid us in differentiating between persistent sexual abusers and lower-risk offenders and non-offenders (Thornton & Laws, 2009).Given the importance attributed to sexual interests in the assessment and treatment of sexual abusers, further development of an information-processing tool that could measure the implicit nature of this factor would be a valuable contribution to theory and practice in work with sexual abusers.

Smith (2009) has noted that the offence-related emotional Stroop tasks that have been tested with offender samples have used word stimuli that are general in nature, or do not test for specific deviance, and are thus unable to differentiate between offender subgroups. Smith (2009) and Price and Hanson (2007) suggest that if the word stimuli were altered to reflect more specific deviant sexual interests then they could be used to assess sexual interests relevant to sexual abusers. Given the failure thus far of the emotional Stroop task to be able to differentiate between different offender groups (i.e., between sexual offenders and violent offenders or between paedophiles and rapists), an amendment to the word stimuli could be an appropriate solution to this dilemma. It could be that the results yielded thus far on the emotional Stroop task with sexual abuser samples are measuring sexual preoccupation or a preference for sexual material rather than specific elements of sexual interest.

This study is an examination of Stroop interference effects experienced by sexual abusers, violent offenders and non-offending controls to offence-related word stimuli. This study makes use of the word stimulus set from Smith and Waterman's (2004) study to examine whether changes in the administration of the Stroop task (i.e., button-press responses instead of voice activated responses) produces similar response patterns across the groups. Additionally, response bias patterns from a new stimulus set derived to reflect sexual interests that are specific to sexual abusers are examined for differences between the groups. Attention is paid to potential participant characteristics that have been known to have an effect on Stroop results. Significant differences in participant age, levels of executive function and level of vocabulary are explored between the groups. Additionally, these variables are examined to verify whether they have an influence on Stroop results.

Study Purpose

Given evidence to support that the emotional Stroop task can be used to assess differences in information-processing between sexual offenders and other groups, the purposes of the study were:

- To compare response patterns to the Smith and Waterman (2004) word stimulus set using a button-press version of the emotional Stroop task between three adult groups: sexual abusers, offending controls, non-offending controls;
- To test a new set of word stimuli that reflect more specific deviant sexual interests and to examine the differences in mean RT's and Stroop bias scores between groups of sexual abusers, offender controls, and non-offending controls;
- To examine whether participant characteristics such as age, executive function and beliefs about children influence Stroop results;
- To examine whether type of Stroop bias (i.e., word category) and participant group interact and have a joint effect on emotional Stroop bias scores.

Method

Participants

Twenty-seven men convicted of sexual offences took part in the study. Of these, there were five exhibitionists, four incest offenders, nine paedophiles, seven rapists and two mixed sexual abusers (had offended against both adult and child victims). An additional 21 men took part in the study that had been convicted of violent offences and 38 non-offending control participants. Eight of the sexual abusers were on

community supervision and recruited through West Midlands Probation, UK; 19 sexual abusers and the 21 violent offenders were recruited through poster recruitment and by research representatives from separate wings at HMP Grendon (Oxfordshire, UK). All of the offending controls had been convicted of either manslaughter or murder. The non-offending controls were recruited through the University of Birmingham research scheme whereby psychology undergraduates, masters and PhD students complete 10 hours of research credits per year to fulfil degree requirements.

Additional Methods

Participants in this study completed two emotional Stroop tasks, the BPVS, Hayling and Brixton tests of executive function and BACS questionnaire. Chapter 2 provides a full description of the apparatus, materials, analysis and procedures used in this study on pages 38-48.

Violations of Parametric Testing

Exploration of the data for participant age, the measures used in the study and the number of errors made during the Stroop tasks indicated that the data were not normally distributed for any of the variables. The assumption of homogeneity of variance was violated for participant age, BACS cognitive distortions, Hayling total score and number of errors made on the Smith and Waterman (2004) Stroop task. Non-parametric analysis of this data using the Kruskal-Wallis test was conducted due to the degree of violations to the assumptions of parametric testing. Post-hoc analyses were conducted using Mann-Whitney tests, and a Bonferroni correction was applied to the data when multiple comparisons were undertaken (described in Chapter 2).

Two participant outliers, with respect to the mean RT's and Stroop bias scores, were identified from the sample of non-offending controls. These participants were removed from the database as they yielded extremely high scores on the BACS subscales and RT's that were two standard deviations above the mean RT's for all of word categories. Therefore, the results from these two participants did not appear to be representative of the non-offending control subgroup. Rather their results seemed to indicate a hypersensitivity to the emotional Stroop task. An additional two participants in the Smith and Waterman stimulus set, and seven participants in the newly-derived stimulus set, were identified as contributing to the skewness of the data. However there was no valid reason for the removal of these participants from the analyses. Data from one non-offending control participant was lost for the newly-derived word stimulus set.

Smith and Waterman Stimulus Set

Further to the removal of the outliers and attempts made at transforming the data (outlined Chapter 2, p. 44), the mean RT data still did not meet assumptions of parametric testing. The data was not normally distributed for the non-offending controls in the neutral word category, $W(36) = .927$, $p < .05$ and the offender controls for the positive words, $W(21) = .866$, $p < .05$. Testing the assumption of homogeneity of variance on mean RT's showed that the assumption was violated for the positive, $F(2, 81) = 7.60$, $p < .05$; negative, $F(2, 81) = 7.10$, $p < .05$; sexual, $F(2, 81) = 7.52$, $p < .05$; and colour words, $F(2, 81) = 3.74$, $p < .05$, based on the mean.

Exploration of the emotional Stroop bias scores also revealed violations of the assumptions of parametric testing. The data was not normally distributed for the offender controls on aggressive Stroop bias scores, $W(21) = .845$, $p < .05$, and sexual

Stroop bias scores, $W(21) = .856$, $p < .05$. Finally, the data for emotional Stroop bias scores violated the assumption of homogeneity of variance on the positive, $F(2, 81) = .33$, $p < .05$, and colour Stroop bias scores, $F(2, 81) = 6.45$, $p < .05$.

Newly-Derived Word Stimuli

Table 1 outlines the instances where the assumption of normally distributed data was violated, and the groups that contributed to the violation of the assumption. Additionally, the assumption of homogeneity of variance was violated for the mean RT data, based on the mean, for all word categories: emotional/personality descriptors (EPD), $F(2, 80) = 9.25$, $p < .05$; matched EPD (MEPD), $F(2, 80) = 4.60$, $p < .05$; sexual actions (SA), $F(2, 80) = 6.97$, $p < .05$; matched SA (MSA), $F(2, 80) = 7.86$, $p < .05$; physical descriptors (PD), $F(2, 80) = 4.17$, $p < .05$; matched PD (MPD), $F(2, 80) = 6.72$, $p < .05$; overall experimental mean RTs (EXP), $F(2, 80) = 8.33$, $p < .05$; and matched EXP (MEXP), $F(2, 80) = 5.03$, $p < .05$. Homogeneity of variance was violated for the EPD Stroop bias scores, $F(2, 80) = 6.10$, $p < .05$, SA Stroop bias scores, $F(2, 80) = 3.50$, $p < .05$, and EXP Stroop, $F(2, 80) = 9.90$, $p < .05$, based on the mean. Attempts made to transform the data (outlined Chapter 2, p. 44) did not result in changes to the violations of the assumptions of parametric testing.

Table 1

Violations to the Assumption of Normally Distributed Data for the Newly-Derived Word Stimulus Set

Word Category	Violations of Assumption	Test Statistic (p-value)
EPD	Offending controls	W(21) = .84 (.003)
MEPD	Sexual abusers	W(27) = .909 (.022)
	Non-offending controls	W(35) = .94 (.039)
SA	Non-offending controls	W(35) = .89 (.002)
MSA	Sexual abusers	W(27) = .90 (.016)
MPD	Offending controls	W(21) = .86 (.007)
MEXP	Sexual abusers	W(27) = .92 (.047)
Stroop EPD	Sexual abusers	W(27) = .86 (.002)
Stroop SA	Offending controls	W(21) = .87 (.011)
	Non-offending controls	W(35) = .83 (.000)
Stroop EXP	Non-offending controls	W(35) = .91 (.005)

Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories. W denotes the test-statistic for the Shapiro-Wilk test of normality.

Results

Significant differences between the groups resulted for participant age, $H(2) = 60.50$, $p < .001$; Hayling test scores, $H(2) = 9.51$, $p<.05$; Brixton test scores, $H(2) = 13.50$, $p < .05$; and number of errors made on the Smith and Waterman (2004) stimulus set, $H(2) = 13.15$, $p < .05$ (see Table 2).

Table 2

Mean Age, Scores on Study Measures and Number of Stroop Errors (SD) by Participant Group

Variable	Adult Non-Offending Controls	Adult Offending Controls	Adult Sexual Abusers
Age	20.47 (2.48)	35.24 (10.06)	43.31 (9.82)
BACS (CD)	3.36 (3.71)	2.80 (3.38)	3.80 (7.21)
BACS (EC)	11.33 (9.28)	7.61 (7.11)	9.52 (11.20)
Hayling	6.33 (.83)	6.10 (.72)	5.58 (1.03)
Brixton	8.69 (1.47)	8.10 (1.48)	6.76 (2.19)
BPVS	146.06 (9.40)	146.33 (10.67)	144.31 (11.58)
Stroop Errors (Smith & Waterman, 2004)	7.83 (4.59)	4.10 (2.30)	5.78 (6.64)
Stroop Errors (Price, 2010)	5.71 (4.55)	3.29 (2.28)	5.04 (4.98)

Post-hoc analysis showed that the sexual abusers were significantly older than the offending control subjects ($U = 115$, $r = .47$), and non-offending controls ($U = 3$, $r = .85$). Sexual abusers scored significantly lower than the non-offending controls on the Hayling test of executive function ($U = 266.5$, $r = .37$), and the Brixton test of executive function ($U = 196.5$, $r = .46$). Despite the medium to large effect sizes yielded for the tests of executive function, it is important to note that the sexual abusers were merely one scaled score lower than the other groups on the Hayling. This difference rendered them a categorisation of "moderate average" and the other groups were categorised as "average". On the Brixton test, the sexual abusers were

categorised as "average" and the other two groups were categorised as "good", just two categories above average. Interestingly, there were no significant differences evident between groups for scores on the subscales of the BPVS, meaning that their level of vocabulary was adequate across all groups. Finally, the sexual abusers made significantly fewer mistakes on the Smith and Waterman (2004) emotional Stroop task than the non-offending controls ($U = 289$, $r = .35$).

Smith and Waterman Stimulus Set

Results from the one-way ANOVAs, with participant group as the between-groups variable, revealed significant differences between the groups for all mean RT's: neutral, $F(2, 81) = 18.12$, $p < .001$; positive, $F(2, 81) = 25.85$, $p < .001$; negative, $F(2, 81) = 27.82$, $p < .001$; colour, $F(2, 81) = 19.15$, $p < .001$; aggression, $F(2, 81) = 20.15$, $p < .001$; and sexual, $F(2, 81) = 26.00$, $p < .001$.

Non-parametric testing, using Kruskal-Wallis tests, was conducted on the neutral, positive, negative, colour and sexual mean RT's due to the violations of the assumptions of parametric testing exhibited in these word categories. Significant differences between the groups were confirmed for all mean RT's run in these analyses: neutral, $H(2) = 25.39$, $p < .001$; positive, $H(2) = 36.19$, $p < .001$; negative, $H(2) = 33.67$, $p < .001$; colour, $H(2) = 26.47$, $p < .001$; and sexual, $H(2) = 33.25$, $p < .001$.

Table 3 displays the mean RT's and emotional Stroop bias scores for each adult group from this stimulus set. Although medians are typically reported when nonparametric testing is conducted, the means have been reported in this case to remain consistent with the reporting in the literature and to display the magnitude of

variability in the data. However, the medians for the RT's and emotional Stroop bias scores between the adult groups for this word set are found in Appendix C.

Multiple comparisons for the mean RT's using a Bonferroni correction show that the sexual abusers took significantly longer to respond to the aggressive word stimuli than the non-offending controls ($p < .05$). Post-hoc analysis on the nonparametric data (i.e., Mann-Whitney tests with a Bonferroni correction) reveal that the sexual abusers took significantly longer to respond to negative word stimuli than offending controls ($U = 171$, $r = .34$) and non-offending controls ($U = 94$, $r = .69$). Additionally, the sexual abusers took significantly longer to respond than non-offending controls to neutral ($U = 149$, $r = .59$), positive ($U = 89$, $r = .69$), colour ($U = 133$, $r = .62$), and sexual ($U = 94$, $r = .69$) word stimuli.

Table 3

Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Smith and Waterman (2004) Word Stimulus Set

Word Category	Adult Group	Mean RT, ms, (SD)	Stroop Bias scores, ms (SD)
Neutral	Adult sexual abusers	765.00 (130.50)	
	Adult offending controls	701.60 (118.66)	
	Adult non-offending controls	593.77 (96.84)	
Colour	Adult sexual abusers	800.68 (149.30)	35.69 (81.09)
	Adult offending controls	724.92 (141.80)	23.31 (62.91)
	Adult non-offending controls	605.78 (92.63)	12.02 (44.67)
Positive	Adult sexual abusers	802.26 (158.19)	37.26 (82.88)
	Adult offending controls	710.16 (136.82)	8.56 (68.80)
	Adult non-offending controls	581.32 (72.41)	-12.45 (51.78)
Negative	Adult sexual abusers	822.67 (162.50)	57.67 (76.99)
	Adult offending controls	709.17 (115.43)	7.57 (67.69)
	Adult non-offending controls	591.69 (84.98)	-2.07 (63.01)
Aggression	Adult sexual abusers	780.05 (126.11)	15.05 (71.37)
	Adult offending controls	720.90 (156.25)	19.29 (78.62)
	Adult non-offending controls	587.31 (97.52)	-6.46 (65.78)
Sexual	Adult sexual abusers	811.41 (154.21)	46.41 (75.58)
	Adult offending controls	720.18 (121.66)	18.57 (72.99)
	Adult non-offending controls	593.69 (84.82)	-.08 (57.29)

Although the groups displayed significant differences in mean RT's, differences in Stroop bias scores demonstrate the word categories that are most salient for each of the groups. One-way ANOVA yielded a significant difference between the groups for negative Stroop bias scores, $F(2, 81) = 6.22$, $p < .05$, $\omega = .48$. Multiple comparisons revealed that the sexual abusers experienced greater Stroop interference than the offending controls and non-offending controls ($p < .05$).

The analysis of the nonparametric data uncovered significant differences between the groups for the positive Stroop, $H(2) = 7.12$, $p < .05$, and sexual Stroop, $H(2) = 7.92$, $p < .05$ bias scores. Figure 1 displays patterns of emotional Stroop interference effects by word category as experienced between the groups. Post-hoc analysis revealed that the sexual abusers displayed significantly larger Stroop bias scores than the non-offending controls for the positive ($U = 304$, $r = .32$) and sexual ($U = 238$, $r = .32$) Stroop bias scores.

Figure 1. *Emotional Stroop Bias Scores: Smith and Waterman Word Stimulus Set*

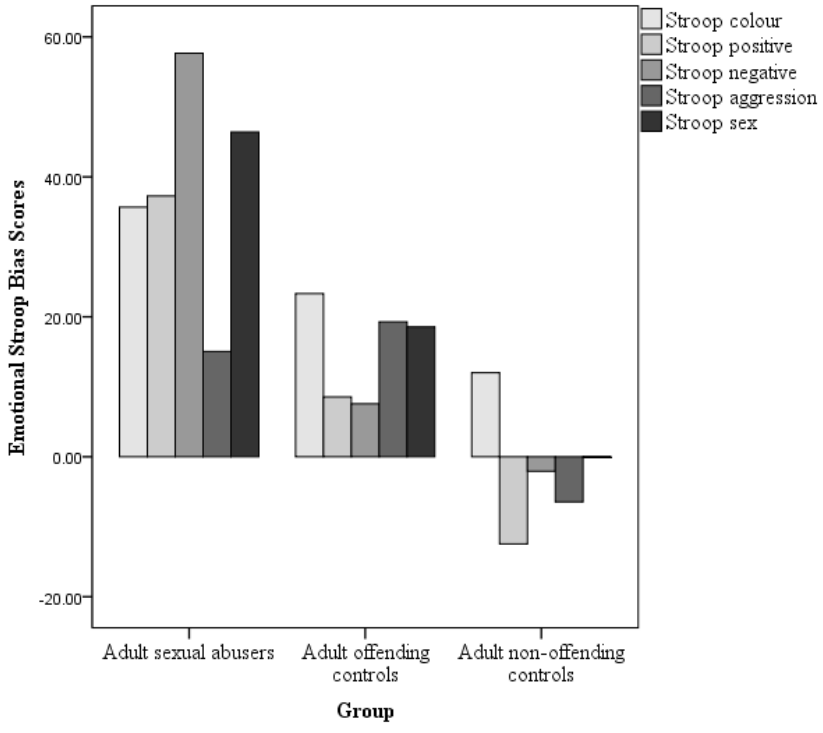
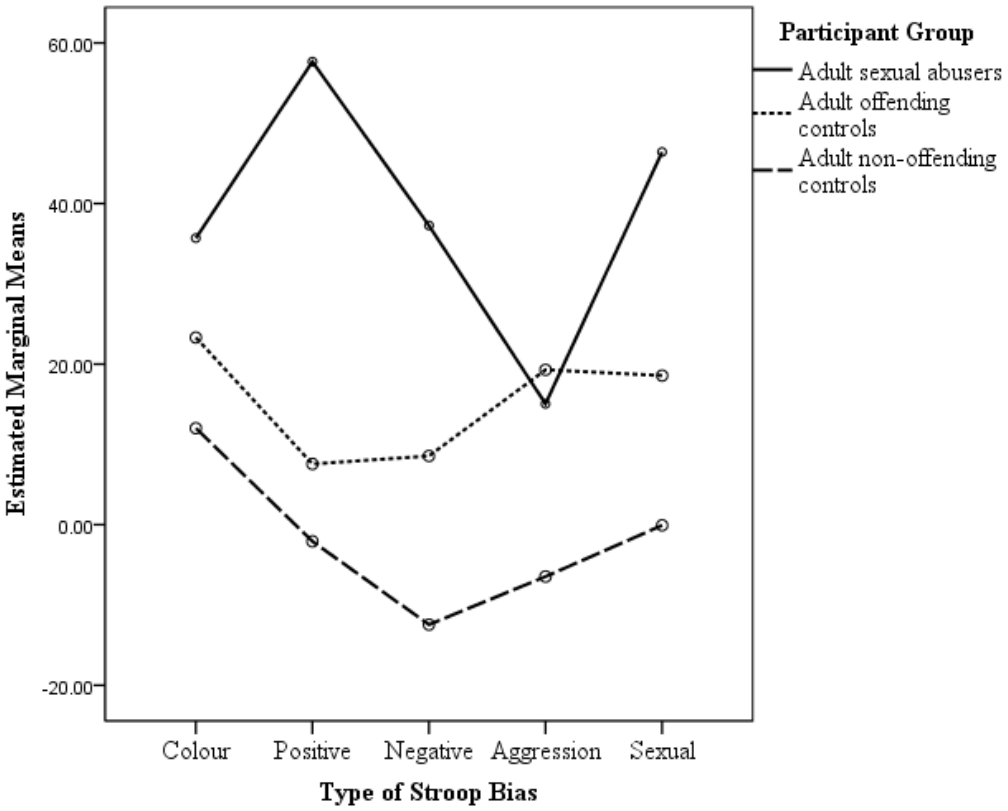


Figure 2 displays the means of the main effects of the two-way mixed ANOVA that was run with type of Stroop bias as the within-subjects variable and participant group as the between-subjects variable. The analysis revealed that the assumption of sphericity was violated therefore the Greenhouse-Geisser correction was used to assess the significance of the F-statistics. The effect of type of Stroop bias was non-significant, $F(3.41, 275.94) = 1.56$, *ns*, as was the effect of the interaction, $F(6.81, 275.94) = 1.44$, *ns*. For the between-groups analysis, the Levene's test for each of the repeated-measures variables indicated that the assumption of homogeneity of variance was violated for the colour Stroop bias and positive Stroop bias scores. Previous attempts to transform the data did not result in the stabilization of the variances between the groups therefore the accuracy of the F-test for participant group is compromised and should be interpreted with caution. However, the test of between-subjects effects yielded a significant main effect of participant group, $F(2, 81) = 4.51$, $p < .05$, indicating that the participant groups elicit different Stroop bias scores regardless of the specific type of Stroop bias. Bonferroni corrected post-hoc comparisons tell us that Stroop bias scores between adult sexual abusers and adult offending controls did not significantly differ ($p = .412$), however Stroop bias scores did significantly differ between the adult sexual abusers and non-offending adult controls ($p < .025$).

Figure 2. *Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman Stimuli*



Analysis of covariance, revealed that the covariate, BACS cognitive distortions, was significantly related to aggression Stroop, $F(1, 77) = 5.05$, $r = .25$. The effect of participant group on aggression Stroop bias scores remained non-significant when BACS (CD) was controlled for. BACS (CD) also displayed a significant relationship with positive Stroop bias scores, $F(1, 77) = 4.01$, $p < .05$, $r = .22$, and the significant effect of participant group on positive Stroop bias scores remained significant after controlling for the effect of BACS (CD), $F(2, 81) = 3.35$, $p < .05$. Finally, the results from the analysis of covariance from this word stimulus set did not yield any significant relationships between the Stroop bias effects and the covariates age, BACS (EC), Hayling and Brixton tests of executive function and BPVS raw scores.

Newly-Derived Stroop Stimulus Set

Non-parametric testing using the Kruskal-Wallis test was conducted on the data from this stimulus set due to the degree of violations made to the assumptions of parametric testing. Analysis of the mean RT's yielded significant differences between the groups for the following word categories: EPD, $H(2) = 31.61$, $p < .05$; MEPD, $H(2) = 24.39$, $p < .05$; SA, $H(2) = 24.17$, $p < .05$; MSA, $H(2) = 25.47$, $p < .05$; PD, $H(2) = 28.45$, $p < .05$; MPD, $H(2) = 31.45$, $p < .05$, EXP, $H(2) = 30.33$, $p < .05$; and MEXP, $H(2) = 29.49$, $p < .05$. Table 4 provides the mean RT's and Stroop bias scores per group (see Appendix D for medians).

Post-hoc analysis revealed that the sexual abusers had slower mean RT's on the EPD and MPD word categories than the offending controls ($U = 172$, $r = .33$; $U = 174$, $r = .33$) and non-offending controls $U = 109$, $r = .66$; $U = 149$, $r = .58$). Additionally, sexual abusers had significantly slower mean RT's than the non-offending controls for the SA ($U = 156$, $r = .57$); MSA, ($U = 152$, $r = .58$); PD ($U = 119.5$, $r = .64$); EXP ($U = 117$, $r = .64$); and MEXP ($U = 114$, $r = .64$) mean RT's.

Table 4

Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Newly-Derived Word Stimulus Set

Word Category	Adult Group	Mean RT, ms, (SD)	Stroop Bias Score, ms (SD)
EPD	Adult sexual abusers	798.10 (197.76)	30.55 (90.05)
	Adult offending controls	683.10 (111.15)	13.83 (48.86)
	Adult non-offending controls	568.36 (84.06)	-14.07 (35.55)
MEPD	Adult sexual abusers	767.55 (164.87)	
	Adult offending controls	669.28 (104.47)	
	Adult non-offending controls	582.43 (92.00)	
SA	Adult sexual abusers	793.81 (186.35)	18.21 (111.55)
	Adult offending controls	685.91 (104.54)	-8.56 (83.89)
	Adult non-offending controls	596.96 (121.72)	17.34 (72.73)
MSA	Adult sexual abusers	775.60 (174.36)	
	Adult offending controls	694.47 (91.41)	
	Adult non-offending controls	579.62 (97.73)	
PD	Adult sexual abusers	782.39 (167.63)	-10.77 (102.32)
	Adult offending controls	704.86 (119.73)	21.02 (73.94)
	Adult non-offending controls	582.61 (98.38)	13.13 (59.88)
MPD	Adult sexual abusers	793.17 (182.51)	
	Adult offending controls	683.84 (108.02)	
	Adult non-offending controls	569.48 (86.50)	

Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories.

Figure 3 demonstrates the patterns in emotional Stroop bias scores for this word stimulus set between the groups. Analysis of the emotional Stroop bias scores revealed a significant difference between the groups for the EPD Stroop bias score, $H(2) = 7.60$, $p < .05$, with the sexual abusers exhibiting significantly larger EPD Stroop effects than the non-offending controls ($U = 310$, $r = .29$). There were no other significant effects for the emotional Stroop bias scores.

Figure 3. *Emotional Stroop Bias Scores: Newly-Derived Word Stimulus Set*

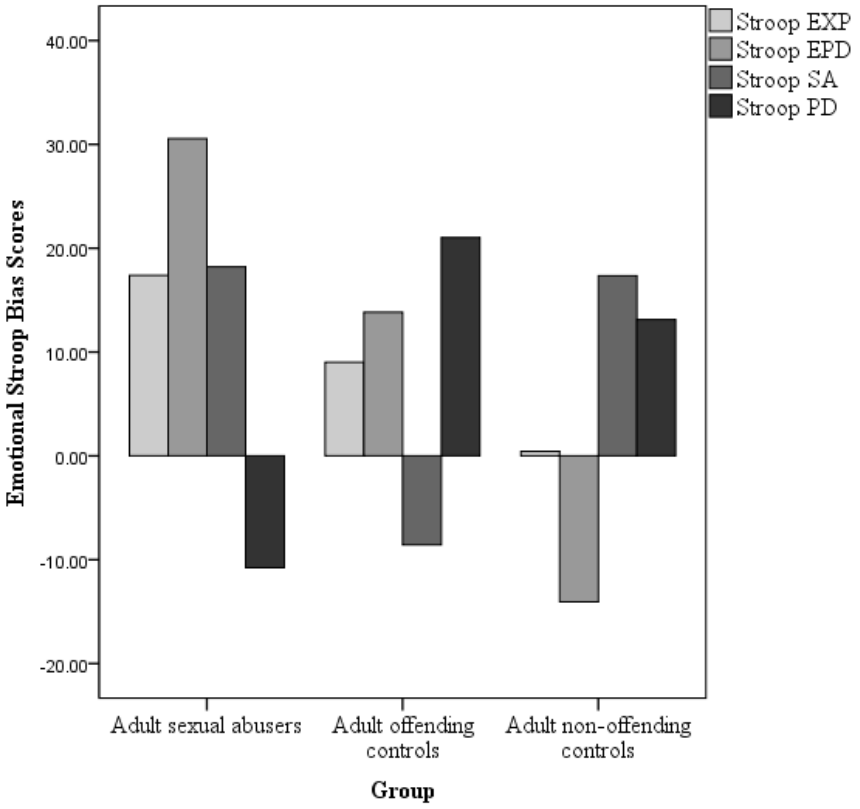
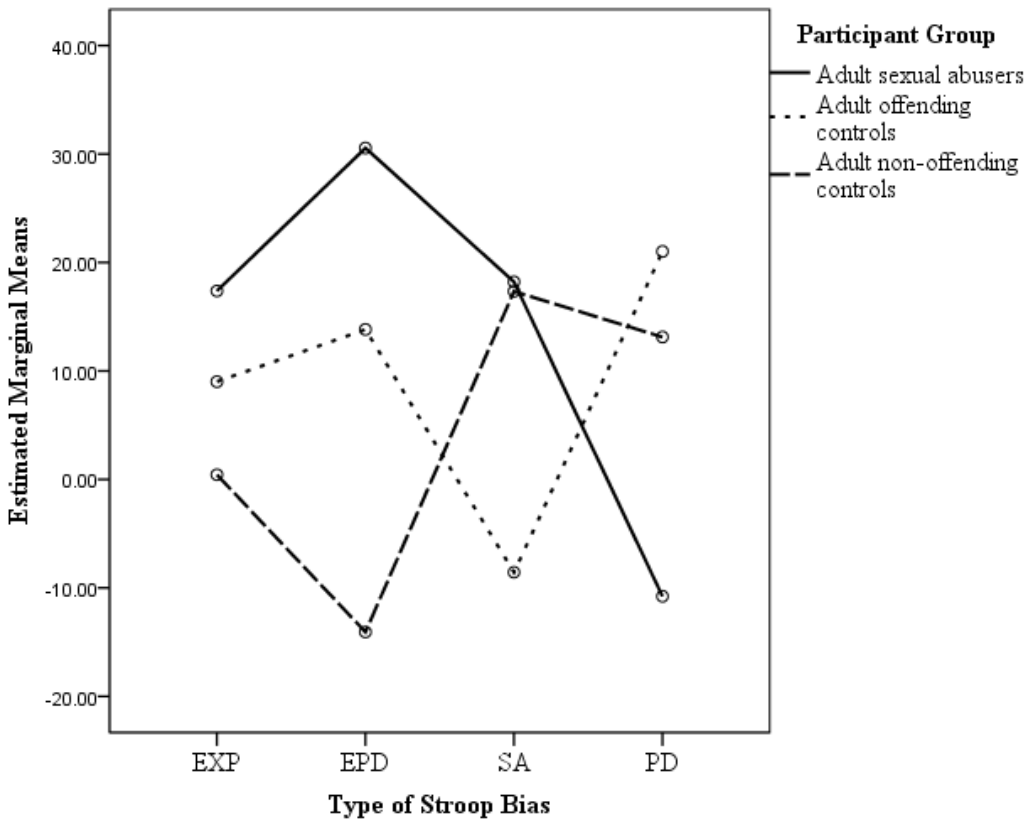


Figure 4 displays the means of the main effects of the two-way mixed ANOVA that was run with type of Stroop bias (i.e., EXP, EPD, SA and PD) as the within-subjects variable and participant group as the between-subjects variable. The assumption of sphericity was again violated therefore the Greenhouse-Geisser correction of the

degrees of freedom was used to assess the significance of the F-statistics. The F-statistic for the effect of type of Stroop bias was less than 1, $F(1.93, 154.16) < 1$, ns, which tells us that there was more error than variance created by the experiment (Field, 2005). Had sphericity been assumed the Type of Stroop Bias x Participant Group interaction would have reached significance, however, with the Greenhouse-Geisser correction this significance was lost, $F(3.85, 154.16) = .50$, ns. The Levene's test for each of the repeated-measures variables indicated that the assumption of homogeneity of variance was violated for the EXP Stroop, EPD Stroop and SA Stroop. The between-subjects main effect of participant group was not significant and was also less than 1, $F(2, 80) < 1$, ns.

Figure 4. *Means for the Main Effects of Type of Stroop Bias and Participant Group: Newly-Derived Word Stimuli*



Analysis of covariance revealed that the covariate, age, was significantly related to the EXP Stroop bias score, $F(1, 77) = 4.193$, $p < .05$, $r = .22$, however the effect of participant group on EXP Stroop bias score remained non-significant when the effect of age was controlled for. Finally, analysis of covariance showed that the covariates, BACS (CD), BACS (EC), Hayling and Brixton tests of executive function and BPVS raw scores) were not significantly related to the Stroop bias effects from the newly-derived word stimulus set.

Discussion

Overview of Results

This study examined differences in Stroop response patterns to two stimulus sets containing sexual, aggressive and sexual interest-themed words. Generally, the sexual abusers took significantly longer to respond to the task than the two comparison groups. Additionally, significant differences for the positive Stroop, sexual Stroop and Stroop bias for emotional/personality descriptors of potential victims were evident between sexual abusers and non-offending control participants.

Discussion of Problems with the Data

The disorganised nature of the data makes it difficult to interpret the results of the Stroop task. Details regarding the violations to the assumptions of parametric testing were provided to display the unsystematic features of RT data, and to aid in the interpretation of the analyses. However it is not unusual to have skewed data, extreme values or numerous violations to the assumptions of parametric testing when dealing with RT data (Gress & Laws, 2009; Miller, 1991). These difficulties with the data are

expected because the results are dependent on the salience of the stimuli to different groups of participants. Additionally, extreme values are often anticipated due to the research question (Gress & Laws, 2009).

Discussion of Differences in Age and Scale Scores

Regardless of the difficulties in the data, significant differences were evident between the groups on age and level of executive functioning. However, when the participant characteristics were taken into account in the analysis, they did not appear to have a significant effect on the outcome of Stroop results. This finding is important because it demonstrates that although significant differences in participant characteristics were evident between the groups, the differences did not appear to be contributing to the significant differences in Stroop interference evident between the groups.

The results of the subscales from the self-report BACS questionnaire did not reflect the anticipated results. One would expect that the sexual abusers would display the highest scores on the scale. Instead, the non-offending control subjects scored similarly to the groups on cognitive distortions and highest on emotional congruence with children. Oddly, the covariate BACS cognitive distortions displayed a relationship with aggression and positive Stroop bias scores. The results of the ANCOVAs should be interpreted with caution due to the nonparametric nature of the data. Moreover, when the influence of BACS (CD) was controlled for, the significant results from the between-groups analysis for Stroop bias scores remained the same despite the influence of the covariate. Had the results of these measures worked as they were intended, the relationships of the subscales to Stroop bias effects would have been considered more authentic.

Discussion of Stroop Findings

The emotional Stroop results from this study are consistent with previous findings from emotional Stroop task studies using offender samples. The sexual abusers in this study took significantly longer to colour-name all word stimuli than the non-offending control subjects, a finding that has also been demonstrated by sexual abuser samples in Price and Hanson (2007) and Smith and Waterman's studies. Patterns in Stroop bias effects were also similar to previous findings in that the sexual abusers from this study yielded significantly larger sexual Stroop bias scores than non-offending controls; and the offender controls (violent offenders) experienced the greatest Stroop interference for aggression-themed word stimuli than the other groups however this finding was not significant. Therefore, the offender groups continue to display interference effects towards stimuli that are relevant to their offending behaviour using the Smith and Waterman (2004) word stimuli, despite the change in Stroop methodology (i.e., button-press recording of responses versus voice-activated response recordings).

Noteworthy is that the Smith and Waterman word stimuli were still not able to differentiate between offender types for the offence-related word stimuli. The results from the two-way ANOVA confirmed that overall that significant differences were evident between the sexual abusers and non-offending controls on Stroop interference. No interaction effect was evident for the Smith and Waterman stimulus set between type of Stroop bias and participant group, suggesting that the groups are responding in a similar way across the different word categories. This finding is not surprising given the concerns regarding the nature of the word stimuli, and the need to develop word stimuli that are more specific to the interests of sexual abusers.

Interestingly, the control subjects in this study generated negative Stroop bias scores for the emotional word categories (i.e., positive, negative, aggression, and sexual), meaning that they took longer to respond to neutral word stimuli than words with emotional content. This result would indicate that the target words from the Smith and Waterman (2004) stimulus set did not produce interference effects for the non-offending control subjects in this study. Importantly, the non-offending controls did experience Stroop interference for the incongruent colour-word stimuli, confirming that they respond in a manner that is consistent with responses we would expect based on literature from the traditional colour-word Stroop task (Stroop, 1935).

What is notable about the patterns of Stroop interference between the groups is that both offender groups experienced greater Stroop interference effects across each of the word categories compared to control subjects. This finding, in combination with the negative Stroop bias scores elicited by the non-offending controls, would suggest that there are clear differences in the way that offender groups process affective information compared to non-offenders.

Interestingly, the sexual abusers in this study yielded greater Stroop interference for negative word stimuli than sexual word stimuli. Mood questionnaires were not administered to participants in this study, therefore one could only propose that their bias towards negative word stimuli is due to the negative stigma often attributed to their offending behaviours, or to mood factors elicited by their surrounding (i.e., prison setting).

For the newly-derived word stimulus set, the sexual abusers were again significantly slower to respond in mean RT's to all word categories than non-offending controls. Although there was only one significant difference between sexual abusers and non-offending controls on Stroop bias scores for this word stimulus set,

the patterns in Stroop bias effects between participants across the word categories are interesting.

For example, the sexual abusers displayed a negative Stroop bias for the physical descriptor word stimuli, while the other two groups experienced interference effects for these types of words. Based on the results from the first Stroop task where sexual abusers displayed interference effects for all of the word categories containing emotional content, one could infer that physical descriptors are not salient or central in the idea of sexual interest for sexual abusers.

Furthermore, the sexual abusers experienced the greatest amount of Stroop interference for the emotional/personality descriptor word category in this stimulus set. In comparison, these word stimuli did not have as much of an effect for offending controls, and did not have an effect at all for non-offending controls. If we reflect on the question that these word stimuli were derived from (i.e., How would you describe who would be most at risk from you?), it could be that the emotional/personality descriptors of potential victims are the more likely elements of sexual interest for sexual abusers.

Finally, the offending control participants displayed negative bias scores for the sexual action words, while the sexual abusers and non-offending controls responded similarly to these words. Of note is the difference in the nature of the sexual action words in the new stimulus set compared to the nature of the sexual words from the Smith and Waterman (2004) words. For example, the new words contain less violent connotations, and are relatively common sexual associations. Therefore, it is interesting that without the aggressive undertones in the word category, the violent offenders are not responding to these word stimuli.

Limitations and Directions for Future Research

Although the methodologies used in this study adhered to suggestions made to control for participant characteristics and experimental variables, this study does suffer some limitations. First, the heterogeneity of the sexual abuser group could have had an impact on the significance of the findings. Not only did the sexual abuser sample vary in offence type, but the length of time that had elapsed since having committed their crimes and the severity of their crimes differed. Some of the sample was on community supervision, while others had been in prison for more than a decade. Despite efforts made to recruit a more homogeneous sample, difficulties with recruitment resulted in a more mixed sample than intended. Additional analysis on the subgroups of this sample would not have elicited enough power to be able to generalise the results for these subgroups, therefore this analysis was not conducted. However, this limitation leaves room for future research to be conducted that concentrates on expanding upon the group sizes in order to obtain more confident results that we would be able to generalize to more specific subgroups of sexual abusers.

The heterogeneity of the group also raises questions about whether risk level and time spent in prison would have an effect on Stroop results. Information on risk level and amount of time spent in prison, or amount of time that had elapsed since the occurrence of the offence was collected when possible. However, since criminal records and risk level information were not accessible to the researcher, analysis of these variables was not possible. Future research efforts should examine whether emotional Stroop biases differ between groups when these factors are considered in the analysis.

Finally, although age was not considered to be a covariate, a non-offending control group of a similar age group to the offenders would have been desired, and perhaps a more appropriate comparison group for the study. Additionally, had time allowed, mood questionnaires could have been advantageous to use.

CHAPTER 4

Measuring Deviant Sexual Interest in Adolescent Samples Using the Emotional Stroop Task

The results presented in the previous chapter summarised the response patterns to emotional Stroop tasks for adult samples. The results provide further evidence that differences in information-processing exist between adult sexual abusers and other adult groups towards emotional stimuli. The Smith and Waterman (2004) stimuli appear to elicit response bias for offender groups on a general level. The different word categories of the newly-derived word stimuli appear to be generating different patterns in response bias across the varying groups more specific to individual interests. It is of note that the results from the adult study in Chapter 3 were not statistically significant for the most part, however, the response patterns between the groups were distinct across the separate word categories for the newly-derived stimuli. These results would suggest that we may be one step closer to deriving words that would be reflective of sexual interests specific to sexual abuser populations.

The aim of this chapter then was to explore the utility of the emotional Stroop task with adolescent samples. The current study reports the results of three adolescent samples to the same emotional Stroop word sets that were used in the adult study. Similarities and differences in response patterns to the Smith and Waterman (2004) stimulus set between the groups are explored. Additionally, this study attempted to determine whether the new word stimuli reflect sexual interests that are specific to adolescent sexual abusers. The value in this is to determine whether the word stimuli from the emotional Stroop task are robust enough to be applied to this subgroup.

Introduction

The prevalence rates of young people aged less than 21 who commit sexually abusive offences ranges between 20% and 50% of all sexual offences committed (Barbaree & Marshall, 2006; Becker, Kaplan, Cunningham-Rathner, & Kavoussi, 1986; Erooga & Masson, 2006; Oliver, 2007). However, adolescent sexual abusers are a heterogeneous group of offenders (Andrade, Vincent & Saleh, 2006; Beckett, 1999, 2006), but they often receive generic treatment services that are modelled on adult sex offender treatment programmes. This tends to result in adolescent sexual abusers receiving treatment that overlooks the specific treatment needs of this group and their heterogeneous nature (Freeman-Longo, Bird, Stevenson, & Fiske, 1995; Gerhold, Browne, & Beckett, 2007; Hunter, Figuerdo, Malamuth & Becker, 2003).

Essential to our understanding of motivations for sexually abusive behaviour is the information gained from the assessment of sexual interests and arousal (Barbaree, 1990). A focus on assessing and treating deviant sexual arousal and interests is typically a component of the treatment services offered to adolescent sexual abusers (Freeman-Longo et al., 1995), and attempts at improving and developing more accurate tools to measure the sexual interests of adolescents are an ongoing task (Abel et al., 2004). Measures that have been used to assess sexual interest in adults such as risk assessments, self-report measures and penile plethysmography (PPG) have also been employed with adolescent samples (Abel et al., 2004; Grant, 2006).

However, the usefulness of risk assessments with adolescent sexual abusers has been brought into question because of low base rates in reoffending, a lack of

adequate information to base decisions on, and a requirement for further empirical evidence to validate existing risk assessment measures (Beckett, 2006; Grant, 2006).

There is also little evidence to support the use of PPG with adolescent samples. PPG's test-retest reliability has been examined by Becker, Hunter, Goodwin, Kaplan and Martinez (1992) using audio-taped stimuli on 20 adolescent sexual abusers. The authors found significant correlations ranging between .46 and .83 in strength for 15 out of 19 audio-taped vignettes. The correlations were strongest when the adolescents had previously engaged in the behaviour similar to that represented in the vignette (Becker et al., 1992). Seto, Lalumière and Blanchard (2000) examined PPG's discriminate validity in being able to identify sexual interest in children for adolescent sexual offenders. It was found that offenders who had only targeted female victims did not significantly differ from the young adult rapist and non-offender comparison groups. However, adolescent perpetrators that had offended against both male and female victims responded more to the stimuli depicting children than to adults (Seto et al., 2000). Issues surrounding the use of phallometric assessment with adolescents include concerns for its use with this age group, unsuitable comparison groups (Seto et al., 2000), and a lack of empirical studies validating the measure for use with adolescents, and matters regarding informed consent (Grant, 2009).

Additional efforts to measure sexual interests of adolescent sexual abusers have been conducted by Abel and colleagues (2004) incorporating visual reaction (VRT) time as measured by the Abel Assessment for Sexual Interest™ (AASI). The AASI combines self-reported ratings of sexual interest, VRT measures to various sexual stimuli and a questionnaire to measure sexual interest of sexual abusers. Research has found evidence to support the criterion validity of the AASI and its ability to resist false accounts of offending behaviours in adult samples (Abel et al.,

2001). In an attempt to validate the measure to assess sexual interest with adolescent sexual abusers Abel et al., (2004) tested 1,170 adolescents who had offended against a child and 534 adolescents that had committed a variety of other types of sexual offences (i.e., voyeurism, rape, exhibitionism). The AASI was able to differentiate between adolescent molesters and non-molesters and the study found evidence for the validity of the AASI in being able to measure sexual interest with adolescent sexual abusers (Abel et al., 2004).

Given the concerns and ethical issues involved in existing assessment measures used with adolescent sexual abusers (Abel et al., 2004; Beckett, 2006; Grant, 2006), it was of interest to examine whether the emotional Stroop task was an accurate measure of sexual interest with this subgroup. The emotional Stroop task has been extensively tested in adult clinical samples, however, the research using the emotional Stroop task with adult forensic samples and adolescent forensic samples is limited (Moradi et al., 1999).

In the general literature on the colour-word Stroop task, adolescent samples have been tested to examine the effects that ADHD, juvenile delinquency and conduct disorder have on the individual's ability to successfully complete the Stroop task, and to measure level of cognitive deficits experienced in these samples. Evidence is mixed regarding the ability of adolescent offender samples to successfully complete the traditional colour-word Stroop task. For example, Herba, Tranah and Rubia (2010) tested both male and female groups of individuals with conduct disorder, and controls without conduct disorder, on three domains of inhibition including the Stroop task: (1) motor response inhibition as measured by a Stop task; (2) verbal inhibition as measured by the Hayling test of executive function; and (3) cognitive interference inhibition as measured by the Stroop task. The study found that individuals with

conduct disorder were impaired on motor inhibitory control but were not impaired in cognitive and verbal inhibitory control. While, Caroll et al. (2006) found that their offender groups did not display deficits in inhibitory control as measured by the colour-word Stroop task in early-onset juvenile delinquents, late-onset juvenile delinquents and non-offending controls. Instead, the non-offending controls displayed the most Stroop interference on the task. Caroll et al. (2006) then concluded that the adolescent offenders do not experience deficits in inhibition control as measured by the Stroop task. Contrary to these findings, Moffitt (1990) has linked antisocial behaviours and impulsivity to poor performance on the colour-word Stroop task. It is important to note that Caroll et al. (2006) did not test their groups on reading ability or level of vocabulary. Therefore, it is possible that the offender groups were not able to experience Stroop interference if they were lacking basic reading skills (Mutter, Naylor, & Patterson, 2005; Schiller, 1966). Additionally, the study used a paper and pencil version of the Stroop task (Golden, 1978) which is a less reliable version of the task than computerised versions (Salo, Henik, & Roberston, 2001).

In research using the emotional Stroop task Moradi et al., (1999) have examined how children and adolescents with PTSD process emotional information for trauma-related stimuli. The study tested 23 young people aged 9-17 who met the criteria for PTSD according to DSM-IV (APA, 1994) and 23 healthy control subjects using an emotional Stroop task with five word categories: happy, neutral, depression-related, general threat-related and trauma-related. Moradi et al. (1999) found that those diagnosed as having PTSD were slower overall on colour-naming and displayed attentional bias to trauma-related word stimuli when compared to healthy controls independent of age. The authors suggest that the study provides support for the idea that young children and adolescents process emotional information in the same

manner as adults do. This suggestion would then render the results from the adult literature using the emotional Stroop task relevant and comparable to how we should expect adolescents to respond to the task.

Gallagher-Duffy et al. (2009) tested whether a fire-related emotional Stroop (pictorial) task could measure interference in information-processing bias for fire-specific stimuli in three adolescent groups: (1) arsonists; (2) offending non-arsonist controls; and (3) non-offending control subjects. Adolescents that exhibited fire setting behaviours displayed the greatest attentional bias for fire-themed stimuli. Additionally, the Stroop biases were negatively correlated with self-reported fire interest, and positively correlated with fire setting frequency. The authors suggest that the fire-Stroop could supplement self-report measures of interest for fire setting behaviour. It should be noted that the task was not able to significantly differentiate between the offender groups. The authors, however, attribute this finding to a broader categorisation of antisocial behaviour between the two offender groups.

Study Purpose

Given that this task has elicited some results from adult samples indicating biases towards sexual interest information, it was of interest to test whether similar results could be reproduced with adolescent samples. The validation of an information-processing tool to measure sexual interest could be of great value with adolescent samples when considering some of the issues surrounding existing measures used to assess sexual interest with this group.

Therefore, the purposes of the study were:

- To examine the response patterns of adolescent sexual abusers, adolescent offending controls and adolescent non-offending controls to the Smith and Waterman (2004) word stimulus set;
- To test a new set of word stimuli that reflect more specific deviant sexual interests and to examine the differences in mean RT's and Stroop bias scores between the sexual abusers, offender controls, and non-offending controls;
- To examine whether participant characteristics such as age, executive function and beliefs about children influence Stroop results with this sample;
- To examine whether type of Stroop bias (i.e., word category) and participant group interact and have a joint effect on emotional Stroop bias scores;
- To contribute to research regarding the measurement of sexual interests of adolescent sexual abusers.

Method

Participants

Three groups of adolescent participants took part in this study: (1) adolescent sexual abusers (n=24); (2) adolescent violent offenders (n=21); and (3) non-offending adolescent controls (n=21). Twenty-two of the sexual abusers were recruited from G-MAP services (Sale, Cheshire), an independent organisation that works with young people who display inappropriate sexual behaviours. All adolescent sexual abusers from this site were residing in secure placement homes. The adolescent offender comparison group and two sexual abusers were recruited through a Youth Offending Team (YOT) organisation in the West Midlands. The non-offending adolescent

controls were recruited from a school in the West Midlands and through recruitment of individuals attending Open Days at the University of Birmingham.

From the sample of sexual abusers 12 were extrafamilial offenders, 8 were intrafamilial offenders, and 4 were mixed (i.e., had offended both within and outside of the family environment). Five of the sexual abusers targeted male victims only, 11 targeted female victims only, and eight had offended against both male and female victims. The adolescent offending controls had committed a variety of offences including property offences (N = 1), theft (N = 5) and violent offences (N = 15). None of the adolescent non-offending controls had a criminal history.

Additional Measures

Participants in this study completed two emotional Stroop tasks, the BPVS, Hayling and Brixton tests of executive function and the BACS questionnaire. Chapter 2 provides a full description of the apparatus, materials, analysis and procedures used in this study on pages 38-48.

Violations of Parametric Testing

Analysis on age, the measures used in the study and the number of errors made during the Stroop tasks show that the data was not normally distributed for any of the variables, and the assumption of homogeneity of variance was violated for participant age, BACS subscales and Hayling test scores.

One outlier was identified in the data set resulting in the removal of this participant from the adolescent offending control group. The participant that was removed yielded the lowest raw score on the BPVS, high scores on the BACS subscales, and RT's of two standard deviations below the mean RT's for all of word

categories indicating an unsuccessful attempt at completing the emotional Stroop task. An additional 12 participants were identified as contributing to the skewness of the data for some of the word category means however there was no valid reason for removing any of these participants from the data. Data from one non-offending control subject on the Smith and Waterman Stroop task was lost, resulting in a sample size of 20 for this group.

Smith and Waterman Stimulus Set

Table 1 lists the violations to the assumption of normally distributed data for the mean RT's and emotional Stroop bias scores based on the Shapiro-Wilk test. The assumption of homogeneity of variance (based on the mean) was violated for the mean RT's of sexual words, $F(2, 61) = 3.33$, $p < .05$, and sexual Stroop bias scores, $F(2, 61) = 4.10$, $p < .05$. Attempts made to transform the data (outlined in Chapter 2, p.44) did not correct the data enough to be normally distributed in the sexual and colour word categories despite the assumption of homogeneity of variance being maintained with the transformed data.

Table 1

Violations to the Assumption of Normally Distributed Data for the Smith and Waterman Stimulus Set

Word Category	Violation to Assumption	Test Statistic (p-value)
Negative	Offending controls	W(20) = .86 (.027)
Sexual	Offending controls	W(20) = .89 (.027)
	Non-offending controls	W(20) = .83 (.003)
Colour	Sexual Abusers	W(24) = .89 (.014)
	Offending controls	W(20) = .87 (.014)
Stroop Colour	Sexual abusers	W(24) = .90 (.018)

Newly-Derived Word Stimulus Set

Table 2 lists the violations to the assumption of normally distributed data for the mean RT's and emotional Stroop bias scores based on the Shapiro-Wilk test for the set of newly-derived word stimuli. Homogeneity of variance was violated for the mean RT's (based on the mean) in the MPD, $F(2, 62) = 6.47$, $p < .05$; and MEXP, $F(2, 62) = 4.78$, $p < .05$ word categories. The assumption of homogeneity of variance was not violated for the emotional Stroop bias scores.

Table 2

Violations to the Assumption of Normally Distributed Data for the Smith and Waterman Stimulus Set

Word Category	Violation to Assumption	Test Statistic (p-value)
EPD	Sexual abusers	W(24) = .90 (.024)
	Offending controls	W(20) = .84 (.004)
MEPD	Offending controls	W(20) = .81 (.001)
SA	Sexual abusers	W(24) = .95 (.049)
	Offending controls	W(20) = .89 (.025)
PD	Offending controls	W(20) = .88 (.016)
MPD	Sexual abusers	W(24) = .89 (.012)
	Offending controls	W(20) = .88 (.017)
EXP	Offending controls	W(20) = .86 (.007)
MEXP	Offending controls	W(20) = .78 (.000)
Stroop EPD	Offending controls	W(20) = 90 (.035)
	Non-offending controls	W(21) = .95 (.008)

Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories.

Attempts were again made to transform the data however these transformations did not correct the data enough to meet the assumptions of parametric testing. Therefore, it was necessary to conduct more rigorous non-parametric analyses on the mean RT's and emotional Stroop bias scores that did not meet the assumptions of parametric testing: EPD, MEPD, SA, PD, MPD, EXP, MEXP, and EPD Stroop.

Results

Significant differences were evident between the groups on participant age, $H(2) = 9.54$, $p < .05$; BACS (CD), $H(2) = 8.0$, $p < .05$; BACS (EC), $H(2) = 24.12$, $p < .001$; and BPVS raw scores, $H(2) = 19.95$, $p < .001$ (see Table 3).

Table 3

Mean Age, Scores on Measures and Number of Stroop Errors (SD) per Adolescent Group

Variable	Non-Offending Adolescent Controls	Adolescent Offending Controls	Adolescent Sexual Abusers
Age	17.05 (0.76)	16.75 (0.91)	15.96 (1.27)
BACS (CD)	6.26 (6.05)	7.90 (6.91)	14.08 (10.64)
BACS (EC)	11.42 (7.17)	3.0 (4.58)	19.04 (13.49)
Hayling	6.10 (0.45)	5.8 (0.77)	5.58 (1.35)
Brixton	8.0 (1.49)	6.80 (1.64)	6.92 (1.82)
BPVS	131.35 (12.59)	112.65 (12.09)	113.38 (13.58)
Stroop Errors (Smith & Waterman, 2004)	7.80 (5.65)	10.90 (10.24)	7.08 (7.45)
Stroop Errors (Price, 2010)	4.71 (3.62)	7.35 (6.85)	6.04 (3.76)

Post-hoc analyses using Mann-Whitney tests and a Bonferroni correction displayed medium to large effect sizes showing that adolescent sexual abusers are significantly younger in age to the non-offending adolescent controls ($U = 122.50$, $r = .44$). For the BACS subscales, the adolescent sexual abusers score significantly higher than the

non-offending adolescent controls on cognitive distortions (U = 120.50, r = .40), and significantly higher than the adolescent offending controls on emotional congruence with children (U = 51.00, r = .68). Finally, the sexual abusers differ significantly from the non-offending adolescent controls (U = 19.95, r = .57) displaying significantly lower BPVS raw scores.

Smith and Waterman Stimulus Set

Mean RT's and emotional Stroop bias scores for the Smith and Waterman (2004) stimulus set are displayed in Table 4 (medians displayed in Appendix E). The one-way ANOVAs that were conducted display significant differences between the groups on the mean RT's for the neutral, $F(2, 61) = 10.88$, $p < .001$, $\omega = .49$; positive, $F(2, 61) = 9.10$, $p < .001$, $\omega = .45$; and aggression, $F(2, 61) = 12.91$, $p < .001$, $\omega = .51$, word stimuli. Post-hoc analyses revealed that the adolescent sexual abusers were significantly slower to colour-name the neutral, positive, and aggression word stimuli than both adolescent offending controls ($p < .05$) and adolescent non-offending controls ($p < .001$).

Non-parametric analyses displayed that the adolescent sexual abusers were also significantly slower to colour-name negative word stimuli, $H(2) = 12.17$, $p < .05$, when compared to adolescent offending controls, (U = 138.00, r = .36), and non-offending adolescent controls (U = 100.00, r = .50). Adolescent sexual abusers were significantly slower to colour-name sexual word stimuli, $H(2) = 19.35$, $p < .05$, than the adolescent offending controls (U = 115.00, r = .44), and adolescent non-offending controls (U = 57.00, r = .65). Finally, there was a significant difference in mean colour RT's, $H(2) = 12.37$, $p < .05$, where adolescent sexual abusers took significantly longer to colour-name than the non-offending controls (U = 93.00, r = .52).

Table 4

Mean RT and emotional Stroop Bias Scores in Milliseconds with (SD) for Smith and Waterman (2004) Word Stimulus Set

Word Category	Adolescent Group	Mean RT, ms, (SD)	Stroop Bias scores, ms (SD)
Neutral	Adolescent sexual abusers	810.15 (119.96)	
	Adolescent offending controls	702.06 (115.74)	
	Adolescent non-offending controls	664.27 (83.51)	
Colour	Adolescent sexual abusers	821.93 (167.52)	11.77 (103.90)
	Adolescent offending controls	731.87 (146.65)	29.81 (86.85)
	Adolescent non-offending controls	666.70 (96.38)	2.43 (68.98)
Positive	Adolescent sexual abusers	812.71 (158.55)	2.55 (110.95)
	Adolescent offending controls	696.48 (118.46)	-5.57 (79.88)
	Adolescent non-offending controls	657.43 (83.70)	-6.83 (48.84)
Negative	Adolescent sexual abusers	798.32 (141.00)	-11.83 (84.06)
	Adolescent offending controls	701.56 (120.15)	-.50 (80.97)
	Adolescent non-offending controls	657.22 (94.13)	-7.05 (55.88)
Aggression	Adolescent sexual abusers	802.20 (131.40)	-7.96 (74.90)
	Adolescent offending controls	691.57 (122.98)	-10.49 (95.55)
	Adolescent non-offending controls	629.52 (78.74)	-34.74 (63.81)
Sexual	Adolescent sexual abusers	861.80 (166.47)	51.65 (122.29)
	Adolescent offending controls	696.48 (142.42)	-5.58 (80.81)
	Adolescent non-offending controls	649.71 (92.71)	-14.55 (63.83)

Figure 1 displays the patterns in emotional Stroop bias effects experienced by the different groups. One-way ANOVA's yielded a significant difference between the groups for the sexual Stroop bias scores, $F(2, 61) = 3.24$, $p < .05$, $\omega = .26$. However this difference was nearing non-significance at $p = .046$. Therefore, post-hoc analysis using Bonferroni correction did not display significant differences between the groups. The more stringent nonparametric testing did not yield significant differences for sexual Stroop bias scores, $H(2) = 3.783$, ns, between the three groups.

Figure 1. *Mean Emotional Stroop Bias Scores: Smith and Waterman Stimulus Set*

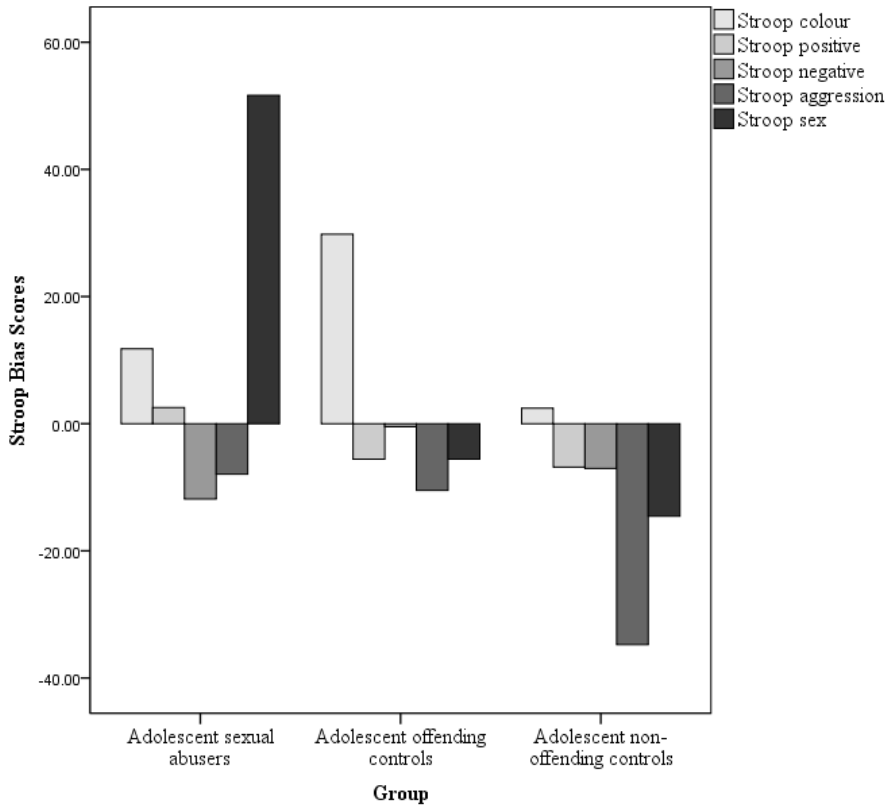
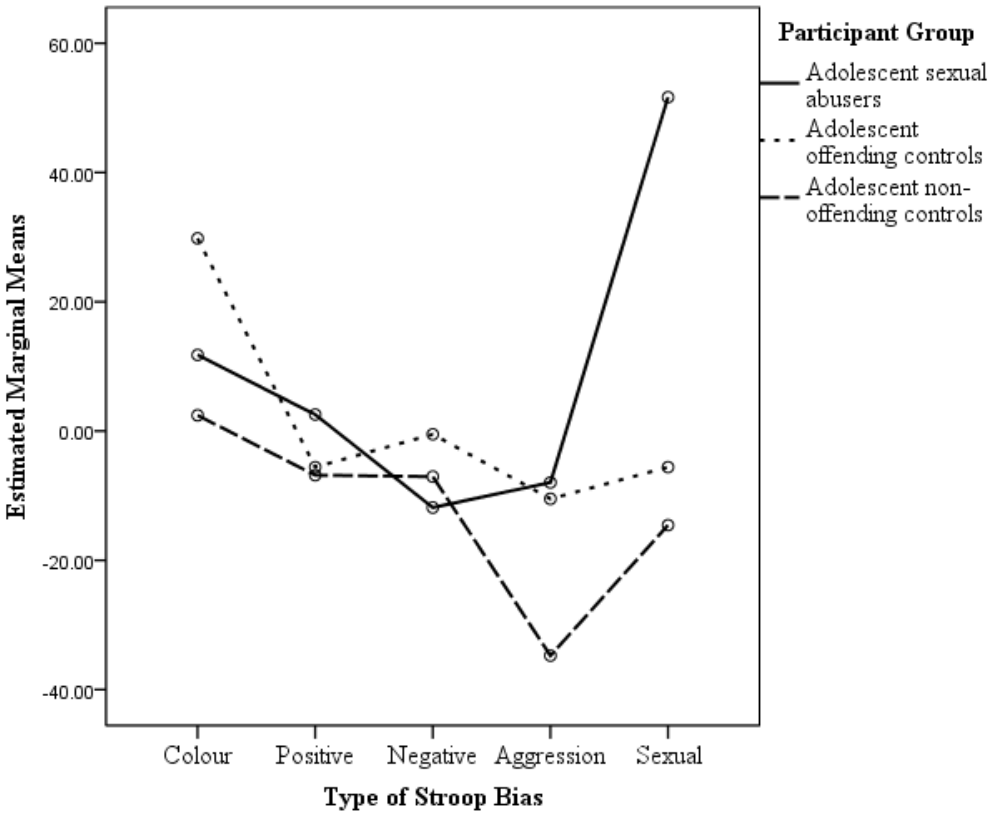


Figure 2 displays the means of the main effects of the two-way mixed ANOVA with type of Stroop bias as the within-subjects variable and participant group as the between-subjects variable. Sphericity Assumed F-ratios were used because the assumption of sphericity was met. There was a significant main effect of type of Stroop bias, $F(4, 244) = 3.00$, $p < .05$, and non-significant effect of the interaction, $F(8,244) = 1.78$, ns. Simple contrasts reveal that the main effect of type of Stroop bias was due to the fact that, overall, the colour Stroop bias scores were significantly higher than negative Stroop bias scores, $F(1, 61) = 5.08$, $p < .05$, $r = .28$ (medium effect), and aggression Stroop bias scores, $F(1, 60) = 11.22$, $p < .05$, $r = .39$ (medium-high effect).

For the between-groups analysis, the Levene's test for each of the repeated-measures variables indicated that the assumption of homogeneity of variance was violated for the sexual Stroop bias scores, therefore, the accuracy of the F-test for participant group is compromised and should be interpreted with caution. The test of between-subjects effects did not in fact yield a significant main effect of participant group, $F(2, 61) < 1$, ns, indicating that there was more error than variance created by the experiment.

Figure 2. *Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman Stimuli*



Finally, the results from the analysis of covariance from this word stimulus set did not yield any significant relationships between the mean RT's and Stroop bias effects and the covariates: age, BACS (CD), BACS (EC), Hayling and Brixton tests of executive function and BPVS raw scores.

Newly-Derived Word Stimulus Set

Mean RT's and emotional Stroop bias scores are displayed in Table 5 (medians in Appendix F) for the newly-derived word stimulus set. One-way ANOVA yielded significant differences for mean RT's between the adolescent sexual abusers and the non-offending adolescent controls for the MSA, $F(2, 62) = 5.65$, $p < .05$ words.

Table 5

Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Newly-Derived Word Stimulus Set

Word Category	Adolescent Group	Mean RT, ms, (SD)	Stroop Bias Score, ms (SD)
EPD	Adolescent sexual abusers	795.83 (161.01)	-1.06 (86.00)
	Adolescent offending controls	694.26 (132.22)	-13.76 (97.69)
	Adolescent non-offending controls	653.40 (103.57)	-.58 (72.53)
MEPD	Adolescent sexual abusers	796.89 (165.14)	
	Adolescent offending controls	708.00 (123.36)	
	Adolescent non-offending controls	654.00 (100.88)	
SA	Adolescent sexual abusers	785.65 (183.15)	20.37 (130.83)
	Adolescent offending controls	703.81 (136.25)	17.33 (94.49)
	Adolescent non-offending controls	654.06 (104.60)	4.24 (97.70)
MSA	Adolescent sexual abusers	765.28(135.37)	
	Adolescent offending controls	686.48 (121.85)	
	Adolescent non-offending controls	649.83 (90.41)	
PD	Adolescent sexual abusers	788.03 (158.49)	13.83 (125.39)
	Adolescent offending controls	717.69 (163.49)	8.84 (103.27)
	Adolescent non-offending controls	661.17 (119.87)	12.78 (100.51)
MPD	Adolescent sexual abusers	774.20 (183.40)	
	Adolescent offending controls	708.85 (125.76)	
	Adolescent non-offending controls	648.39 (77.87)	

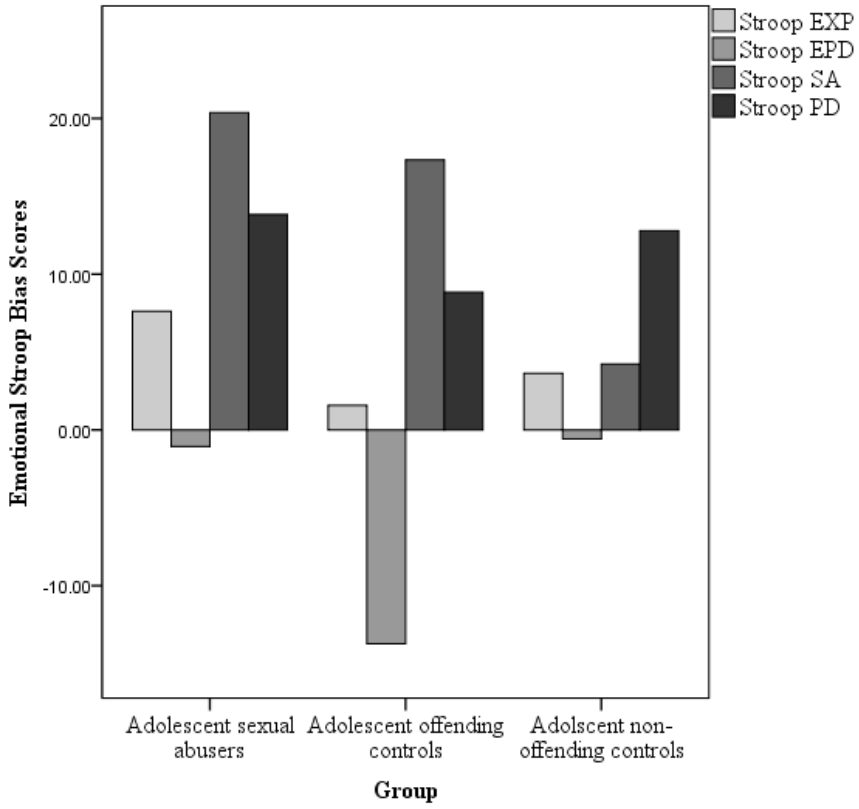
Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories.

Nonparametric testing yielded a significant difference for mean EPD RT's, H(2) = 11.92, with adolescent sexual abusers taking significantly longer to respond to the stimuli than adolescent offending controls (U = 135.00, r = .37) and adolescent non-offending controls (U = 111.00, r = .48). Adolescent sexual abusers also took significantly longer to respond to the Stroop task when compared to the adolescent non-offending controls for mean MEPD (U = 127.00, r = .42); SA, (U = 135.00, r = .40); PD (U = 134.00, r = .40); and MPD (U = 153.00, r = .34) mean RT's.

Overall, significant differences were found between the groups for the experimental RT's, H(2) = 10.75, $p < .05$ and the matched RT's, H(2) = 9.89, $p < .05$. Post-hoc analysis revealed that the adolescent sexual abusers took significantly longer to color-name when compared to adolescent non-offending controls for mean experimental RT's (U = 114.00, r = .47), and mean matched RT's (U = 119.00, r = .45).

No significant differences were yielded for the emotional Stroop bias scores. However, Figure 3 displays the patterns displayed between the groups. It appears that the emotional/personality descriptors developed from adult sexual abusers' responses to what they would describe as most at risk from them are not eliciting any Stroop bias, whereas the other word categories are yielding similar patterns across the groups.

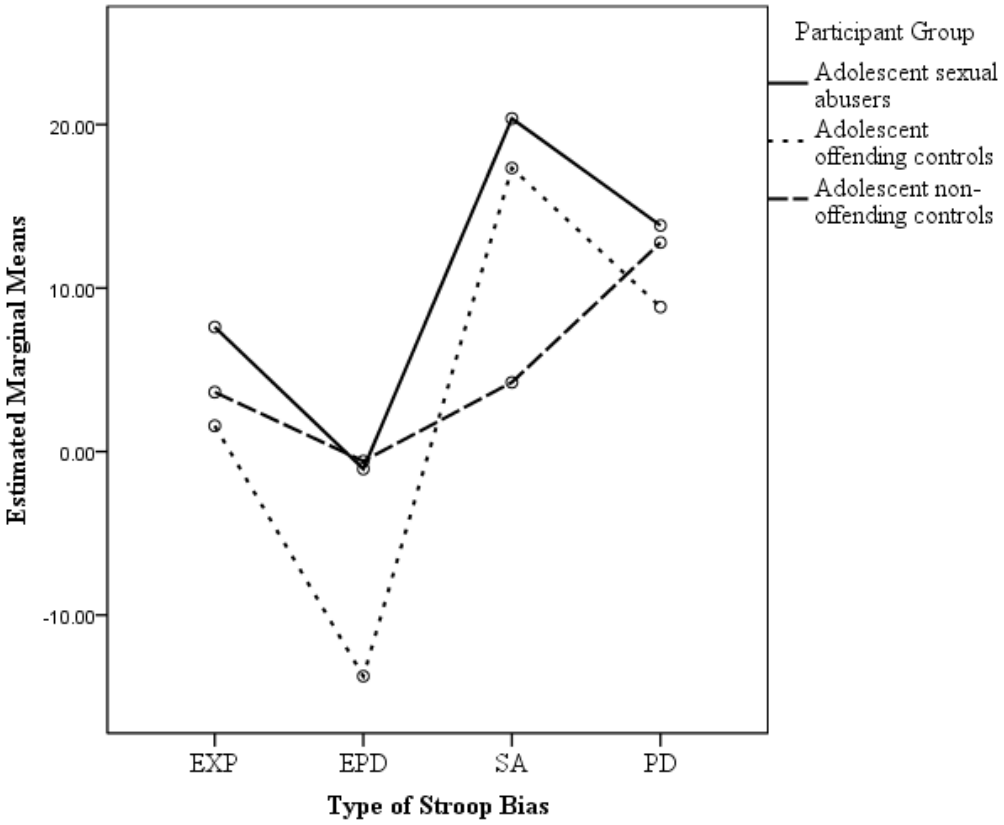
Figure 3. *Emotional Stroop Bias Scores for Newly-Derived Word Stimuli*



The means of the main effects of the two-way mixed ANOVA that was run with type of Stroop bias as the within-subjects variable, and participant group as the between-subjects variable, are displayed in Figure 4. Mauchly's test of sphericity was significant indicating that the assumption of sphericity was violated and the Greenhouse-Geisser correction of the degrees of freedom was used to assess the significance of the F-statistics. There was a non-significant main effect of type of Stroop bias, $F(1.81, 111.92) < 1$, ns, and a non-significant interaction effect, $F(3.61, 111.92 < 1$, ns. Since both F-ratios were less than 1 there appeared to be more error than variance created by the experiment. The Levene's test for each of the repeated-measures variables was not significant indicating that the assumption of homogeneity of variance was not violated for the test of between-subjects effects. The test of

between-subjects effects did not in fact yield a significant main effect of participant group, $F(2, 62) < 1$, ns, indicating that there was more error than variance created by the experiment for this analysis as well.

Figure 4. *Means for the Main Effects of Type of Stroop Bias and Participant Group: Newly-Derived Word Stimuli*



Finally, analysis of covariance showed that the covariates (i.e., age, BACS (CD), BACS (EC), Hayling and Brixton tests of executive function and BPVS raw scores) were not significantly related to the mean RT's or Stroop bias effects from the newly-derived word stimulus set.

Discussion

Very little research has been conducted with adolescent offender samples and the emotional Stroop task. The current study represents the first study using these offence-specific word stimuli to examine biases towards sexual interest information in adolescent offender samples. The RT and Stroop bias outcome data for the adolescent samples appears to be more unsystematic and weaker than the adult data reported in Chapter 3. Although it is not unusual to observe skewed data, extreme values, or numerous violations to the assumptions of parametric testing when dealing with RT data (Gress & Laws, 2009; Miller, 1991), it was difficult to identify reliable patterns in response bias for this sample. Therefore, the results appear arbitrary and are mixed regarding whether the task was actually able to measure sexual interests in the adolescent samples. Specifically, the task was unable to differentiate between the groups on most of the Stroop word categories. However, some response patterns were observed across the word categories, albeit not significant in nature for the most part.

The significant differences in mean RT's displayed medium to high effect sizes between the groups for most of the word categories. However, this finding simply tells us that the adolescent sexual abusers took consistently longer across the word categories to colour-name than the two control groups. This finding is consistent with the results from Chapter 3, and with the findings from Smith and Waterman (2004) and Price and Hanson (2007). The reasons why the sexual abusers (and offender groups more generally) consistently respond more slowly to the word stimuli in this task are unknown. While the factors that were tested in this study such as the Hayling and Brixton tests of executive function and the level of vocabulary as tested by the BPVS did not display a significant relationship with the mean RT's for the word categories. It is not surprising that the analysis of covariance did not yield

significant results considering the lack of Stroop effects experienced overall, and given the nonparametric nature of the data. However, this would lead us to the conclusion that there is a separate reason why the sexual abusers take the longest to respond that was not controlled for in this study. For example, it could be that neurological impairments such as impulsivity, learning difficulties (Smith, 2009), or dysfunctions in the frontal cortex that have been linked with difficulties with behavioural inhibition on tests of executive function (Dolan, Millington, & Park., 2002; Langevin, Lang, Wortzman, Frenzel, & Wright, 1989; Ponseti et al., 2001; Stone & Thompson, 2001; Valliant et al., 2000) are interacting with stimulus salience.

The groups exhibited some Stroop interference for incongruent colour stimuli, indicating the ability to experience Stroop interference effects that are consistent with what we would expect of colour-word Stroop interference (Stroop, 1935). There was also a significant difference between the adolescent sexual abusers, and the non-offending controls, for sexual Stroop bias score however this effect was small, and did not hold up in the post-hoc analysis between the groups. Aside from the small effect of the sexual word stimuli between the sexual abusers and non-offending controls, none of the adolescent groups appeared to be reacting to the words presumably containing emotional content. This is confirmed by the lack of a main effect of group conducted in the two-way ANOVA.

The response patterns from the newly derived word set were not significant. However, it is interesting to note that none of the groups experienced Stroop interference for the emotional/personality descriptors. The emotional/personality descriptors displayed the greatest effect in Chapter 3 between the adult sexual abusers and the non-offending controls. The finding from the adolescent study might then

further suggest that the word stimuli from the emotional/personality descriptor category are more specific to the sexual interests of adult sexual abusers.

Interestingly, all of the adolescent groups experienced a degree of bias for the remaining word categories (i.e., sexual actions and physical descriptors of individuals). The adult groups also displayed Stroop interference for the sexual action and physical descriptor word categories. This might suggest that the word stimuli from these categories are of a more general sexual interest to individuals.

The overall experimental effect (i.e., Stroop EXP) of the newly-derived word stimuli was very small across all of the adolescent groups and the adult groups from Chapter 3. This would suggest that, overall, the word stimuli are not uniform enough when analysed together to be able to discriminate between groups of individuals.

Limitations and Directions for Future Research

This study suffered similar limitations to those outlined in Chapter 3. This sample of sexual abusers was also heterogeneous in nature and this could have contributed to the mixed findings of the emotional Stroop task. Similarly, the variation in living environment and risk level of the offenders could have had an impact on Stroop effects. Had time allowed, the use of mood questionnaires could have been advantageous to test whether the results from the mood questionnaires correlated with the Stroop results on the positive and negative word categories.

This study was the first to test these subgroups of adolescents using the emotional Stroop task to measure sexual interests. Therefore, to completely abandon future efforts using this task to assess sexual interest is not recommended. Instead, future research should consider amending the stimulus words to suit younger respondents or amending the type of stimulus to cater to and control for the additional

needs adolescent offending samples may require. For example, Boonstra, Oosterlaan, Sergeant, & Buitelaar (2005) conducted a meta-analysis on ADHD's effect on Stroop results. Impairments in the naming of the colours were found for those with ADHD, rather than on interference effects. The same results were observed for adolescent samples (van Mourik, Oosterlaan and Sergeant, 2005). Geurts, Verte, Oosterlaan, Roeyers, and Sergeant (2004) have reported that individuals with ADHD were less sensitive to interference caused by word meaning. Although this study tested two levels of executive function and level of vocabulary understanding of the adolescents it could have been beneficial to measure whether a diagnosis of ADHD had an effect on the Stroop results. Unfortunately, confirmed diagnoses of ADHD in these samples were not obtainable.

It would have been useful to explore developmental issues that may have had an effect on Stroop results. Adolescents have fewer life experiences, and their cognitive and affective abilities may be less developed than adults (Grant, 2009). A solution to this problem may be to employ a variant of the Stroop task that uses picture stimuli rather than word stimuli. This may be a more appropriate methodology to use with this group because they may have weaker emotional or cognitive associations and fewer information-processing pathways established to be able to respond to the word stimuli that were used in this study.

CHAPTER 5

Comparison of Measurements of Deviant Sexual Interest Between Adult and Adolescent Samples Using the Emotional Stroop Task

Chapter 3 examined differences in mean RT's and emotional Stroop bias scores between adult sexual abusers, adult offending controls and adult non-offending controls. Chapter 4 introduced the results of three adolescent samples to the same emotional Stroop word sets. The aim of this chapter was to compare the datasets from the adolescent and adult studies, and to provide evidence for similarities and differences in response bias to sexual word stimuli amongst the groups.

Introduction

One of the initial aims of this thesis was to examine Stroop interference effects of offence-related and sexual interest information experienced by participants from different age groups, offending groups and levels of executive function. Chapter 3 presented the results from three adult samples. For the Smith and Waterman (2004) word stimulus set, the adult sexual abusers took significantly longer to colour-name the word stimuli from all of the word categories than the non-offending control subjects. Significant Stroop interference effects were evident between the adult sexual abusers and the adult non-offending controls for the positive, negative and sexual words. Negative Stroop bias scores for the target words (i.e., positive, negative, aggressive and sexual) indicated that Stroop interference effects were not experienced at all by the non-offending control subjects. The adult sexual abusers were again slowest to respond in mean RT's for the newly-derived word stimuli for all of the word categories. There was a significant difference between sexual abusers and non-offending controls for the emotional/personality descriptor words that were derived for the purpose of tapping into more specific elements of sexual interest for sexual abusers. However, the task was not able to differentiate between offender types for any of the word stimuli.

The results from the adolescent study in Chapter 4 were weaker than the results from the adult study in Chapter 3. However, the adolescent sexual abuser sample was also consistently slower in responding to word stimuli than the control groups. Adolescent sexual abusers and non-offending controls significantly differed for the sexual Stroop bias. Although the effect was small and it did not hold up in the post-hoc analysis. Overall, none of the adolescent groups appeared to be reacting to

the Smith and Waterman (2004) word stimuli. The adolescents' response patterns to the newly-derived word set were interesting but non-significant. Overall, the experimental effects from this word stimulus set were very small for the adolescent groups and the task was again unable to differentiate between the groups for most of the Stroop word categories.

Study Purpose

Thus far, the adolescent and adult samples have been tested separately to examine the differences between the sexual abusers and respective comparison groups. This chapter combines the datasets from Chapters 3 and 4 to allow for a cross-comparison between the adult and adolescent samples. The aims of this chapter then, are:

- To provide empirical evidence of the similarities and differences in emotional Stroop response patterns between the adult and adolescent groups;
- To determine to what degree adolescent sexual abusers and adult sexual abusers respond similarly to this task

Method

Participants

For the purpose of cross-comparison, the data from the adult samples from Chapter 3 and the adolescent samples from Chapter 4 were combined and used as the participants for this chapter. Therefore, the sample consisted of 24 adolescent sexual abusers, 21 adolescent offending controls, 20 adolescent non-offending controls, 27 adult sexual abusers, 21 adult offending controls and 35 non-offending adult controls.

Additional Measures

See Chapter 2 for a full description of the apparatus, materials, analysis and procedures used in this study (pp 38-48). Six groups of participants were compared in the post-hoc analyses. Therefore, it was necessary to be careful in selecting the post-hoc comparisons made between the groups. Five post-hoc comparisons were conducted for each Kruskal-Wallis test made on the dependent variables: (1) adolescent sexual abusers compared to adult violent offenders; (2) adolescent sexual abusers compared to adult non-offending controls; (3) adult sexual abusers compared to adolescent offending controls; (4) adult sexual abusers compared to adolescent non-offending controls; and (5) adult sexual abusers compared to adolescent sexual abusers. The critical value from the Bonferroni corrections used in the post-hoc analyses was .01, making the analysis of significant differences between the groups quite restrictive.

Violations of Parametric Testing

The data regarding participant age, the measures used in the study and the number of errors made on the separate Stroop tasks were not normally distributed. The assumption of homogeneity of variance was violated for participant age, the Hayling test of executive function and the BACS subscales. Therefore, significant differences between the groups for these variables were tested using Kruskal-Wallis tests.

Stroop Task Violations of Parametric Testing

Violations to the assumption of normally distributed data and homogeneity of variance, for both Stroop tasks on the mean RT's and Stroop bias scores have been

outlined in Chapters 3 (pp 55-57) and 4 (pp 84-86). Therefore, the findings were not repeated here.

Results

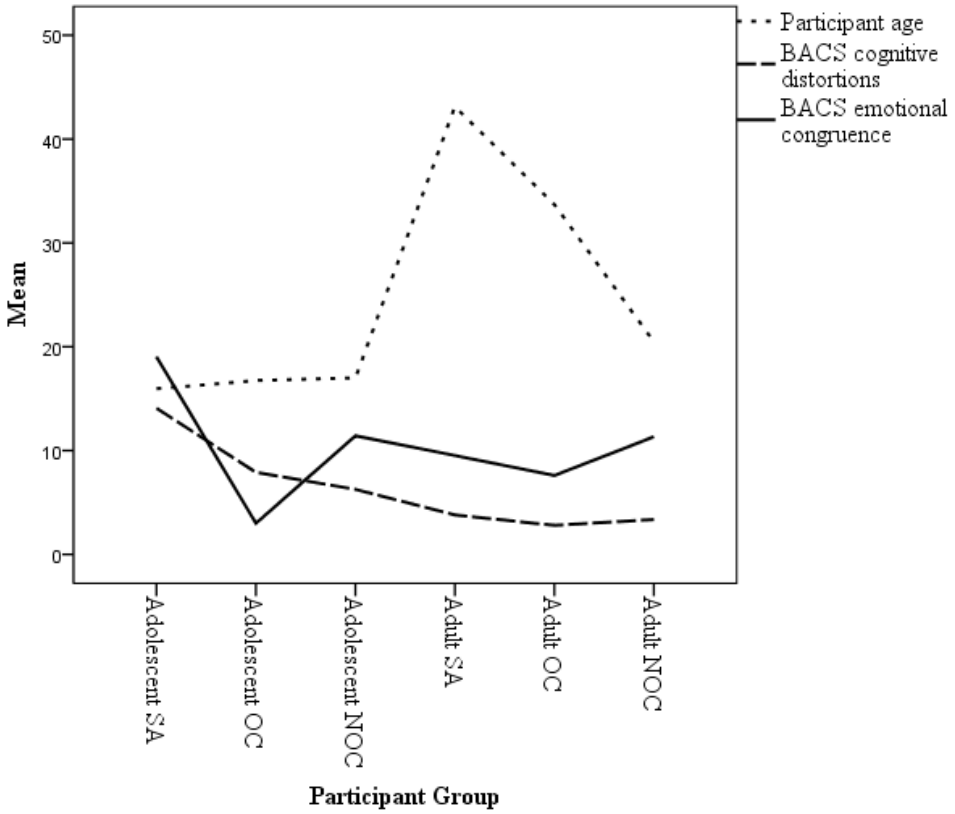
Significant differences were evident between the groups on the following variables and measures: participant age, $H(5) = 126.68$, $p < .001$; BACS (CD), $H(5) = 38.68$, $p < .001$; BACS (EC), $H(5) = 28.65$, $p < .001$; Hayling scores, $H(5) = 13.85$, $p < .05$; Brixton scores, $H(5) = 26.97$, $p < .001$; BPVS, $H(5) = 81.72$, $p < .001$; number of Stroop errors (Smith and Waterman stimulus set), $H(5) = 21.11$, $p < .001$; and number of Stroop errors (newly-derived stimulus set), $H(5) = 10.45$, $p < .05$. Chapter 3 (pp 58) and Chapter 4 (pp 88) provide summaries of the means and standard deviations for these variables.

Figure 1 is provided to demonstrate the differences between the groups on mean age and self-reported beliefs about children. Post-hoc analyses revealed that adolescent sexual abusers were significantly younger than the adult offending controls ($U = 0$, $r = .86$), adult non-offending controls ($U = 6$, $r = .84$), and adult sexual abusers ($U = 0$, $r = .86$). The adult sexual abusers were significantly older than the adolescent offending controls ($U = 0$, $r = .85$), and adolescent non-offending controls ($U = 0$, $r = .86$).

Additionally, adolescent sexual abusers scored significantly higher on BACS cognitive distortions than the adult offending controls ($U = 62.5$, $r = .63$), adult non-offending controls ($U = 127.5$, $r = .60$), and adult sexual abusers ($U = 99.5$, $r = .58$). The adult sexual abusers scored significantly lower on cognitive distortions than the adolescent offending controls ($U = 107.5$, $r = .50$), and adolescent non-offending controls ($U = 131$, $r = .39$). Finally, on the BACS emotional congruence subscale, the

adolescent sexual abusers scored significantly higher than the adult offending controls ($U = 114$, $r = .45$), while the adult sexual abusers scored significantly lower than the adolescent sexual abusers ($U = 164$, $r = .39$) on emotional congruence with children.

Figure 1. *Mean Age, BACS (CD) and BACS (EC): Adolescent and Adult Groups*

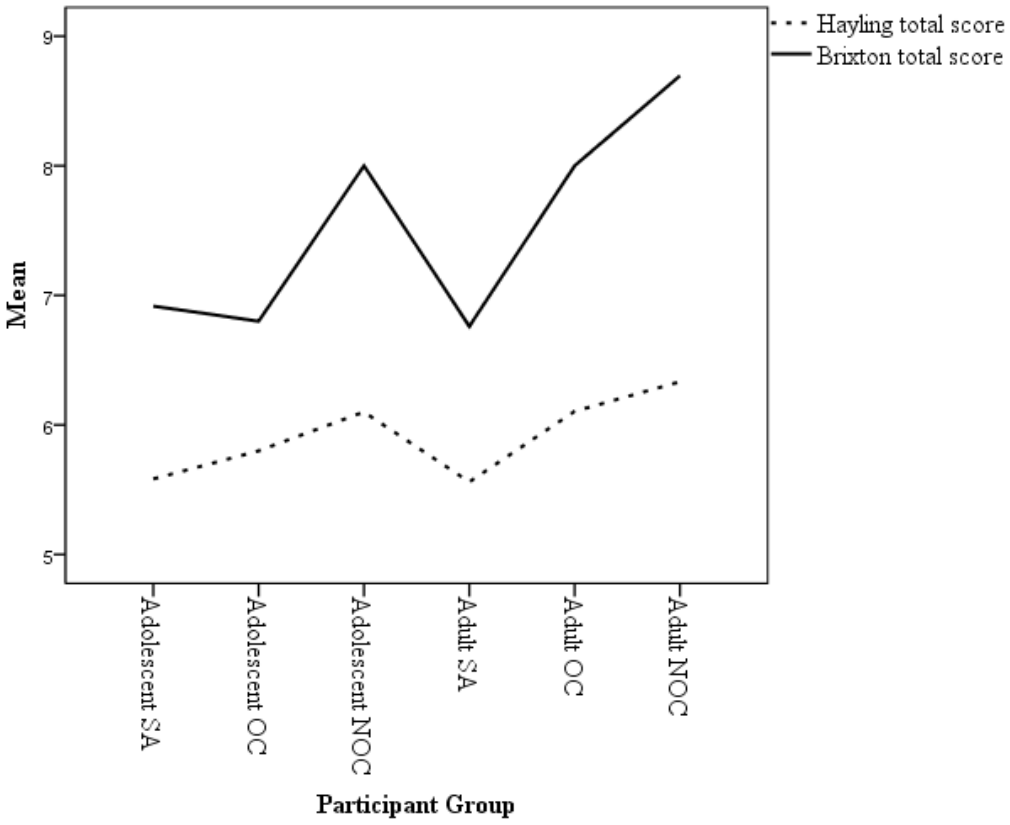


Note. SA indicates sexual abusers, OC indicates offending controls, and NOC indicates non-offending controls.

Figure 2 is provided to demonstrate where significant differences between the groups were evident on the tests of executive function. Post-hoc analysis revealed that the adolescent sexual abusers scored significantly lower on the Brixton test than adult non-offending controls ($U = 201.5$, $r = .46$), and post-hoc analyses from the one-way ANOVA (i.e., non-parametric post-hoc was rendered non-significant after Bonferroni

correction) showed that the youth sexual abusers also had significantly lower Hayling test scores than non-offending adult controls.

Figure 2. *Tests of Executive Function: Adolescent and Adult Groups*



Note. SA indicates sexual abusers, OC indicates offending controls and NOC indicates non-offending controls.

Additionally, the BPVS raw scores were significantly lower for the adolescent sexual abusers when compared to adult offending controls ($U = 21$, $r = .78$), adult non-offending controls ($U = 29$, $r = .79$) and adult sexual abusers ($U = 34.5$, $r = .76$), while adult sexual abusers scored significantly higher on the BPVS than the adolescent offending controls ($U = 29.5$, $r = .75$) and adolescent non-offending controls ($U = 96.5$, $r = .53$). Finally, the adult sexual abusers made significantly fewer

mistakes than the adolescent offending controls (U = 124.5, r = .46) on the Smith and Waterman stimulus set; and the adolescent sexual abusers made far more errors than the adult offending controls (U = 139.5, r = .38) on the new stimulus set.

Smith and Waterman Stimulus Set

Summary tables of the mean RT's and Stroop bias scores are provided in Chapters 3 (pp 61) and 4 (pp 89) for the varying age groups. One-way ANOVA displayed a significant difference between the groups for mean aggression RT's, $F(5, 142) = 13.783$, $p < .001$. Multiple comparisons reveal that the adolescent sexual abusers took significantly longer to colour-name aggressive word stimuli than the adult non-offending controls ($p < .05$). Additionally, the adult sexual abusers took significantly longer to respond to aggressive word stimuli than the adolescent non-offending controls ($p < .05$).

The Kruskal-Wallis tests found significant differences on mean RT's between the adolescent and the adult groups for neutral words, $H(5) = 46.99$, $p < .001$; positive words, $H(5) = 54.27$, $p < .001$; negative words, $H(5) = 49.75$; colour words, $H(5) = 42.72$, $p < .001$; and sexual words, $H(5) = 56.06$, $p < .001$.

Table 1 displays the results from the non-parametric post-hoc comparisons. Generally, adolescent sexual abusers took significantly longer to respond to the individual word category stimuli than the adult non-offending control subjects. Additionally, the adult sexual abusers typically took significantly longer to respond to word stimuli than the adolescent non-offending controls.

The principal finding from the analysis of the mean RT's was the finding from the analysis of the sexual word category between the groups. Adolescent sexual abusers took significantly longer to respond to the sexual word stimuli than both the

adult offending controls ($U = 125$, $r = .43$) and adult non-offending controls, ($U = 47$, $r = .75$). Additionally, the adult sexual abusers to significantly longer to respond to the sexual word stimuli than both the adolescent offending controls ($U = 148$, $r = .56$) and adolescent non-offending controls ($U = 91$, $r = .56$). Interestingly, significant differences were not evident between the two sexual abuser groups for mean sexual RT's and the effect sizes from the significant post-hoc comparisons are quite large.

Table 1

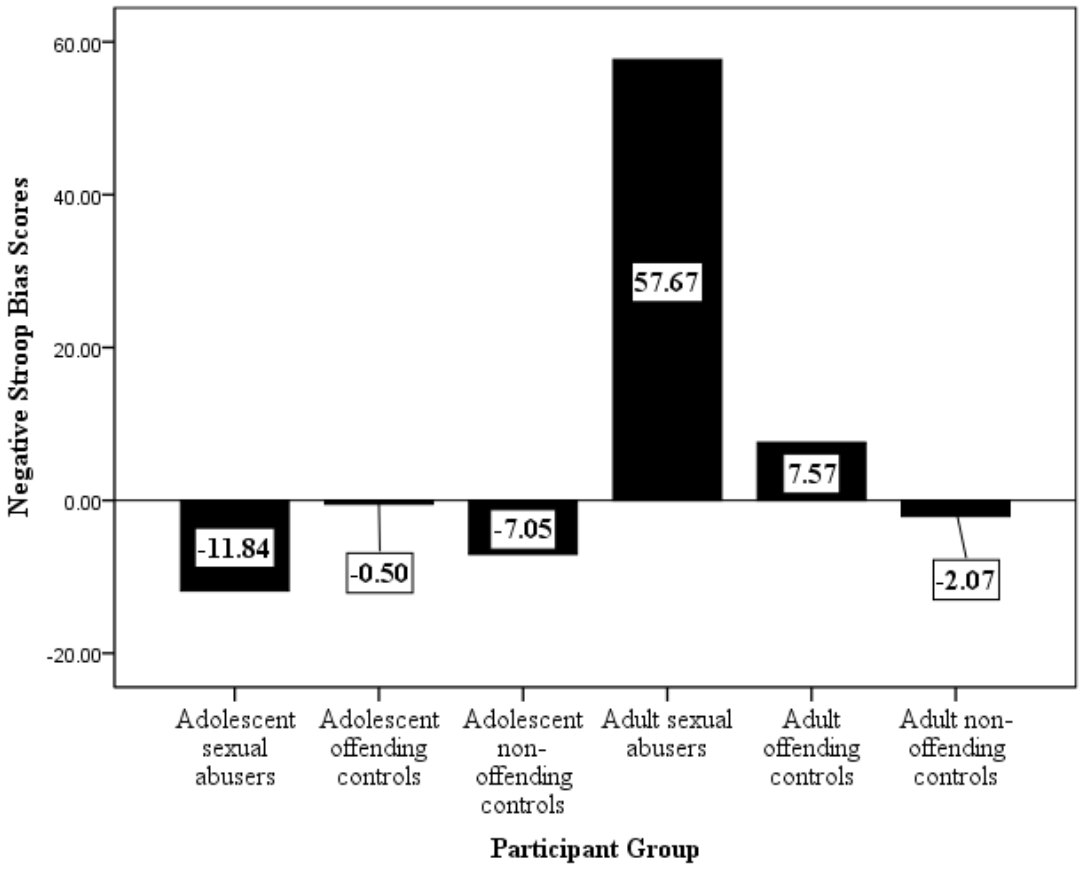
Significant Differences from Non-Parametric Post-Hoc Comparisons of Mean RT's

Post-hoc comparisons		Word Categories				
		Neutral *U* (*r*)	Positive *U* (*r*)	Negative *U* (*r*)	Colour *U* (*r*)	Sexual *U* (*r*)
Adolescent SO	Adult OC	125 (.43)	148 (ns)	164 (ns)	166 (ns)	125 (.43)
Adolescent SO	Adult NOC	71 (.70)	62 (.72)	78 (.69)	82 (.68)	47 (.75)
Adult SO	Adolescent SO	251 (ns)	307 (ns)	300 (ns)	313 (ns)	276 (ns)
Adult SO	Adolescent OC	193 (ns)	156 (ns)	151 (ns)	192 (ns)	148 (.56)
Adult SO	Adolescent NOC	139 (.41)	131 (.44)	111 (.50)	130 (.44)	91 (.56)

Note. SO indicates sexual offenders, OC indicates offending controls and NOC indicates non-offending controls. *U* indicates the test statistics from the Mann-Whitney post-hoc tests and *r* indicates the effect size for the comparison.

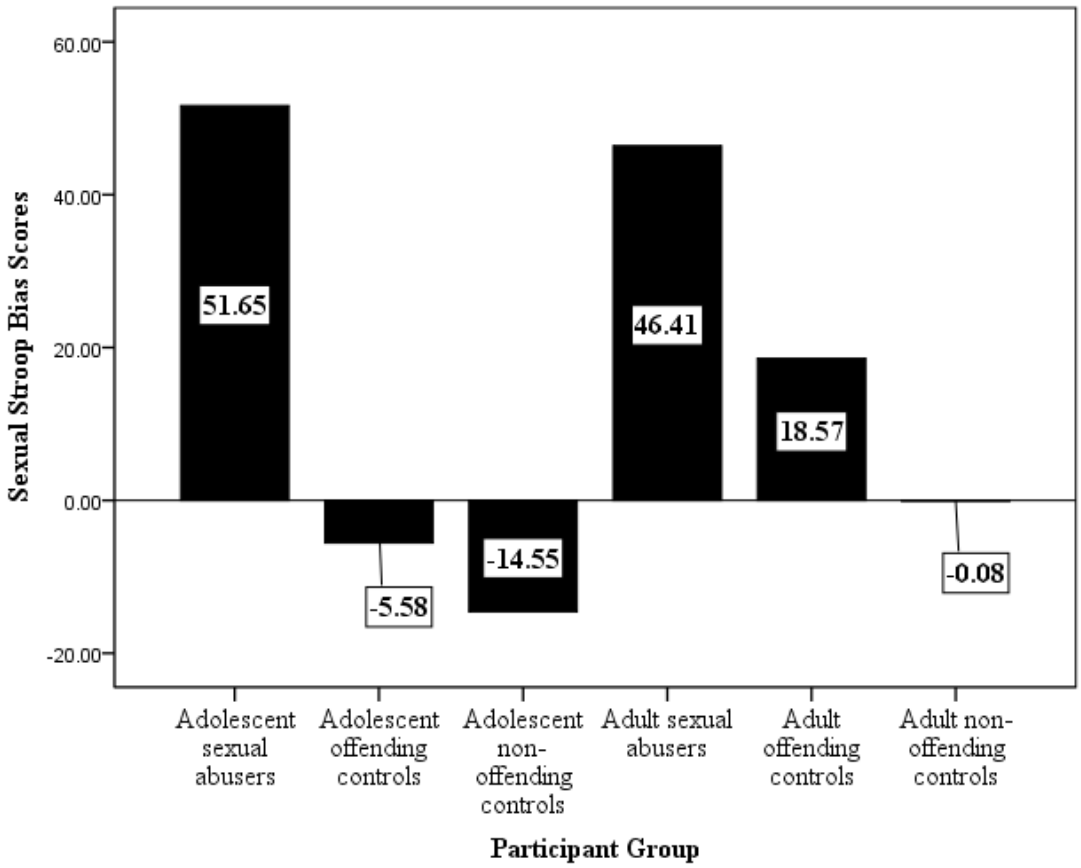
For the analysis of the Stroop bias scores, significant differences were evident between the groups for the negative, $H(5) = 14.28$, $p < .05$ and sexual Stroop bias scores, $H(5), = 13.13$, $p < .05$. Figure 3 displays the mean negative Stroop bias scores for the groups. Post-hoc analysis showed that the adult sexual abusers had significantly larger Stroop effects for negative word stimuli than the adolescent sexual abusers ($U = 166$, $r = .42$) and adolescent non-offending controls ($U = 138$, $r = .41$).

Figure 3. *Negative Stroop Bias Scores by Participant Group*



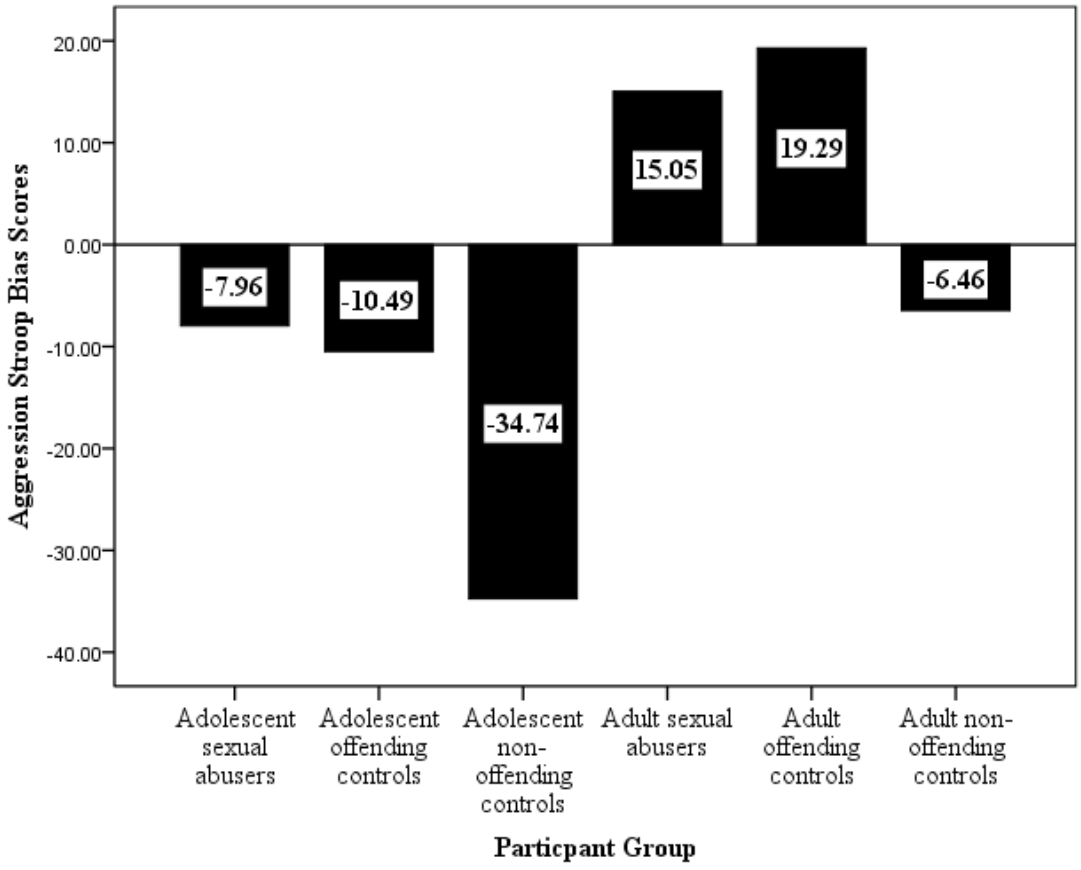
For the sexual Stroop bias scores (see Figure 4), post-hoc comparisons revealed that the adult sexual abusers had significantly larger Stroop effects than the adolescent non-offending controls ($U = 145$, $r = .39$). Post-hoc comparisons did not elicit any additional significant differences between the groups for the sexual Stroop bias scores. However, it is interesting to see that these stimuli are most salient for the sexual abuser groups, and that the control groups (with the exception of the violent offending controls) did not elicit Stroop bias for the sexual word stimuli.

Figure 4. *Sexual Stroop Bias Scores by Participant Group*



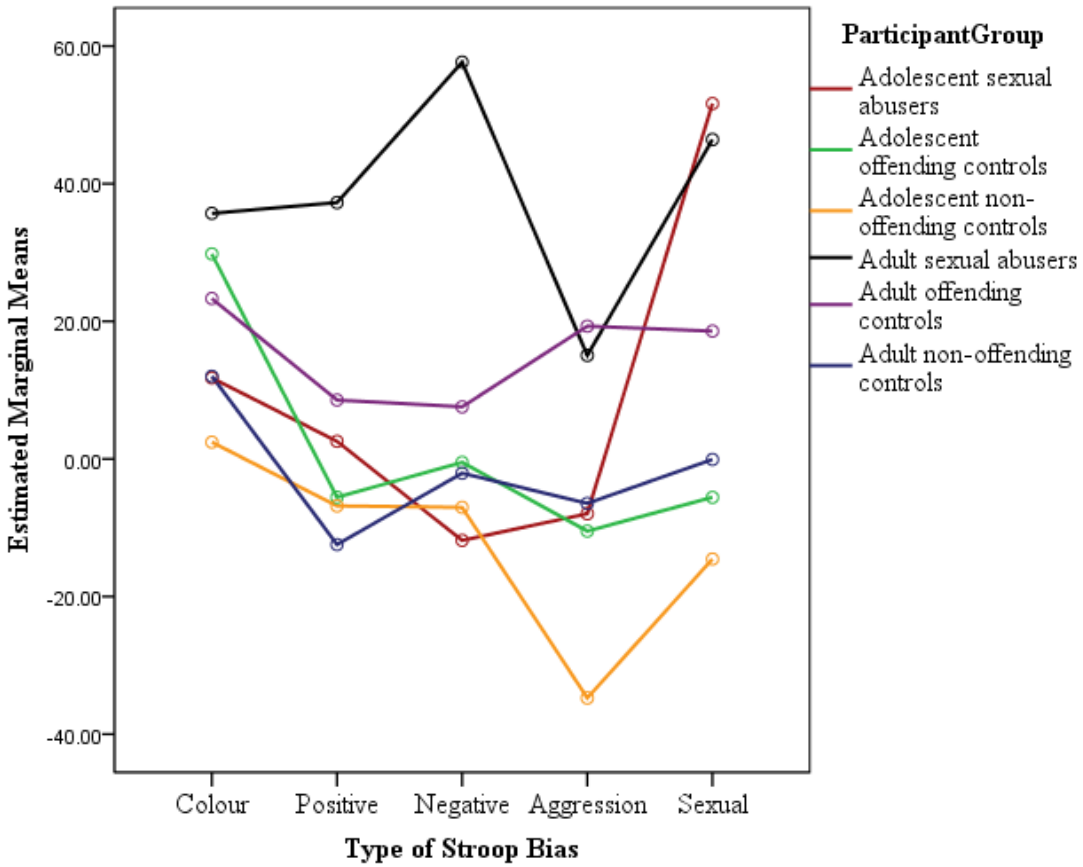
Finally, although there were no significant differences between the groups for the aggression Stroop bias scores it is interesting to note the patterns in response that were evident for this comparison (see Figure 5). For example, these word stimuli did not elicit Stroop bias for any of the adolescent groups. Additionally, both the adult violent offenders and sexual abusers were similarly affected by these word stimuli.

Figure 5. *Aggression Stroop Bias Scores by Participant Group*



Two-way mixed ANOVA, with type of Stroop bias as the within-subjects variable and participant group as the between-subjects variable, violated the assumption of sphericity ($p = .037$). Therefore, the Greenhouse-Geisser correction was used to assess the significance of the F-statistics. The main effect of type of Stroop bias was significant, $F(3.76, 533.63) = 4.40$, $p < .05$. However, the effect of the interaction was non-significant, $F(18.79, 533.63) = 1.58$, ns. Therefore, the mean RT's within the groups were significantly different across the types of Stroop bias. Post-hoc analysis with Bonferroni correction revealed that the colour Stroop bias scores significantly differed from positive Stroop bias scores, $F(1, 142) = 5.95$, $r = .20$, negative Stroop bias scores, $F(1, 142) = 4.02$, $r = .17$, and aggression Stroop bias scores, $F(1, 142) = 15.39$, $r = .31$. Despite this significant effect of type of Stroop bias, it should be noted that the effect sizes are considered to be small. The main effect of participant group was not significant, $F(5, 142) = 2.25$, ns, indicating that the participant groups did not elicit different Stroop bias scores. Figure 6 displays the means of the main effects of the two-way mixed ANOVA.

Figure 6. *Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman Stimuli*



Newly-Derived Stimulus set

Significant differences between the adolescent and adult samples for mean RT's were again evident for all word categories in this stimulus set: emotional/personality descriptors (EPD), $H(5) = 48$, $p < .001$; matched EPD, $H(5) = 39.42$, $p < .001$; sexual actions (SA), $H(5) = 34.21$, $p < .001$; matched SA, $H(5) = 37.27$, $p < .001$; physical descriptors (PD), $H(5) = 40.36$, $p < .001$; matched PD, $H(5) = 43.71$, $p < .001$; experimental words (EXP), $H(5) = 46.31$, $p < .001$; and matched EXP, $H(5) = 43.71$, $p < .001$.

Tables listing the mean RT's and Stroop bias scores for this stimulus set are provided in Chapters 3 (pp 67) and 4 (pp 93) for the varying age groups. Table 2 provides a summary of the results from the non-parametric post-hoc comparisons. The adolescent sexual abusers were consistent in taking significantly longer to respond to the word stimuli than the adult non-offending controls, while the adult sexual abusers typically took significantly longer to respond to the word stimuli than the adolescent non-offending controls. The results are weaker than the post-hoc comparisons from the Smith and Waterman (2004) word stimulus set. However, medium effect sizes were yielded.

Table 2

Significant Differences from Non-Parametric Post-Hoc Comparisons of Mean RT's: Newly-Derived Word Stimuli

Post-hoc comparisons		Word Categories							
		EPD	MEPD	SA	MSA	PD	MPD	EXP	MEXP
		U (*r*)	*U* (*r*)	*U* (*r*)	*U* (*r*)	*U* (*r*)	*U* (*r*)	*U* (*r*)	*U* (*r*)
Adolescent SO	Adult OC	139 (.38)	61 (.39)	166 (ns)	175 (ns)	180 (ns)	172 (ns)	141 (ns)	149 (ns)
Adolescent SO	Adult NOC	61 (.72)	104 (.63)	142 (.56)	111 (.62)	104 (.63)	112 (.62)	83 (.68)	100 (.64)
Adult SO	Adolescent SO	316 (ns)	289 (ns)	316 (ns)	320 (ns)	305 (ns)	299 (ns)	315 (ns)	315 (ns)
Adult SO	Adolescent OC	181 (ns)	204 (ns)	182 (ns)	192 (ns)	194 (ns)	190 (ns)	188 (ns)	190 (ns)
Adult SO	Adolescent NOC	150 (.39)	177 (ns)	156 (.38)	159 (.37)	156 (.38)	141 (.43)	154 (.39)	152 (.39)

Note. Abbreviations for participant groups in post-hoc comparisons represent sexual offenders (SO), offending controls (OC), and non-offending controls (NOC). *U* is the test statistics from the Mann-Whitney post-hoc tests and *r* indicates the effect size for the comparison. Word Category abbreviations represent the emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories.

Unfortunately, the analysis of the Stroop bias scores for this word stimulus set did not yield any significant differences between the groups. This result is not surprising given the weak results from Chapters 3 and 4. However, Figures 7, 8 and 9 are provided to display the response patterns between the groups for the separate word categories. Figure 7 shows that the adolescent groups did not exhibit salience to the EPD emotional/personality descriptors, while the adult offending groups did elicit response biases to these word stimuli.

Figure 7. *EPD Stroop Bias Scores by Participant Group*

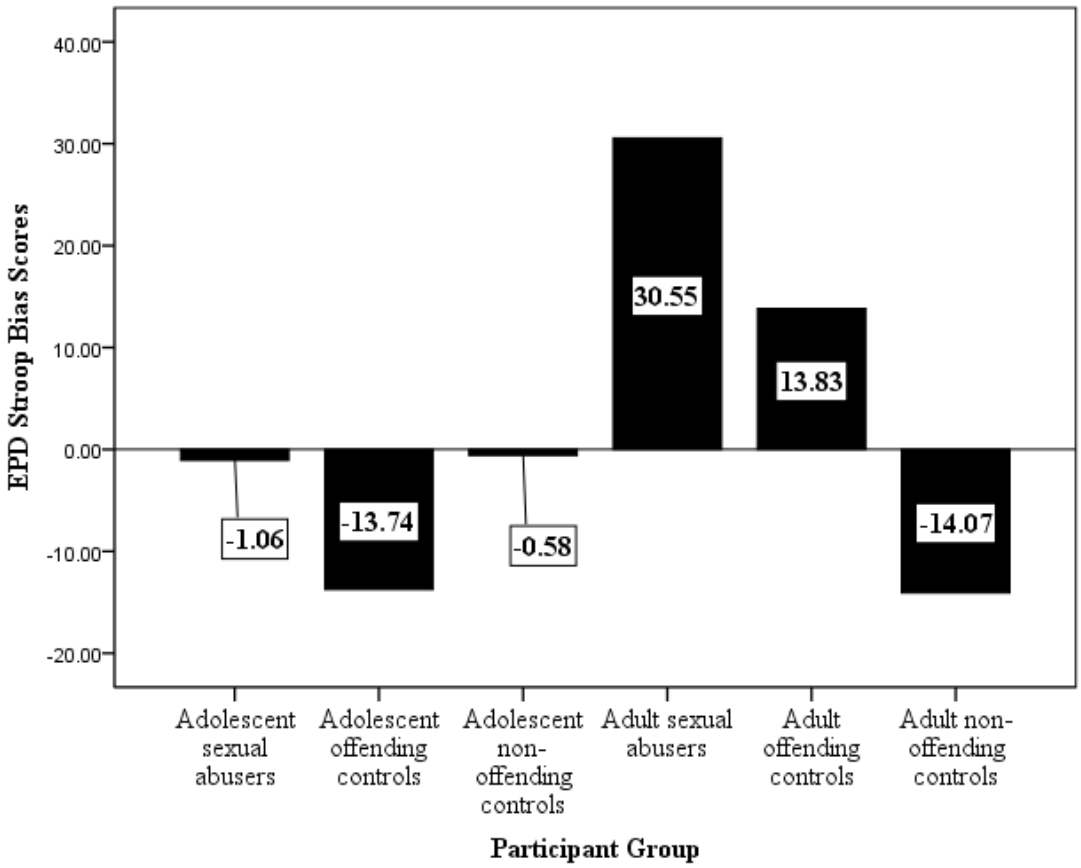
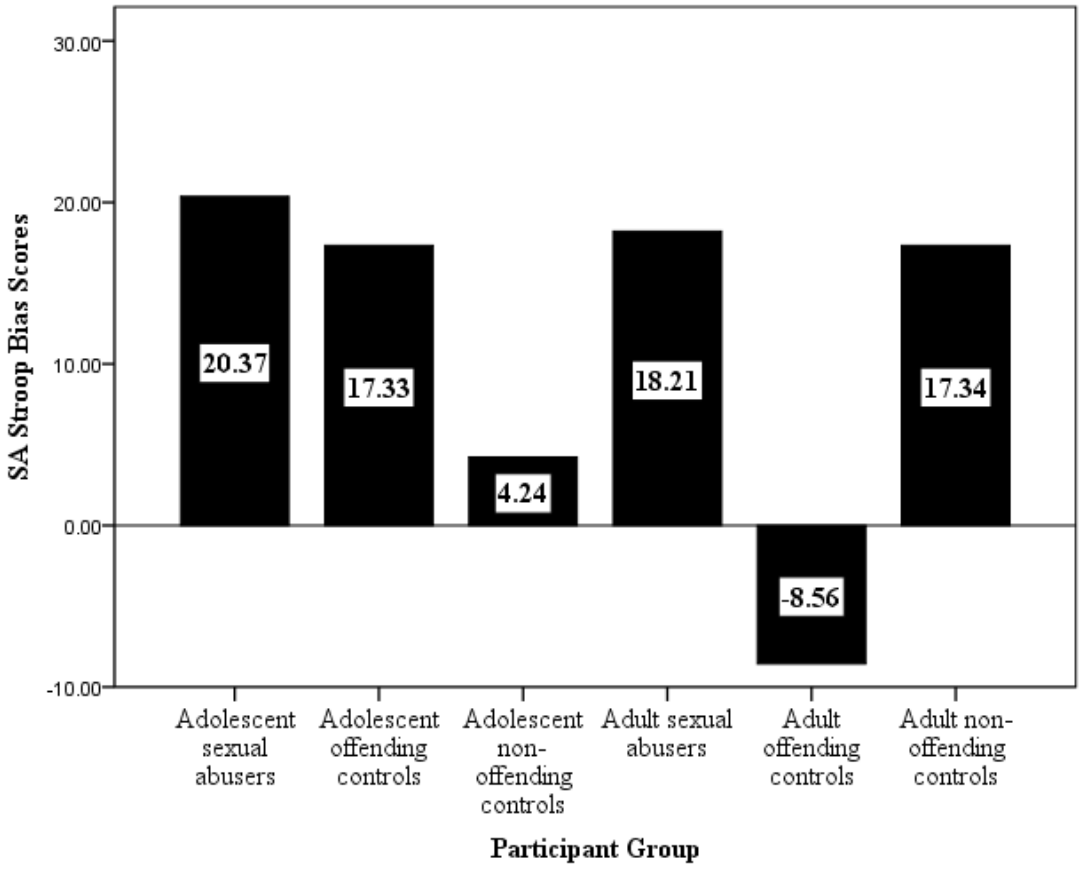


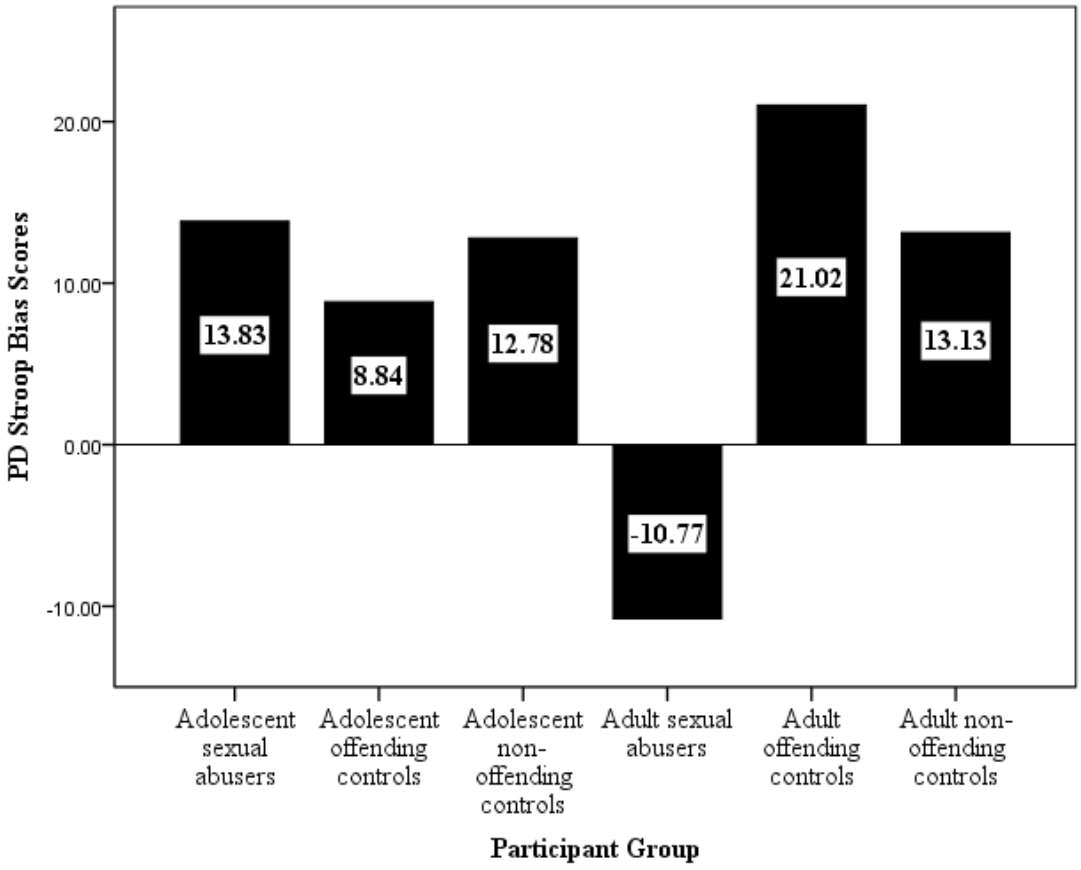
Figure 8 displays interference in information-processing of the sexual action words from this word stimulus set between the individual groups. These word stimuli were salient to all of the participant groups, with the exception of the adult offending control subjects.

Figure 8. *SA Stroop Bias Scores by Participant Group*



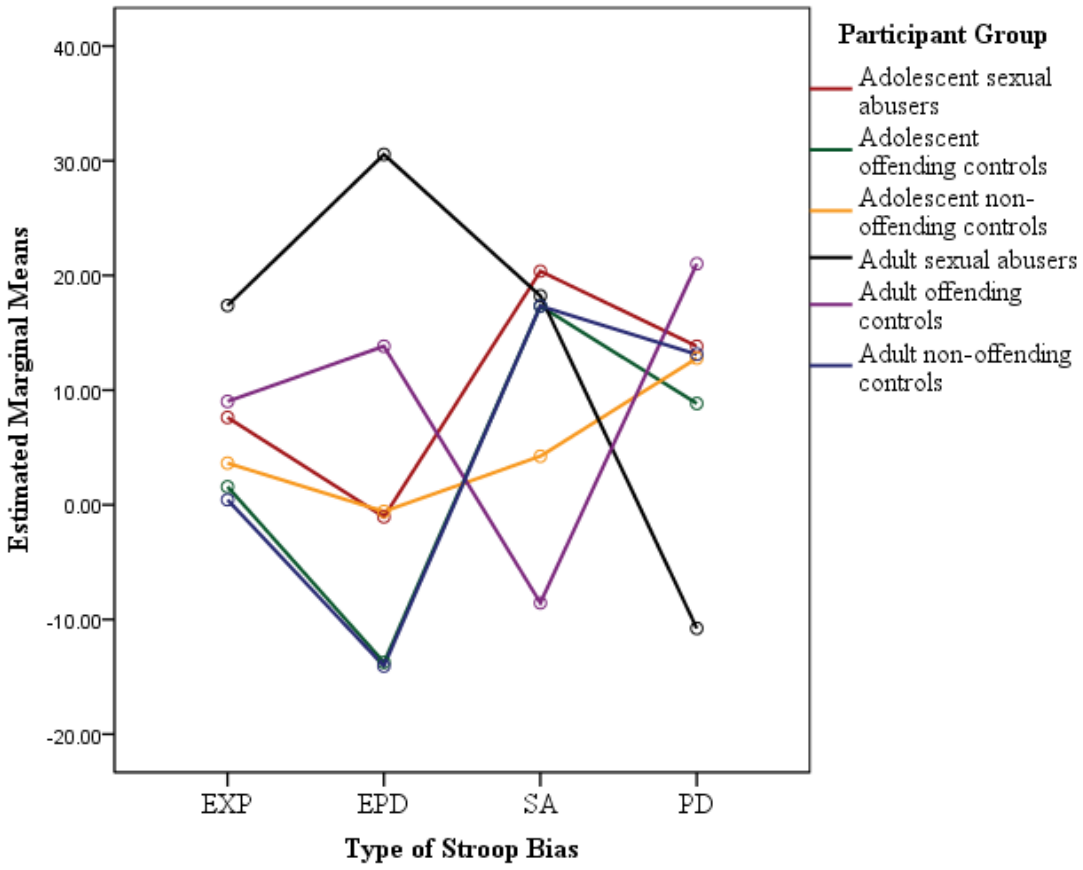
Finally, all participant groups, barring the adult sexual abusers, experienced interference in information-processing for the words depicting physical descriptors of potential victims (see Figure 9).

Figure 9. *PD Stroop Bias Scores by Participant Group*



All main effects from the two-way mixed ANOVA were non-significant: type of Stroop bias, $F(1.88, 266.61) < 1$, ns; the interaction, $F(9.39, 266.61) < 1$, ns; and participant group, $F(5, 142) < 1$, ns. Additionally, all of the F-ratios were less than 1 indicating that there was more error than variance created by the experiment. Figure 10 displays the means of the main effects of the two-way mixed ANOVA.

Figure 10. *Means for the Main Effects of Type of Stroop Bias and Participant Group: Newly-Derived Word Stimuli*



Discussion

This study provides empirical evidence of the similarities and differences in emotional Stroop response patterns between adolescent and adult offending and control participants towards emotional, offence-related and sexual interest word stimuli. With respect to similarities between the two sexual abuser groups, both groups were consistently slowest to respond (followed by the violent offender groups) in mean RT's to the separate word categories for both word stimulus sets. This finding is consistent with the findings from Price and Hanson (2007). However, Smith and Waterman (2004) found that their violent offenders were consistently slowest to respond to the task, with the sexual abuser sample following closely behind their responses. Smith (2009) notes that this tendency for offender samples to take longer to respond could be due to the presence of an interaction between the effects of the stimulus salience and possible neurological impairments such as impulsivity. Additionally, it has been recommended that this task not be administered to individuals diagnosed with learning disabilities (Smith, 2009). The presence of diagnosed learning disabilities was not obtained for the studies comprised in this thesis. This was due to restrictions or limited access in obtaining reliable information that could confirm diagnoses of learning disabilities. Nevertheless, there were few significant differences in mean RT's between the offender groups in this study, and they consistently responded more slowly across all word categories. Therefore, it would appear that the offender groups in this study were experiencing a general slowing effect that was caused by a variable that was not measured or controlled for in this study.

Both sexual abuser groups experienced greater Stroop interference effects for the sexual word stimuli compared to the other groups. This result was significant

between the sexual abusers and the non-offending control groups. Importantly, the adolescent sexual abusers yielded negative Stroop bias scores (i.e., responded more slowly to neutral word stimuli than emotional word stimuli) for the remaining word categories that contained emotional content. These findings combined would suggest that the groups are in fact displaying effects that are specific to an element of the sexual word stimuli from the Smith and Waterman (2004) stimulus set.

The importance of appropriate word selection and the inability of word stimulus sets to be able to differentiate between offender groups thus far have been highlighted in the research (Price & Hanson, 2007; Smith, 2009; Smith & Waterman, 2004). The offending controls in this study also exhibited Stroop bias for the sexual words. Additionally, the sexual abusers displayed a response bias towards the aggressive word stimuli. A possible explanation for these findings relates to the nature of the sexual word stimuli (Smith, 2009). For example, as previously mentioned in Chapters 3 and 4, the words from this stimulus set do not denote a specific deviance or sexual interest. Rather, they are fairly general in nature and some of the words are quite aggressive in content as well (i.e., bitch, slap, force, rape, slut). Therefore, the offending controls could be eliciting Stroop bias towards the words that are more violent nature within the sexual word category. It is just as likely that some of the adult sexual abusers have a preference for aggressive sexual behaviours. Therefore, it is possible that instead of measuring sexual interest specifically, the contents of this word category are more reflective of sexual preoccupation or a preference for sexual material in general for the sexual abusers and aggressive word content for the offending controls.

The adult sexual abusers continued to display larger Stroop interference for negatively-themed stimuli when compared to the adolescent groups. It remains

unclear why the negative word stimuli were so salient to this group. It has been recommended that mood questionnaires be administered and that diagnoses of mental disorders be measured and controlled for when conducting emotional Stroop tasks with offender samples (Smith, 2009). Had this recommendation been feasible for this study, it would have been possible to investigate whether the scores on the mood questionnaires were related to the outcomes of the positive and negative Stroop bias scores. This practice could have also been beneficial in controlling for the potential effects that the experimental setting (i.e., prison) could have on the offender samples.

If we look at the overall Stroop patterns across the groups for the Smith and Waterman word categories, the adult offender groups appear to be the only groups eliciting any type of response for the emotional word stimuli. Barring the adolescent sexual abusers exhibiting sexual Stroop bias, all of the adolescent groups displayed negative Stroop effects across the emotional word categories, indicating that the word stimuli were not salient to this age cohort or that there was an inability to process the word stimuli at all. This could be due to developmental factors that should be considered in the assessment of adolescent samples. For example, their limited life experience alone could explain why they do not respond to these stimuli. It is possible that they have not developed cognitive or affective associations pertaining to the stimuli presented (Grant, 2006). Additionally, adolescents may not have had the time yet to develop or learn what is of sexual interest to them. Even so, it is clear from the results from all of the adolescent groups that the word stimuli used in this study are not appropriate for these individuals to assess sexual interest.

Although the results from the newly-derived word stimulus set were not statistically significant it seems key to note some of the observations in response patterns. For example, all of the adolescent groups and the non-offending adult

controls displayed a negative Stroop effect towards the emotional/personality descriptors of potential targets of victimisation. In contrast, all of the groups, except for the adult sexual abusers, yielded a response to the physical descriptors. Instead, the adult sexual abusers displayed negative Stroop bias for this word category. This finding is crucial because the EPD word category is comprised of the main words that were specifically derived from the accounts of adult sexual abusers and individuals working with adult sexual offenders to describe who would be most at risk from them. Therefore, this would appear to provide further confirmation that this word category reflects sexual interests more specific to adult sexual abusers.

Finally, most of the groups exhibited a Stroop bias towards the sexual action word stimuli. However, the offending (violent) control subjects displayed negative Stroop effects for the sexual action words. Interestingly, this finding is contrary to the interference caused to the offending controls for the sexual word stimuli from the Smith and Waterman word set. If we review the nature of the words comprised in this word category they lack violent connotations (i.e., playing, pretending, sucking, stroking). Therefore, it is not surprising that the words from this category did not elicit a response from the offending controls when we consider both the type of crimes that they have committed and the evidence from this study that the adult offending groups are experiencing Stroop effects to offence-related word stimuli.

Conclusions

From the separate analysis of the adolescent and adult samples in Chapters 3 and 4 we were able to observe clear differences in response patterns to the emotional Stroop task between the two age cohorts. The formal comparison of these separate groups has confirmed that sexual abusers respond quite differently to the task from non-offending

control subjects. Additionally, there is evidence that adolescent groups as a whole differ quite significantly in the processing of emotional and sexual word stimuli compared to adult samples for these word stimulus sets. This study provides further proof that additional considerations should be made in assessment efforts of adolescent samples. Concomitantly, it is important not to generalise assessment methodologies across adolescent and adult sexual offender groups, but to treat them differently in the assessment of sexual interests. The response patterns of the adults lead us to believe that this tool may be used to assess the salience of sexual abusers' cognitions towards sexual word stimuli, and more specific sexual interests of this subgroup. In contrast, the response patterns of the adolescent samples to this themed emotional Stroop task would suggest that another variant of the task may be more appropriate to measure specific sexual interests of this age cohort.

CHAPTER 6

An Emotional Stroop Task Between Offenders and Controls and the Identification of Areas of Brain Activation Through the Use of Functional Magnetic Resonance Imaging (fMRI).

This study aimed to create a better understanding of the cognitive processes of sexual offenders, and how they differ from control subjects in processing sexual material using an emotional Stroop task paradigm. This study links behavioural findings of the emotional Stroop task (i.e., differences in emotional Stroop interference effects in adult samples) with brain activation patterns measured through functional magnetic resonance imaging (fMRI).

fMRI was used to identify areas of brain activation that are present during the presentation of an emotional Stroop task. The study discusses the fMRI results from control subjects using the Smith and Waterman (2004) word stimulus set and uses the results from the control subjects to form a baseline of brain activation areas present during the completion of this specific task. The study then compares the brain activation patterns of the control subjects to the results from a group of sexual abusers in order to identify differences in brain activation patterns between the groups. Difficulties in the interpretation of results due to similarities in response activation patterns between cognitive and affective tasks are identified and discussed. The implications of the findings are discussed and future research considerations in this area suggested.

Note. The following article has been submitted to Archives of Sexual Behavior for review and is authored by Shelley Price, Anthony Beech, Harriet Allen, and Ian Mitchell. The following article was also awarded the Graduate Research Award from the Association for the Treatment of Sexual Abusers (ATSA) for the 29th Annual Research and Treatment Conference (2010, October). This research was supported in part by the National Organisation for the Treatment of Abusers (NOTA) and the School of Psychology, The University of Birmingham, Birmingham, West Midlands, United Kingdom. Some of the background content overlaps with previous chapters due to the need to include previous material in the manuscript submitted.

Editorial Manager(tm) for Archives of Sexual Behavior
Manuscript Draft

Manuscript Number:

Title: An Emotional Stroop Task Between Offenders and Controls and the Identification of Areas of Brain Activation Through the Use of Functional Magnetic Resonance Imaging (fMRI).

Article Type: Article

Keywords: Emotional Stroop task; fMRI; sexual offending; sexual interest

Corresponding Author: Shelley A. Price, M.A.

Corresponding Author's Institution: The University of Birmingham

First Author: Shelley A. Price, M.A.

Order of Authors: Shelley A. Price, M.A.;Anthony R. Beech, Prof;Harriet A. Allen, PhD;Ian Mitchell, PhD

Abstract: This study aimed to create a better understanding of the cognitive processes of sexual offenders, and how they differ from control subjects in processing sexual material using an emotional Stroop task paradigm. The current study examined differences in emotional Stroop interference effects between adult sexual offenders, adult violent offenders and non-offending adult controls. A second experiment was conducted using functional magnetic resonance imaging (fMRI) to identify differences in brain activation patterns between adult male sexual abusers and controls during the completion of an emotional Stroop task. Results are consistent with previous findings with sexual abusers experiencing the greatest Stroop interference effects for sexual word stimuli compared to controls. fMRI results indicate different areas of activation for emotional word content compared to neutral stimuli in control subjects. Differences in brain activation patterns between controls and sexual abusers are evident in areas of the brain that have been linked with attention and emotion, including the prefrontal cortex, precentral cortex, anterior cingulate cortex, occipital cortex and temporal lobe. Implications of the findings are discussed and future considerations for research in this area are suggested.

An Emotional Stroop Task Between Offenders and Controls and the Identification of Areas of Brain Activation Through the Use of Functional Magnetic Resonance Imaging (fMRI).

Shelley A. Price, Anthony R. Beech, Harriet A. Allen, and Ian Mitchell

The University of Birmingham

Birmingham, United Kingdom

Author Note:

This research was supported in part by the National Organisation for the Treatment of Abusers (NOTA) and the School of Psychology, The University of Birmingham, Birmingham, West Midlands, United Kingdom.

Corresponding Author:
Shelley A. Price, School of Psychology, The University of Birmingham, Edgbaston, U.K., B15 2TT
Email: s

Abstract

This study aimed to create a better understanding of the cognitive processes of sexual offenders, and how they differ from control subjects in processing sexual material using an emotional Stroop task paradigm. The current study examined differences in emotional Stroop interference effects between adult sexual offenders, adult violent offenders and non-offending adult controls. A second experiment was conducted using functional magnetic resonance imaging (fMRI) to identify differences in brain activation patterns between adult male sexual abusers and controls during the completion of an emotional Stroop task. Results are consistent with previous findings with sexual abusers experiencing the greatest Stroop interference effects for sexual word stimuli compared to controls. fMRI results indicate different areas of activation for emotional word content compared to neutral stimuli in control subjects. Differences in brain activation patterns between controls and sexual abusers are evident in areas of the brain that have been linked with attention and emotion, including the prefrontal cortex, precentral cortex, anterior cingulate cortex, occipital cortex and temporal lobe. Implications of the findings are discussed and future considerations for research in this area are suggested.

Keywords: Emotional Stroop task, fMRI, sexual offending, sexual interest

Introduction

Deviant sexual interest is a factor that is considered essential to include in actuarial risk measures used in the assessment of sexual offenders (Kalmus & Beech, 2005). The measurement of these interests is difficult because they are implicit in nature. Traditionally, sexual interest has been measured using the penile plethysmograph (PPG) however, difficulties concerning the administration of the test (i.e., cost and ethical considerations) and questions surrounding its reliability and validity for measuring sexual interest have been raised (Kalmus & Beech, 2005; Laws, 2009; Marshall & Fernandez, 2003).

A more recent focus on the assessment of sexual interest at an information-processing level has resulted in a body of literature of possible cognitive approaches for the assessment of sexual interest. Researchers, for example, have used: the Implicit Association Test (IAT; Gray et al., 2005; Nunes, Firestone & Baldwin, 2007); the Rapid Serial Visual Presentation (RSVP) test (Beech et al., 2008); and the emotional Stroop task (Price & Hanson, 2007; Smith & Waterman, 2004). However, there is question regarding whether sexual interest can be measured using these approaches. It is thought that the RSVP is an attentional method of measuring sexual interest; that the IAT may be measuring attitudes; and that the emotional Stroop task may be measuring social-affective functioning (Flak, Beech & Fisher, 2006), however it has also been described as an attentional method and as an information-processing measure of sexual interest with both cognitive and affective components (Smith, 2009). The predictive utility of these cognitive measures on sexual offending behavior is still a question that needs to be examined, therefore, attempts to solidify the research surrounding available cognitive measures as potential alternatives to physiological measures of sexual interest is valuable to the field (Ó Ciardha & Gormley, 2009).

Previous emotional Stroop studies encompassing a broad range of psychological research have consistently shown that subjects take longer to respond to individual word

stimuli if they hold greater cognitive or emotional association with the stimuli presented (e.g., Egloff & Schmukle, 2004; Foa, Feske, Murdock, Kozek & McCarthy, 1991; Gross Jarvik & Rosenblatt, 1993; Holle, Neely & Heimberg, 1997; Kramer & Goldman, 2003; Smith & Waterman, 2003, 2004, 2005; Waters & Feyerabend, 2000; Wertz & Sayette, 2001). It is this strength in findings of previous research in psychology that has fuelled the use of this tool to examine deviant sexual interests of sexual abusers.

Two studies thus far have been conducted using the emotional Stroop paradigm with sexual abusers in an attempt to measure level of deviant sexual interests. Smith and Waterman (2004) conducted an emotional Stroop task using aggressive and sexually-themed word stimuli with a sample of sexual offenders, violent offenders, non-violent offenders and undergraduates. They reported that sexual offenders and violent offenders were significantly slower than undergraduates to color-name sexual word stimuli, and that violent offenders and violent sexual offenders took significantly longer to color-name aggression-themed words than undergraduates.

In a replication of the Smith and Waterman study, Price and Hanson (2007) examined emotional Stroop interference effects of sexual and aggressive word stimuli between rapists, child molesters, violent offenders, non-violent non-sexual offenders and community controls. Price and Hanson (2007) found similar results to Smith and Waterman (2004), in that rapists were significantly slower than community controls to color-name sexually-themed words. However, the significant results for the aggression-themed words were not duplicated. Despite the lack of significance in the findings of this replication study, similar response latency patterns were displayed by the offender groups between the two studies. The offender groups typically took longer to respond to target emotional/offending-related words than did the controls (Price & Hanson, 2007).

Although the findings of the two studies using sexual offender populations has yielded significant results, these results have not yet been able to differentiate between different offender groups (i.e., between sexual offenders and violent offenders or between pedophiles and rapists). However, since the word list (see the Appendix) used in these studies has been able to elicit significant differences between offenders and controls thus far, it is clear that there are differences in information-processing between sexual offenders and non-offenders (Price & Hanson, 2007). It is then worth further examination to attempt to identify what mechanisms are responsible for these significant differences between offenders and controls using this particular paradigm.

Functional Magnetic Resonance Imaging

It is through functional MRI (fMRI) techniques that we are able to identify which brain regions are most active while completing specified tasks. fMRI research using the traditional color-word Stroop task has identified several brain regions thought to be relevant to interference effects experienced during the task (Bush et al., 1998; Bench et al., 1993; Carter, Mintun, & Cohen, 1995; Leung, et al., 2000; Pardo et al., 1990; Taylor et al., 1994). There is some debate surrounding the structures most responsible for resolving color-Stroop interference. Leung et al. (2000) found the anterior cingulate cortex (ACC), insula, inferior frontal and middle frontal (prefrontal cortex; PFC), parietal, and mid-temporal regions to be most activated during incongruent trials (i.e., color word and ink color do not match). Bench et al. (1993) and Taylor et al. (1994) attribute Stroop interference to activations in the inferior frontal gyrus (IFG) and basal ganglia, while Pardo et al. (1990) believe the ACC to be the most relevant region. Finally, Veen and Carter (2005) have found differences in activation patterns between congruent and incongruent stimuli in the ACC, PFC and posterior parietal cortex (PPC). Because the traditional Stroop task requires a participant to selectively attend to

color information and suppress word information, the brain regions listed above that are involved in processing are thought to be part of the brain's attentional mechanisms (Compton, et al., 2003; Corbetta & Shulman, 2002; Van Veen, & Carter, 2005)

A limited number of studies have been conducted using variants of an emotional Stroop task and fMRI, although similar brain structures have been found to be involved with this variant of the task. For example, Ersche et al. (2010) used a drug-word emotional Stroop task with a sample of stimulant-dependent individuals compared to healthy controls. The authors identified various parts of the brain activated during drug-word stimulus presentation which included parts of the PFC and precentral cortex (PCC), parietal cortex, inferior temporal cortex, left thalamus and caudate nucleus, ACC and posterior cingulate. Again, these are areas that have been implicated in attention control networks (Corbetta & Shulman, 2002). Additionally, Ersche et al. (2010) found that attentional bias for drug stimuli was associated with abnormal activation in the IFG. Abnormal IFG activation has been identified in disorder-related Stroop studies using patient samples with diagnoses of post-traumatic stress disorder, obsessive-compulsive disorder, bipolar disorder and panic disorders (Malhi, et al., 2005; Shin, et al., 2001; van den Heuvel, et al., 2005).

Wingenfeld, et al. (2009) conducted an individual emotional Stroop task study using negative-themed words compared between patients with borderline personality disorder (BPD) and healthy controls. Wingenfeld, et al. (2009) compared activation to neutral words, negative emotion words and negative emotion words individualized to their patients. The authors found that healthy control subjects exhibited increased activation in the left ACC, right medial frontal gyrus and right precentral gyrus, compared to BPD patients during the presentation of negative emotion words versus neutral words. Furthermore, in the comparison of individualized negative words and neutral words, increased activation in controls was exhibited bilaterally in the ACC, right frontal lobe areas, left medial temporal gyrus, right

hippocampus and left cuneus. The superior temporal gyrus and medial frontal gyrus were implicated with the emotional and attentional control networks of the brain, while the ACC and pre-post central gyrus were involved in target processing, and the IFG with memory. Therefore, Wingenfeld et al. (2009) concluded that the patients with BPD in their study failed to show a fronto-limbic pattern of activation as their controls did with increased activation in the ACC (Wingenfeld et al., 2009).

The Current Study

The current study attempted to expand on previous research (see Price & Hanson, 2007; Smith & Waterman, 2004) conducted using the emotional Stroop task as a measure used to explore the psychological nature of deviant sexual interests with offender populations. The present study used the same word list used by Smith and Waterman (2004) and Price and Hanson (2007) in two separate emotional Stroop experiments. The first experiment examined emotional Stroop interference effects of emotional words between sexual abusers, violent offenders and controls using a button-press version of the task.

Due to the unclear nature of what the emotional Stroop task is really measuring, a second experiment was conducted using fMRI. The purpose of this experiment was to examine emotional Stroop interference for the Smith and Waterman word stimulus set at a more explicit information-processing level. It was the aim of this part of the study to be able to identify specific brain structures involved in the processing of this particular word stimulus set, in hopes of gaining a better understanding of why these words have elicited significant differences between offenders and controls. It was thought that having the ability to examine the cognitive mechanisms behind the emotional Stroop could bring us closer to answering the question of whether this particular emotional Stroop task is able to measure deviant sexual interest.

Experiment 1

Based on the results from Smith and Waterman (2004) and Price and Hanson (2007) it was hypothesized that the offenders would demonstrate slower reaction times overall on the emotional Stroop task. Additionally, it was anticipated that sexual abusers would elicit response bias to sexually-themed stimuli compared to other groups in the study, and that violent offenders would demonstrate a response bias to aggressive word stimuli.

Method

Participants. Twenty-seven men convicted of sexual offences, 21 men convicted of violent offences and 35 non-offending control participants took part in the study. The sample of sexual offenders included 4 incest offenders, 9 pedophiles, 7 rapists, and 2 mixed sexual abusers (had offended against both adult and child victims). Eight of the sexual abusers were on community supervision and recruited through West Midlands Probation, UK. Nineteen sexual abusers and the sample of violent offenders were recruited through poster recruitment and research representatives of separate wings at HMP Grendon (Oxfordshire, UK). The non-offending controls were recruited through the University of Birmingham research scheme whereby psychology undergraduates, masters and PhD students complete 10 hours of research credits per year to fulfill degree requirements. Sexual abusers had a mean age of 43.31 years (SD=9.82), violent offenders a mean age of 35.24 years (SD=10.06) and non-offending controls a mean age of 20.49 (SD=2.51). A significant difference in participant age, $F(2, 79) = 69.35$, $p < .001$, was found between all groups.

Apparatus and materials. Information concerning participants' age, presence of a learning disability, handedness, and whether they had an existing criminal record was

collected. Information regarding offenders' index offense, prior offences, victim age, victim gender and relationship to the offender was collected in order to aid in the sub-categorization of offender groups.

A computerized version of the emotional Stroop task was presented on a Toshiba laptop on a 12"x9" screen. Word stimuli were presented using version 2.0 of E-Prime software (Psychology Software Tools, Inc.). The color-identification response latencies for each trial were detected and recorded by a five-button serial response box (Psychology Software Tools, Inc., model 200A) with four task-specific colored buttons identified (green, red, blue and white). This study used button-press response recordings due to the noisy environments the research was often carried out in (i.e., prisons, group treatment settings), and to maintain consistency across the two experiments. It was necessary to use button-press response recordings in Experiment 2 because participants completed the task within an MRI machine and button-press methods reduce participants' movement within the scanner, minimizing the chance of noise in the data. It was of interest to the authors to examine whether similar response latency patterns between offender and control groups would be evident, despite the change in response recording methodology (i.e., the previous two studies used voice-activated response recordings).

In addition to the emotional Stroop task, participants completed the British Picture Vocabulary Scale (BPVS-II; Dunn, Dunn, Whetton, & Burley, 1997); the Hayling and Brixton tests of executive function (Burgess & Shallice, 1997); and a shortened version of the Beliefs About Children Scale (BACS; Beckett; 1987). Scoring items on the BACS yields 2 subscale scores: cognitive distortions (CD) and emotional congruence with children (EC).

The BPVS was used to confirm that participants had an appropriate level of understanding of the words that were presented in the emotional Stroop task. The Hayling Sentence Completion test is a measure of response initiation and response suppression. It

consists of two sets of 15 sentences each missing the last word. In the first section the examiner reads each sentence aloud and the participant has to simply complete the sentences, yielding a simple measure of response initiation speed. The second part of the Hayling requires subjects to complete a sentence with a nonsense ending word (and suppress a sensible one), giving measures of response suppression ability and thinking time. The Brixton test is a visuospatial sequencing task that measures the ability to detect rules in sequences of stimuli. It takes between five and ten minutes to administer, and yields an easily understood scaled score between 1 and 10 (Burgess & Shallice, 1997).

Procedures. Participants were provided with an information sheet that outlined the procedures of the experiment and were then asked to sign consent forms agreeing to take part in the study. Participants were required to learn color-mappings of the response buttons prior to the presentation of target words. Participants completed four practice blocks (25 words per block) of neutral word stimuli in order to memorize the color response mappings on the serial response box. Participants then completed the emotional Stroop task. Word stimuli were presented randomly and prior to each word a fixation "x" appeared at the centre of the screen for 500ms. Participants were asked to press the button corresponding to the color of ink in which the word was presented and to ignore the semantic meaning of the word. No opportunity to correct mistakes was provided because once a response was provided the next word stimulus was prompted. Following the emotional Stroop task, participants completed the BPVS, Hayling and Brixton tests, and shortened BACS questionnaire.

Data analysis. The data was explored for outliers using boxplots and this resulted in the removal of 3 non-offending control cases in the analysis. All three cases had presented as outliers on all mean reaction times (RT) and tests of executive function.

Pre-analysis was completed on the mean scores on the tests of executive function, vocabulary, and errors made on the Stroop task and were entered into separate one-way ANOVAs with participant group as the between-groups variable. Sexual abusers yielded significantly lower scores on the Hayling test of executive function than the non-offending controls $F(2, 78) = 5.75$, $p < .05$, and significantly lower scores on the Brixton test of executive function than both the violent offenders and non-offending controls $F(2, 77) = 9.53$, $p < .001$. It is important to note that the sexual abusers were merely one scaled score lower than the other groups on the Hayling, rendering them a categorization of "moderate average" and the other groups being categorized as "average". On the Brixton test, the sexual abusers were categorized as "average" and the other two groups were categorized as "good", 2 categories above average. Interestingly, there were no significant differences evident between groups for scores on the subscales of the BPVS, meaning that their level of vocabulary was adequate across all groups. Finally, a significant difference was found in number of errors made on the Stroop task, $F(2, 80) = 4.137$, $p < .05$, between violent offenders and controls with control subjects making the most mistakes on the task (see Table 1 for summary).

Approximate location of Table 1: Scale responses and Stroop errors per group with (Standard Deviations).

Emotional Stroop bias scores were calculated by subtracting the mean RT of neutral words from the mean scores of target words. This resulted in 5 bias scores to be compared across the three groups: positive, neutral, color, aggression, and sexual. Mean RT and emotional Stroop bias scores were entered into separate one-way ANOVAs with participant group as the between-groups variable.

In addition to testing for significant differences in Stroop effects between the groups it was of interest to test whether age presented as a covariate and have an influence on the results. To control for age analysis of covariance (ANCOVA) was carried out on the dependent variables (i.e., mean RT's and Stroop bias scores).

Ethical approval. The British Psychological Society code of ethical practice was adhered to in the design of the research projects. Ethical approval was gained from The School of Psychology Ethics Committee at the University of Birmingham, West Midlands Probation, and the Grendon Research Advisory Group (GRAG).

Results

Table 2 lists the mean RTs (with SDs) and mean emotional Stroop bias scores per group from Experiment 1. Positive bias scores indicate that participants took longer to respond to the task for the target emotional word category (i.e. emotional Stroop interference was present for that word category). Significant differences were evident for the mean RT between groups for all word categories.

Approximate location for Table 2: Mean response times (RT) and bias scores with (Standard Deviations) for the emotional Stroop task.

Post-hoc analysis showed that for the neutral mean RTs, sexual abusers were significantly slower to color-name than the non-offending controls, and non-offending controls were significantly faster at color-naming neutral words than both offending groups, $F(2, 80) = 15.87$, $p < .001$. All three groups significantly differed in their mean RTs to both positive, $F(2, 80) = 22.91$, $p < .001$, and negative emotion words, $F(2, 80) = 25.81$, $p < .001$,

where sexual abusers were slowest to respond to the task, and the non-offending controls displayed the fastest RTs. For the mean color RTs, sexual abusers significantly differed from the non-offending controls, and the non-offending controls were significantly quicker to respond than both groups $F(2, 80) = 17.77$, $p < .001$. For mean RTs to the aggression-themed words, the sexual abusers took significantly longer to respond than the non-offending controls, and the non-offending controls were significantly faster to respond than both offending groups, $F(2, 80) = 18.97$, $p < .001$. Finally, significant differences were evident for mean RTs on the sexual word stimuli between all three groups, $F(2, 80) = 22.86$, $p < .001$, with sexual abusers taking the longest to respond to the word stimuli, followed by the violent offenders, then the non-offending controls.

For the emotional Stroop bias scores significant differences were found for the positive and negative emotion bias scores and the sexual bias scores. For positive emotional Stroop bias scores sexual abusers exhibited a significantly larger effect than non-offending controls, $F(2, 80) = 4.32$, $p < .05$. All three groups significantly differed from each other on the negative emotional Stroop bias scores, $F(2, 80) = 6.71$, $p < .005$, with sexual abusers displaying the largest bias score, followed by the violent offenders, and non-offending controls. Finally, sexual abusers significantly differed to non-offending controls on sexual bias scores, $F(2, 80) = 3.883$, $p < .05$. No significant differences were found between the offender groups on sexual bias scores. Finally, the results from the analysis of covariance did not yield a significant relationship between the Stroop bias effects and age.

Discussion

The significant differences found in Experiment 1 are somewhat consistent with the findings from Smith and Waterman (2004) and Price and Hanson (2007). A clear difference between the findings is that the non-offending controls were significantly quicker to respond

across all word categories when compared to the offender samples. Importantly, since response latency differences were evident across all word categories between participants, bias scores may represent a more pure reflection of the emotional Stroop bias that participant groups experienced with the added emotional content of the experimental stimuli.

It is important to note that various groups of sexual offenders have displayed difficulties with behavioral inhibition on tests of executive function (Dolan, Millington, & Park, 2002; Langevin et al., 1989; Ponseti et al., 2001; Stone & Thompson, 2001; Valliant et al., 2000). These difficulties have been linked with frontal cortex dysfunction (Graber et al., 1982) and temporal lobe structures (Graber et al., 1982; Hucker et al., 1986; Kolárský, Freund, Machek, & Polák, 1967). Therefore, although the sexual abuser group was of a heterogeneous nature and was analyzed as one experimental group, there is evidence that there are commonalities in cognitive processing for different types of sexual abusers.

Some may want to attribute the significant differences in mean RTs to the age of the participants considering the significant differences in age between the groups. However, previous research using the traditional color-word Stroop task suggests that the effects of age on and individual's ability to successfully complete the Stroop task are not evident until an individual is over the age of 60 years (Mutter, Naylor, & Patterson, 2005). Johnson, Bouchard, Segal, Keyes, and Samuels (2003) in fact report an increased ability on the Stroop task until the age of 60 when this phenomenon begins to reverse. Verhaeghen and Meersman (1998) attribute a general slowing effect to be responsible for any perceived effects of age on the Stroop task and report that Stroop interference is not age sensitive. Additionally, age was not found to be a covariate in this study.

Two studies have been conducted using emotional Stroop paradigms examining the effects of age. Wurm, Labouvie-Vief, Aycock, Rebucal, and Koch (2004) found that younger participants responded significantly faster to all stimuli compared to older adults in an

emotional Stroop task with words rated as high, medium and low in level of arousal. Wurm et al., (2004) also found that arousal value of words was related to level of Stroop interference in older participants, but not in younger participants. Ashley and Swick (2009) examined age effects towards negative word stimuli, and found that both younger and older participants were slower to color name negatively-themed color words, however the younger participants exhibited a 'persistent slowing' effect, explained as a carry-over effect from the negatively-themed words onto other word categories.

Although the above studies were examining potential age effects in the emotional Stroop research it is important to note that the average age of the older participants in these studies was 74 years and 70 years, ages that are significantly older than any of the participants included in this study. Additional research is required to examine whether there are significant age effects involved in emotional Stroop processing across other age categories rather than at extreme ends of the age spectrum. Further discussion of these results is presented in combination with the results from Experiment 2 in the 'Summary and Concluding Discussion section'.

Experiment 2

Since this is the first fMRI study conducted using this word stimulus set comparing sexual abusers to control subjects, all differences in brain activation patterns were of interest to the authors. Based on research that would suggest that the ACC plays a role in affective processing it was anticipated that differences in activation patterns in the ACC would be evident between control subjects and sexual offenders in the fMRI portion of the study due to the emotional content of the experimental word categories (Bush, Luu, & Posner, 2000; Whalen, et al., 1998). Furthermore, it was anticipated that abnormal activations in the PFC would be expected between controls and sexual abusers based on current fMRI literature

available using emotional Stroop tasks (Bremner et al., 2004; Ersche, et al., 2010; Malhi, et al., 2005; Shin, et al., 2001; Van den Heuvel, et al., 2005; Wingenfeld, et al., 2009).

Method

Participants. Thirteen paid adult controls were recruited through a combination of participant lists, recruitment of individuals associated professionally through the University of Birmingham and poster recruitment. Data from three participants was excluded from the analysis due to technical difficulties. Three unpaid sexual abuser participants that were recruited from West Midlands Probation in the first experiment agreed to take part in Experiment 2. Two had been convicted of exhibitionism, and one of incest offences. The mean age for control participants was 42.90 years (SD=8.80), while the mean age for the sexual abusers was 47.33 (SD=10.50). The sexual offenders that did participate in this study would be considered low-risk offenders, and these types of offenders displayed the least effects in the emotional Stroop task from Experiment 1.

Unfortunately, interest to participate by the sexual abuser sample in the community setting was difficult to obtain due to the inability of the researchers to pay them to participate (based on ethical guidelines of organizations involved). This resulted in a very small sample size of sexual abusers. Participants in this experiment were matched on age and all other variables outlined in the Apparatus and materials section from Experiment 1. All participants gave written informed consent in accordance with the ethical procedures of the Birmingham University Imaging Centre (BUIC) in Birmingham, UK (available at: http://www.buic.bham.ac.uk.html). All participants had normal or corrected-to-normal vision.

Apparatus and materials. All materials used in Experiment 1 were also used in Experiment 2, however, in Experiment 2 participants lay in a 3T Philips Achieva MRI

scanner and viewed the Stroop stimuli via a tilted mirror on the eight-channel SENSE head coil with (P-reduction=2). The color-identification response latencies for each trial were detected and recorded through two five-button serial response boxes (Psychology Software Tools, Inc.) with two buttons on each response box (green, red, and blue, white) identified as responses for the purpose of this study.

A T1-weighted high resolution anatomical scan (1 mm^3) was acquired during each session. Functional data was acquired using a T2* - weighted echoplanar imaging sequence (with 32 ascending slices, a repetition time of 2000 ms, a time to echo of 35 ms, a flip angle of 85˚ and a resolution of 2.5 mm^3, restricting the field of view to the lower portion of the cortex) to measure blood oxygenation level-dependent (BOLD) signal.

Procedures. Participants were placed in the MRI scanner and completed 100 practice trials to learn the color-mappings of the response buttons prior to the presentation of target words. Participants were then presented with an emotional Stroop task (using the Smith and Waterman stimulus set) that was presented in a partially-blocked format for analysis purposes. A word category block consisted of five words from the specified word category (neutral, aggression, positive, negative, sexual and incongruent colors), and the order of word category block presentation was counterbalanced. Prior to each word a fixation 'x' appeared on the screen for 500ms and each word was presented on the screen for 2500ms, with participants providing their ink color response as quickly as possible within that time frame. After every six blocks of word presentation a 6 second break was given where participants viewed a blank screen.

Data analysis. Due to the small number of sexual abusers that took part in this study the non-parametric version of the independent samples t-tests (Mann-Whitney Test) was run

on the scale responses and number of errors made during the task. No significant differences between the two groups were evident (see Table 3). Mean RTs and Stroop bias scores between the controls and the sexual abusers were analyzed using the Mann-Whitney test.

Approximate location of Table 3: Scale responses and Stroop errors per group with (Standard Deviations) for the fMRI emotional Stroop task.

FMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following lower-level preprocessing was applied to each participant's data: motion correction using MCFLIRT (absolute mean displacements per scan averaged across participants was 0.26) (Jenkinson, Bannister, Brady & Smith, 2002); slice-timing correction; non-brain removal using BET (Smith, 2002); spatial smoothing (5mm full width at half maximum Gaussian kernel); grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 45.0s). ICA-based exploratory data analysis was carried out using MELODIC (Beckmann & Smith, 2004), in order to investigate the possible presence of unexpected artifacts or activation. No components were deleted. Each participant's functional data set was registered with their anatomical image (T1) and then transformed into Montreal Neurological Institute (MNI) space using FLIRT (Jenkinson, et al., 2001; Jenkinson & Smith, 2001).

Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). Contrasts of interest included comparing activation during neutral trials to the activation during experimental word

category trials (i.e., aggression-neutral, positive-neutral, negative-neutral, sexual-neutral and incongruent colors-neutral). Regressors describing participant motion were included as regressors of no interest. Higher-level analysis for each participant and group analysis was carried out using FLAME (FMRIB's Local Analysis of Mixed Effects) stage 1 with automatic outlier detection (Beckmann, Jenkinson & Smith, 2003; Woolrich, 2008; Woolrich, Behrens, Beckmann, Jenkinson & Smith, 2004). Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and a (corrected) cluster significance threshold of $P=0.05$ (Worsley, 2001).

Ethical approval. The British Psychological Society code of ethical practice was adhered to in the design of the research projects. Ethical approval was gained from The School of Psychology Ethics Committee at the University of Birmingham and the Birmingham University Imaging Centre (BUIC).

Results

Mean RTs and bias scores for the emotional Stroop task between the controls and sexual abusers are displayed in Table 4. No significant differences were evident for this analysis. The lack of significant differences is likely attributable to the small sample size of the samples used in this study.

Approximate location of Table 4: Mean response times (RT) and bias scores with (Standard Deviations) for the fMRI emotional Stroop task.

Averaged imaging results for the control group are listed in Table 5 to provide a summary of the brain structures that were most activated during each word category of the

Smith and Waterman word stimulus set (*x, y, z* coordinates of location provided). These results may be considered a baseline of the brain structures involved in processing this particular version of the emotional Stroop task.

Approximate location of Table 5: Areas of activation for control subject per word category.

Comparisons made on the contrasts of interest (i.e., aggression – neutral, sexual – neutral, etc.) for control subjects elicited significant clusters of activation for the aggression, positive, sexual and color contrasts thus reflecting added activation caused by the added emotional content of the target words compared to neutral words (see Table 6 and Figure 1). Areas of peak activation for these contrasts were in the occipital lobe, PCC, PFC, and temporal lobe. No significant clusters for the negative – neutral contrast resulted.

Approximate location of Table 6: Areas of Significant Clusters of Experimental Target Word Contrasts: Controls.

Approximate location of Figure 1: Activation Areas for Contrasts of Interest for Control Subjects.

Contrasts of interest were compared between the two samples in the study. No significant areas of activation were present for the sexual abusers versus controls. However, Table 7 outlines the areas of activation that were present for the control subjects between contrasts but were not activation areas for the sexual abusers. The most notable finding for this comparison is the presence of five significant peak clusters of activation exhibited by the

controls and not the sexual abusers for the sexual – neutral contrast (see Figure 2). Importantly, one of those clusters is located in the ACC in the left and right anterior division of the cingulate gyrus (0, 24, 22), an area involved in cognitive and emotional processing (Bush, et al., 2000). Significant clusters were also evident in the intracalcarine cortex (4, -76, 12); left frontal orbital cortex (-50, 22, -12); right frontal orbital cortex (30, 16, -20); and right planum polare extending to the temporal gyrus (56, 4, -16; 48, 0, 18).

Approximate location for Table 7: Areas of Significant Clusters of Experimental Target Word Contrasts for Controls > Sexual Offenders.

Approximate location for Figure 2: Significant Clusters for Contrast Sexual > Neutral Words: Controls > Sex offenders.

The significant areas of activation for the remaining contrasts of interest are displayed in Figure 3. In the aggression – neutral contrast increased areas of activation were only evident in the occipital lobe. For positive emotion and negative emotion contrasts, controls exhibited increased activation in areas linked with emotion, such as the paracingulate gyrus (0, 44, 26) and insular cortex (36, 6, 4) (Augustine, 1996). Interestingly, the color-neutral contrast showed increased activation at the temporal pole, an area involved in cognitive processing and decision making (Leung, et al., 2000).

Approximate location for Figure 3: Significant Clusters for Contrasts of Interest: Controls > Sex offenders.

Discussion

Being that this is the first fMRI emotional Stroop task conducted testing differences between sexual abusers and controls it was of interest to establish a baseline outlining the most significant areas of activation experienced per word category for control subjects. Analysis conducted on the contrasts of interest (i.e., aggression – neutral) for the control subjects resulted in significant areas of activation in the occipital, PFC, PCC, left temporal and dorsal ACC areas. These findings were also evident when the analysis was conducted comparing activation areas in controls versus sexual abusers, with control subject exhibiting increased, non-overlapping areas of activation in these regions over sexual abusers. Therefore, for the word categories eliciting emotional content significant clusters of activation appear in structures that are consistent with previous emotional Stroop studies (Bremner et al., 2004; Ersche, et al., 2010; Malhi, et al., 2005; Shin, et al., 2001; Van den Heuvel, et al., 2005; Wingenfeld, et al., 2009).

The areas of activation that are present during the contrasts of interest for control subjects independently and when compared to sexual abusers appear to be areas most associated with cognitive functions of the brain. For example, there is some indication that cognitive tasks increase activation in areas within the dorsolateral region of the PFC (i.e., parts of the middle frontal gyrus and inferior frontal gyrus) and the dorsal ACC. Whereas, emotional tasks elicit increased activation in regions within the orbitofrontal (OFC) cortex and the ventral ACC (Drevets & Raichle, 1998).

Previous research has implicated the inferior and middle frontal gyrus with the need to maintain a focus on task-relevant information when faced with more salient distracting information (i.e., emotional word content) (Banich, et al. 2000, Banich, et al., 2001; Compton, et al., 2003; Milham et al., 2002), and the potential role of the PFC in filtering task-relevant from irrelevant information (Thompson-Schill, D'Esposito, Aguirre, & Farah,

1997). Additionally, increases in brain activity in visual-processing areas of the brain (i.e., occipital lobe) are consistent with the literature (Lane et al., 1999; Lane, Reiman et al, 1997; Lang et al., 1998). Therefore, it would appear that the emotional word content and incongruent color information are activating areas involved in attentional control, and cognitive (executive function) processing.

It would appear that the dorsal parts of the ACC were more activated in control subjects when compared to sexual abusers. Activation patterns within the ACC are especially important considering evidence that the ACC is divided into two subdivisions: dorsal (cognitive) and rostral (affective). The precise location and components of these subdivisions within the ACC is unclear, but it is thought that the dorsal part of the ACC includes the anterior cingulate gyrus and the paracingulate gyrus (Heckers et al., 2004). However, Wingenfeld et al. (2009) have observed increased activation patterns in both dorsal (cognitive) and rostral (affective) subdivisions of the ACC during a negatively-themed emotional Stroop task.

Despite all of the research supporting the activation of structures involved in cognitive function, it should not be assumed that all of the areas that were most activated were only processing word category information on a cognitive level, rather than an affective level. For example, control subjects experienced similar reaction times to neutral and emotional words, exhibiting negative Stroop bias scores for positive, negative and aggressive word stimuli, making it appear that they were not responding to the emotional content of the words. However, significant clusters of increased activation were present in the PFC, PCC and dorsal ACC during the presentation of those word types and not during the presentation of neutral stimuli. This could imply that the structures involved in processing the emotional words were not simply processing the words on a cognitive attentional level, because if this were true then one might expect the same structures be involved in the processing of the

neutral words as well. Additionally, one might expect positive emotional Stroop bias scores to result if participants were simply allocating more attention to the emotional words.

The controls in this study also displayed increased activation in areas that have been linked with increased activation for arousal-related responses (Kesinger & Schacter, 2006); and emotion conditions for negative stimuli (Compton et al., 2003; Lane et al., 1999; Lane, Reiman et al, 1997; Lang et al., 1998). These studies note that the dorsolateral PFC is engaged during conditions where it is more difficult to ignore task-irrelevant information (i.e., emotional word stimuli), and that words that capture attention, compared to neutral words are more attentionally demanding (Compton et al., 2003).

Cognition and emotion are described in the literature as complexly intertwined, with emotion potentially having an influence on the process of attention. In studies of emotion, the medial frontal cortex has been implicated in attending to internal emotional states (Lane, Fink, Chau, & Dolan, 1997), and generating emotional responses to pictures (Teasdale, et al., 1999). Needless to say, it is clear in the available literature on cognition and emotion that the assumptions regarding the neurobiological basis for behaviors are mixed. Spinella et al., (2006) suggest that pre-frontal systems connect reciprocally with the limbic and other sub-cortical structures. It is then not surprising that cognitive and emotional tasks have displayed areas of common activation, thus making the distinction between differential processing systems a difficult question to answer.

Summary and Concluding Discussion

Across the two Experiments, sexual abusers and offenders in general took longer to respond to the emotional Stroop stimuli across word categories. In Experiment 1 these differences were significant, and are somewhat consistent with findings from Smith and Waterman (2004) and Price and Hanson (2007). As expected, sexual abusers exhibited a

significant response bias toward sexual word stimuli, despite the change in method of recording response latencies. This result is consistent with the results of Smith and Waterman (2004) and Price and Hanson's (2007) studies using the same word stimulus set, and is further support for the hypothesis that sexual abusers do experience response bias towards sexual stimuli. Also consistent with previous findings is that the sexual word stimuli were not able to differentiate between offender types at a significant level. These findings could be due to the heterogeneous nature of the sexual abuser group. Future studies should attempt to increase sample sizes of the different types of sexual abusers where possible, to be able to examine potential sub-group differences.

Although violent offenders experienced the greatest response bias for the aggressive words, this result was not a significant one. In fact, the sexual abusers from Experiment 1 yielded a response bias towards the aggressive words that was similar to that of the violent offenders. This result could be due to the heterogeneous nature of the sexual abuser group, or could be due to the nature of the violent offenders' crimes. For example, some of the violent crimes that had been committed were impulsive in nature, and the participants did not necessarily have a significant history of violent offences or behavior. Therefore, it is possible that the aggressive word stimuli were not an accurate reflection of their cognitions. Further research is required that would examine the differences in Stroop responses to violent/aggressive stimuli and level of history of violence.

Finally, significant emotional Stroop bias for positive and negative words was evident for the sexual abusers compared to controls, with sexual abusers experiencing greater interference. In fact, control subjects did not seem to experience emotional Stroop bias for any of the emotional word categories, other than a slight bias for the sexual words in Experiment 2. This result was not expected, and cannot likely be attributed to an age effect because the control subjects in both experiments exhibited similar response patterns (i.e.,

negative bias scores for the emotional words) however, in Experiment 2 control subjects were matched with the offenders for age. Of note is that the negative response biases displayed by the control subjects were not large negative bias scores, thus their response latency patterns to the emotional words were very similar to their responses to neutral words. Participant levels of anxiety and depression were not controlled for in this study and could have been useful in accounting for some of the findings (Smith, 2009).

Despite some of the mixed results found in the emotional Stroop data, the differences in mean RT and emotional Stroop bias scores between offenders and controls would suggest that differences in word salience or general ability to successfully complete the task are present between group types.

Experiment 2 was conducted in an attempt to pinpoint the mechanisms that may be responsible for these key differences between the groups. Control subjects displayed significant areas of activation in the occipital, PFC, PCC, left temporal and dorsal ACC areas, for the contrasts of interest. When compared to sexual abusers, control subjects exhibited increased, non-overlapping areas of activation in these regions as well. Therefore, for the word categories eliciting emotional content controls exhibited significant clusters of activation within structures that are consistent with previous emotional Stroop studies (Bremner et al., 2004; Ersche, et al., 2010; Malhi, et al., 2005; Shin, et al., 2001; Van den Heuvel, et al., 2005; Wingenfeld, et al., 2009).

Despite the small number of sexual abusers in the study, the group exhibited decreased levels of activation (and reduced inhibition) compared to controls in areas of the brain that are linked with both cognitive and affective functions. Importantly, these results would suggest that the sexual abusers in this study process emotional information differently than controls. This study provides a baseline for research regarding the processing of sexual material and control subjects. However, future research should aim to re-test these findings

using larger samples of sexual abusers to examine whether the differences in activation patterns remain consistent across studies. This study is an important contribution that may serve as an attempt to begin to piece together the cognitive mechanisms involved in the processing of sexual and aggressive information during an emotional Stroop task. Additional research using MRI and emotional Stroop tasks could render us better prepared to answer the question of whether we can use an emotional Stroop paradigm to measure sexual interest adequately.

Finally, it is important to note that the results from the fMRI study are not intended to be generalized about the brain functioning of all sexual abusers. Although the sexual abusers in this study exhibited patterns of deactivation in structures linked to cognition and emotion, given the small sample of sexual abusers in this study and the heterogeneity of the group generalization would not be possible. However, there is currently a clear focus in the literature on cognitive approaches for the assessment of sexual interest, the emotional Stroop task being one of these cognitive approaches. Research using this task with sexual abusers has yielded significant results between sexual abusers and controls suggesting that there are differences in information-processing between sexual offenders and non-offenders. Specifically, these differences are evident when participants are presented with the word stimulus set from Smith and Waterman (2004). Therefore, this research is important for establishing a baseline of brain activation patterns for a stimulus set that has consistently yielded significant differences between controls and sexual abusers regarding the processing of sexual material.

References

Ashley, V., & Swick, D. (2009). Consequences of emotional stimuli: Age differences on pure and mixed blocks of the emotional Stroop. *Behavioral and Brain Functions, 5,* 14-25.

Augustine, J.R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research Reviews, 22*, 229-244.

Banich, M.T., Milham, M.P., Atchley, R., Cohen, N.J., Webb, A., Wszalek, T., …Magin, R. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience, 12,* 988-1000.

Banich, M.T., Milham, M.P., Jacobson, B.L., Webb, A., Wszalek, T., Cohen, N.J., & Kramer, A.F. (2001). Attentional selection and the processing of task-irrelevant information: Insights from fMRI examinations of the Stroop task. In C. Casanova & M. Ptito (Eds.), *Vision: From neuroses to cognition* (Progress in Brain Research, Vol. 134, pp. 459-470). Amsterdam: Elsevier.

Beblo, T., Driessen, M., Mertens, M., Wingenfeld, K., Piefke, M., Rullkoetter, N., Silva-Saavedra, A., …Woermann, F.G. (2006). Functional MRI correlates of the recall of unresolved life events in borderline personality disorder. *Psychological Medicine, 36*, 845-856.

Beckett, R.C. (1987). *Beliefs About Children Scale (BACS)*. Oxford: Oxford Forensic Service.

Beckmann, C., Jenkinson, M., & Smith., S.M. (2003). General multi-level linear modelling for group analysis in FMRI. *NeuroImage, 20*, 1052-1063.

Beckmann, C.F. & Smith, S.M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging, 23*, 137-152.

Beech, A.R., Kalmus, E., Tipper, S.P., Baudouin, J.-Y., Humphreys, G.W., & Flak, V.E. (2008). Children induce an enhanced attentional blink in child molesters. *Psychological Assessment, 20,* 397-402.

Bench, C.J., Frith, C.D., Grasby, P.M., Friston, K.J., Paulesu, E., Frackowiak, R.S.J., & Dolan, R.J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia, 31,* 907-922.

Bremner, J.D., Vermetten, E., Vythilingam, M., Afzal, N., Schmahl, C., Elzinga, B., & Charney, D.S. (2004). Neural correlates of the classic color emotional Stroop in women with abuse-related post-traumatic stress disorder. *Biological Psychiatry, 55,* 612-620.

Burgess, P.W., & Shallice, T. (1997). *The Hayling and Brixton Tests*. London: Harcourt Assessment.

Bush, G., Luu, P., & Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Science, 4,* 215-222.

Bush, G., Whalen, P.J., Rosen, B.R., Jenike, M.A., McInerney, S.C., & Rauch, S.L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging-validation study with functional MRI. *Human Brain Mapping, 6,* 270-282.

Carter, C.S., Mintun, M., & Cohen, J.D. (1995). Interference and facilitation effects during selective attention: An H215O study of Stroop task performance. *Neuroimage, 2,* 264-272.

Compton, R.J., Banich, M.T., Mohanty, A., Milhem, M., Herrington, J., Miller, G.A., …Heller, W. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional Stroop tasks. *Cognitive, Affective and Behavioral Neuroscience, 3,* 81-96.

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G., & Petersen, S.E. (1991). Selective and divided attention during visual discrimination of shape, color and speed. *Journal of Neuroscience, 11,* 2383-2402.

Corbetta, M., & Shulman, G.L. (2002). Control of goal-directed and stimulus driven attention in the brain. *Nature Reviews, 3,* 201-215.

Dolan, M., Millington, J., & Park I. (2002). Personality and neuropsychological function in violent, sexual, and arson offenders. *Medicine, Science, and the Law, 42*, 34–43.

Dunn, L.M., Dunn, L.M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale (BPVS).* Hampshire, UK: Pardy and Son.

Drevets, W.C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interaction between emotion and cognition. *Cognition and Emotion, 12,* 353-385.

Egloff, B., & Schmukle, S.C. (2004). Gender differences in implicit and explicit anxiety measures. *Personality and Individual Differences, 36*, 1807-1815.

Ersche, K.D., Bullmore, E.T., Craig, K.J., Shabbir, S.S., Abott, S., Müller, U., …Robbins, T.W. (2010). Influence of compulsivity of drug abuse on dopaminergic modulation of attentional bias in stimulant dependence. *Archives of General Psychiatry, 67*, 632-644.

Foa, E.B., Feske, U., Murdock, T.B., Kozak, M.J., & McCarthy, P.R. (1991). Processing of threat-related information in rape victims. *Journal of Abnormal Psychology, 100*, 156-162.

Flak, V., Beech, A.R., & Fisher, D. (2006). Forensic assessment of deviant sexual interests: The current position. *Issues in Forensic Psychology, 6,* 70-83.

Graber, B., Hartmann, K., Coffman, J. A., Huey, C.J., Golden, C.J. (1982). Brain damage among mentally disordered sex offenders. *Journal of Forensic Sciences, 27*, 125–34.

Gray, N.S., Brown, A.S., MacCulloch, M.J., Smith, J. & Snowden, R.J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology, 114,* 304-8.

Gross, T.M., Jarvik, M.E., & Rosenblatt, M.R. (1993). Nicotine abstinence produces content-specific Stroop interference. *Psychopharmacology, 110*, 333-336.

Heckers, S., Weiss, A.P., Deckersbach, T., Goff, D., Morecraft, R.J., & Bush, G. (2004). Anterior cingulate cortex activation during cognitive interference in Schizophrenia. *American Journal of Psychiatry, 161,* 707-715.

Holle, C., Neely, J. H., & Heimberg, R. G. (1997). The effects of blocked versus random presentation and semantic relatedness of stimulus words in response to a modified Stroop task among social phobics. *Cognitive Therapy and Research,* 6, 681-697.

Hucker, S., Langevin, R., Wortzman, G., Bain, J., Handy, L., Chambers, J., et al. (1986). Neuropsychological impairment in pedophiles. *Canadian Journal of Behavioral Science, 18*, 440–8.

Jenkinson, M., Bannister, P., Brady, M. & Smith, S. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage, 17*, 825-841.

Jenkinson, M., & Smith, S.M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5*, 143-156.

Johnson, W., Bouchard, T. J., Segal, N. L., Keyes, M., & Samuels, J. (2003). The Stroop Color-word test: Genetic and environmental influences; Reading, mental ability and personality correlates. *Journal of Educational Psychology, 95,* 58-65.

Kalmus, E., & Beech, A.R. (2005). Forensic assessment of sexual interest: A review. *Aggression and Violent Behavior, 10*, 193-217.

Kessinger, E.A., & Schacter, D.L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective and Behavioral Neuroscience, 6,* 110-126.

Kolárský, A., Freund, K., Machek, J., Polák, O. (1967). Male sexual deviations: association with early temporal lobe damage. *Archives of General Psychiatry, 17*, 735–43.

Kramer, D. A., & Goldman, M. S. (2003). Using a modified Stroop task to implicitly discern the cognitive organization of alcohol expectancy. *Journal of Abnormal Psychology, 112*, 171-175.

Lane, R.D., Fink, G.R., Chau, P.M-L., & Dolan, R.J. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport, 8,* 3969-3972.

Lane, R.D., Reiman, E.M., Bradley, M.M., Lang, P.J., Ahern, G.L., Davidson, R.J., & Schwartz, G.E., (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia, 35,* 1437-1444.

Lang, P.J., Bradley, M.M., Fitzsimmons, J.R., Cuthbert, B.N., Scott, J.D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex : An fMRI analysis. *Psychophysiology, 35,* 199-210.

Langevin, R., Lang, R. A., Wortzman, G., Frenzel, R. R., & Wright P. (1989). An examination of brain damage and dysfunction in genital exhibitionists. *Annals of Sex Research, 2*, 77–87.

Laws, D.R. (2009). Penile plethysmography: Strengths, limitations, innovations. In D. Thornton & D.R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 7-29). Chichester, UK: Wiley

Leung, H., Skudlarski, P., Gatenby, J.C., Peterson, B.S., & Gore, J.C. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cerebral Cortex, 10,* 552-560.

Malhi, G.S., Lagopoulos, J., Sachdev, P.S., Ivanovski, B., & Shnier, R. (2005). An emotional Stroop functional MRI study of euthymic bipolar disorder. *Bipolar Disorder, 7* (suppl 5), 58-69.

Marshall, W.L., & Fernandez, Y.M. (2003). Sexual preferences: Are they useful in the assessment and treatment of sexual offenders? *Aggression and Violent Behavior, 8*, 131-143.

Milhem, M.P., Erikson, K.L., Banich, M.T., Kramer, A.F., Webb, A., Wszalek, T., & Cohen, N.J. (2002). Attentional control in the aging brain: Insights from an fMRI study of the Stroop task. *Brain and Cognition, 49,* 277-296.

Mutter, S. A., Naylor, J. C., & Patterson, E. R. (2005). The effects of age and task context on Stroop task performance. *Memory and Cognition, 33,* 514-530.

Nunes, K.L., Firestone, P., & Baldwin, M.W. (2007). Indirect assessment of cognitions of child sexual abusers with the implicit association test. *Criminal Justice and Behaviour, 34,* 454-475.

Ó Ciardha, C., & Gormley, M. (2009). Comparing two implicit cognitive measures of sexual interest: A pictoral modified Stroop task and the implicit associations test. In D. Thornton & D.R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 177-201). Chichester, UK: Wiley

Pardo, J.V., Pardo, P.J., Janer, K.W., & Raichle, M.E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences USA, 87,* 256-259.

Paulesu, E., Frith, C.D., & Frackowiak, R.S.J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362,* 342-345.

Ponseti, J., Bosinski, H. A., Wolff, S., Peller, M., Jansen, O., Mehdorn, H. M., et al. (2006). A functional endophenotype for sexual orientation in humans. *NeuroImage*, *33*, 825–33.

Price, S. A., & Hanson, R. K. (2007). A modified Stroop task with sexual offenders: Replication of a study. *Journal of Sexual Aggression, 13,* 203-216.

Shin, L.M., Whalen, P.J., Pitman, R.K., Bush, G., Macklin, M.L., Lasko, N.B., …Rauch, S.L. (2001). An fMRI study of anterior cingulate function in posttraumatic stress disorder. *Biological Psychiatry, 50*, 932-942.

Smith, P. (2009). Assessing sexual interest with the emotional Stroop test. In D.Thornton & D.R. Laws (Eds.), Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders (pp. 159-176). Chichester, UK: Wiley

Smith, P., & Waterman, M. (2005). Sex differences in processing aggression words using the Emotional Stroop task. *Aggressive Behavior, 31,* 271-282.

Smith, P., & Waterman, M. (2004). Procession bias for sexual material: The emotional Stroop and sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 16,* 163-171.

Smith, P., & Waterman, M. (2003). Processing bias for aggression words in forensic and nonforensic samples. *Cognition and Emotion, 17,* 681-701.

Smith, S. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*, 143-155.

Spinella, M., White, J., Frank, M. L., & Schiraldi, J. (2006). Evidence of orbitofrontal dysfunction in sex offenders. *International Journal of Forensic Psychology, 1,* 62-68.

Stone, M. H., Thompson, E. H., (2001). Executive function impairment in sexual offenders. *Journal of Individual Psychology, 57*, 51–9.

Taylor, S.F., Kornblum, S. Minoshima, S., Oliver, L.M. & Koeppe, R.A. (1994). Changes in medial cortical blood flow with a stimulus response compatibility task. *Neuropsychologia, 32,* 249-255.

Teasdale, J.D., Howard, R.J., Cox, S.G., Ha, Y., Brammer, M.J., Williams, S.C.R., & Checkley, S.A. (1999). Functional MRI study of cognitive generation of affect. *American Journal of Psychiatry, 156,* 209-215.

Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., & Farah, M.J. (1997). Role of the left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, 94,* 14792-14797.

Van den Heuvel, O. A., Veltman, D. J., Groenewegen, H. J., Witter, M. P., Merkelbach, J., Cath, D. C.,…Dyck, R. (2005). Disorder-specific neuro-anatomical correlates of attentional bias in obsessive-compulsive disorder, panic disorder and hypochondriasis. *Archives of General Psychiatry, 62*, 922-933.

Van den Heuvel, O.A., Veltman, D.J., Groenewegen, H.J., Witter, M.P., Merkelbach, J., Cath, D.C.,…Dyck, R. (2005). Disorder-specific neuro-anatomical correlates of attentional bias in obsessive-compulsive disorder, panic disorder and hypochondriasis. *Archives of General Psychiatry, 62*, 922-933.

Veen, V., & Carter, C.S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage, 27,* 497-504.

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging, 13,* 120-126.

Waters, A.J., & Feyerabend, C. (2000). Determinants and effects of attentional bias in smokers. *Psychology of Addictive Behaviors,* 14, 111-120.

Wertz, J. A., & Sayette, M. A. (2001). Effects of smoking opportunity on attentional bias in smokers. *Psychology of Addictive Behaviors,* 15, 268-271.

Whalen, P.J., et al. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry, 44,* 1219-1228.

Wingenfeld, K., Rullkoefter, N., Mensebach, C., Beblo, T., Mertens, M., Kreisel, S., …Woermann, F. (2009). Neural correlates of the individual emotional Stroop in borderline personality disorder. *Psychoneuroendocrinology, 34,* 571-586.

Woolrich, M.W. (2008). Robust group analysis using outlier inference. *NeuroImage 41*, 286-301.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., & Smith, S.M. (2004). Multi-level linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage 21*, 1732-1747.

Woolrich, M.W., Ripley, B.D., Brady J.M., & Smith S.M. (2001). Temporal autocorrelation in univariate linear modelling of fMRI data. *NeuroImage 14*, 1370-1386.

Worsley, K.J. (2001). Statistical analysis of activation images. In P. Jezzard, P.M. Matthews & S.M. Smith (Eds.), *Functional MRI: An Introduction to Methods (Chapter 14)*. Oxford: Oxford University Press.

Wurm, L.H., Labouvie-Vief, G., Aycock, J., Rebucal, K.H., & Koch, H.E. (2004). Performance in auditory and visual emotional Stroop tasks: A comparison of older and younger adults. *Psychology and Aging, 19,* 523-535.

Appendix

Emotional Stroop Words by Category (Smith & Waterman, 2004)

Neutral	Aggression	Positive	Negative	Sexual	Color
Door	Rage	Devotion	Abandoned	Rape	Purple
Group	Anger	Affection	Abused	Lust	Grey
Chair	Tear	Admire	Afraid	Slap	Flesh
Telephone	Assault	Euphoric	Aggressive	Victim	Pink
Dog	Kick	Fond	Agony	Slut	Orange
Coat	Shout	Grateful	Angry	Whore	Scarlet
Sofa	Punch	Tolerant	Arrogant	Fuck	Maroon
Bag	Hate	Affectionate	Bad	Bitch	Crimson
Diary	Argue	Amused	Bitchy	Child	White
Newspaper	Temper	Love	Crazy	Woman	Black
Eat	Fight	Joy	Cruelty	Force	Cyan
Oven	Kill	Proud	Detest	Penetrate	Brown
Floor	Punish	Fond	Envy	Control	Tin
Shopping	Annoyed	Funny	Fear	Schoolgirl	Bronze
Umbrella	Guilt	Glad	Frightened	Power	Mauve
Windy	Scream	Comfortable	Fury	Dominance	Gold
Radio	Crush	Beloved	Glum	Force	Silver
Painting	Slash	Calm	Greed	Oral	Jade
Milk	Smash	Peace	Hateful	Abuse	Topaz
School	Cut	Daring	Spiteful	Incest	Emerald
Ball	Wound	Cheerful	Suspicious	Molest	Purple
Pencil	Injure	Warm	Rage	Man	Grey
Cigarette	Threaten	Protective	Pain	Grope	Flesh
Football	Knife	Hope	Lose	Prostitute	Pink
Shoe	Wound	Lively	Misery	Anal	Orange

Table 1

Scale Responses and Stroop Errors per Group with (Standard Deviations)

	Sexual offenders (n=27)	Violent offenders (n=21)	Non-offenders (n=35)
BACS (CD)	3.80 (7.21)	2.80 (3.38)	3.49 (3.91)
BACS (EC)	9.52 (11.20)	7.60 (7.11)	11.74 (9.82)
Hayling	5.58 (1.03)	6.10 (0.72)	6.34 (0.84)
Brixton	6.76 (2.19)	8.10 (1.48)	8.71 (1.45)
BPVS raw score	144.31 (11.53)	146.33 (10.67)	146.31 (9.39)
Errors	5.78 (6.64)	4.10 (2.30)	7.94 (4.60)

Table 2

Mean Response Times (RT) and Bias Scores with (Standard Deviations) for the Emotional Stroop Task

Word Class	Group	RT, ms (SD)	Bias score, ms (SD)
Neutral	Sexual abusers	765.00 (130.50)	
	Violent offenders	701.60 (118.66)	
	Controls	599.54 (103. 32)	
Positive emotion	Sexual abusers	802.26 (158.19)	37.26 (82.88)
	Violent offenders	710.16 (136.82)	8.56 (68.80)
	Controls	586.25 (84.21)	-13.28 (50.65)
Negative emotion	Sexual abusers	822.67 (162.50)	57.67 (77.00)
	Violent offenders	709.17 (115.44)	7.57 (67.69)
	Controls	594.38 (91.36)	-5.15 (62.43)
Color	Sexual abusers	800.68 (149.30)	35.69 (81.09)
	Violent offenders	724.92 (141.80)	23.31 (62.91)
	Controls	608.36 (97.75)	8.82 (44.05)
Aggression	Sexual abusers	780.05 (126.11)	15.05 (71.37)
	Violent offenders	720.90 (156.26)	19.29 (78.62)
	Controls	589.03 (101.34)	-10.51 (64.34)
Sexual	Sexual abusers	811.41 (154.21)	46.41 (75.58)
	Violent offenders	720.18 (121.66)	18.57 (72.99)
	Controls	598.75 (96.30)	-0.78 (52.82)

Table 3

Scale Responses and Stroop Errors Per Group with (Standard Deviations) for the fMRI Emotional Stroop Task

Variable	Sexual Offenders	Controls
BACS (CD)	1.0 (1.73)	5.6 (6.36)
BACS (EC)	0.67 (0.58)	14.30 (10.45)
Hayling	6.33 (0.57)	6.0 (1.16)
Brixton	5.57 (3.21)	7.6 (2.28)
BPVS	147.33 (4.16)	157.30 (3.77)
Number of errors	3.00 (1.00)	2.90 (2.47)

Table 4

Mean Response Times (RT) and Bias Scores with (Standard Deviations) for the fMRI Emotional Stroop Task

Word Class	Group	RT, ms (SD)	Bias score, ms (SD)
Neutral	Sexual abusers	907.16 (115.05)	
	Controls	825.50 (171.77)	
Positive emotion	Sexual abusers	916.24 (177.53)	9.08 (62.56)
	Controls	809.40 (146.08)	-16.10 (55.64)
Negative emotion	Sexual abusers	884.53 (183.81)	-22.63 (71.54)
	Controls	791.27 (151.43)	-34.23 (51.89)
Color	Sexual abusers	957.61 (207.17)	50.44 (92.44)
	Controls	884.69 (175.24)	59.19 (81.01)
Aggression	Sexual abusers	897.95 (179.70)	-9.21 (66.53)
	Controls	795.84 (125.63)	-29.66 (77.32)
Sexual	Sexual abusers	1010.30 (201.62)	103.14 (95.18)
	Controls	864.87 (158.10)	39.37 (35.24)

Table 5

Areas of Activation for Control Subjects per Word Category

Contrast	Structure	Location x, y, z (mm)	Z score	Volumes (Voxels)
Neutral	R Occipital Pole	24, -98, -4	4.79	2905
	L Occipital Pole	-34, -94, -12	4.96	2488
Aggression	R/L Occipital Pole, Lateral Occipital Cortex	-36,-96,-10	5.01	6203
	L Precentral Gyrus, inferior frontal Gyrus	-62, 4, 16	3.72	577
Positive	R/L Occipital Pole	-36, -94, -12	5.0	5486
	L Supramarginal gyrus, Anterior Division extending to:	-42, -36, 40	3.5	402
	- Parietal Operculum Cortex	-42, -34, 26		
	- Postcentral Gyrus	-50, -20, 30		
	L Precentral Gyrus	-62, 4, 16	3.67	399
Negative	R Occipital Pole, Lateral Occipital Cortex, Inferior Division	18, -104, -2	4.82	2881
	L Occipital Pole, Lateral Occipital cortex, Inferior Division	-34, -92, -14	4.87	2675
	L Precentral gyrus, extending to:	-46, 0, 30	3.86	466
	- Inferior Frontal Gyrus	-50, 8, 6		
Sex	L Lateral Occipital Cortex, Inferior Division, extending to:	-30, -90, -16	4.87	5331
	- R Occipital Pole	20, -100, 6		
	L Precentral Gyrus extending to:	-44, 6, 30	3.65	557
	- L Middle Frontal Gyrus	-48, 12, 38		
Color	L Lateral Occipital Cortex, Inferior Division, extending to:	-30, -90, -16	4.87	5331
	- R Occipital pole	20, -100, 6		
	L Precentral Gyrus	-44, 6, 30	3.65	557

Table 6

Areas of Significant Clusters of Experimental Target Word Contrasts: Controls

Contrasts	Structure	Location x, y, z (mm)	Z score	Volume (voxels)
Aggression – Neutral	L Occipital Pole, Lingual Gyrus, extending to:	-8, -100, -4	3.61	1097
	- R Occipital Fusiform Gyrus	14, -86, -18		
	L Frontal Medial Cortex, extending to:	-2, 48, -14	3.06	367
	- R Paracingulate Gyrus	14, 40, -6		
	L Inferior Frontal Gyrus, extending to:	-58, 12, 20	3.06	307
	- L Precentral Gyrus	-60, 6, 34		
Positive - Neutral	L/R Intracalcarine Cortex	-6, -90, 0	3.63	2202
	R Frontal Pole, extending to:	10. 72, 12	3.33	395
	R/L Paracingulate Gyrus	0, 52, 10; -4, 52, 10		
Sexual - Neutral	L Inferior frontal gyrus, extending to:	-52, 22, 8	3.27	785
	- L Frontal orbital cortex	-30, 20, -4; -30, 20, -18		
	- L Temporal Pole	-30, 20, -18		
	R/L Intracalcarine cortex	4, -78, 8	3.23	668
	L Frontal pole, extending to:	-6, 62, 32	3.19	419
	- R Paracingulate Gyrus	-10, 38, 28; 2, 54, 20; 6, 44, 26		
Colour - Neutral	L Lingual gyrus, extending to:	-16, -80, 0	3.02	541
	- R Occipital Pole	8, -94, 6		
	L Inferior frontal gyrus	-48, 32, 20	3.23	417

Table 7

Areas of Significant Clusters of Experimental Target Word Contrasts for Controls > Sexual Offenders

Contrasts	Structure	Location x, y, z (mm)	Z score	Volume (voxels)
Aggression – Neutral	L Occipital Pole, extending to:	-6, -102, -6	3.92	571
	- R Occipital fusiform gyrus	14, -88, -16		
Positive - Neutral	R Paracingulate gyrus	0, 44, 26	3.25	582
	R Lingual gyrus, extending to:	8, -68, 4	3.55	422
	- R Cuneal Cortex	10, -80, 24; 12, -72, 26		
Negative - Neutral	Insular cortex, extending to:	36, 6, 4	3.64	713
	- R Planum polare	58, -4, 0; 54, 0, -2		
	-Frontal operculum cortex	44, 18, 0		
	- Central Opercular cortex	48, -2, 6		
Sexual - Neutral	Intracalcarine cortex	4, -76, 12	3.29	667
	L Frontal orbital cortex, extending to:	-50, 22, -12	3.44	660
	- L Inferior Frontal Gyrus	-54, 22, 8; -58, 24, 20		
	- L Middle frontal gyrus	-48, 34, 18		
	L/R Cingulate gyrus, anterior division	0, 24, 22	3.75	578
	R frontal orbital cortex, extending to:	30, 16, -20	3.97	485
	- R Putamen	18, 8, -8		
	- R Caudate	12, 10, 0		
	R planum polare, extending to:	44, 6, -12	3.91	461
	- R temporal pole	46, 10, -18		
	- Superior temporal gyrus, anterior division	56, 4, -16; 48, 0, -18		
Colour - Neutral	R Temporal pole, extending to:	48, 8, -12	3.28	367
	- R Frontal orbital cortex	42, 16, -14; 50, 26, -18		
	- Planum polare	46, 0, -10		
	- Superior temporal gyrus	54, 2, -18		
	L/R Intracalcarine cortex	0, -78, 8	3.29	339

colour figure

Figure 1

Activation Areas for Contrasts of Interest for Control Subjects

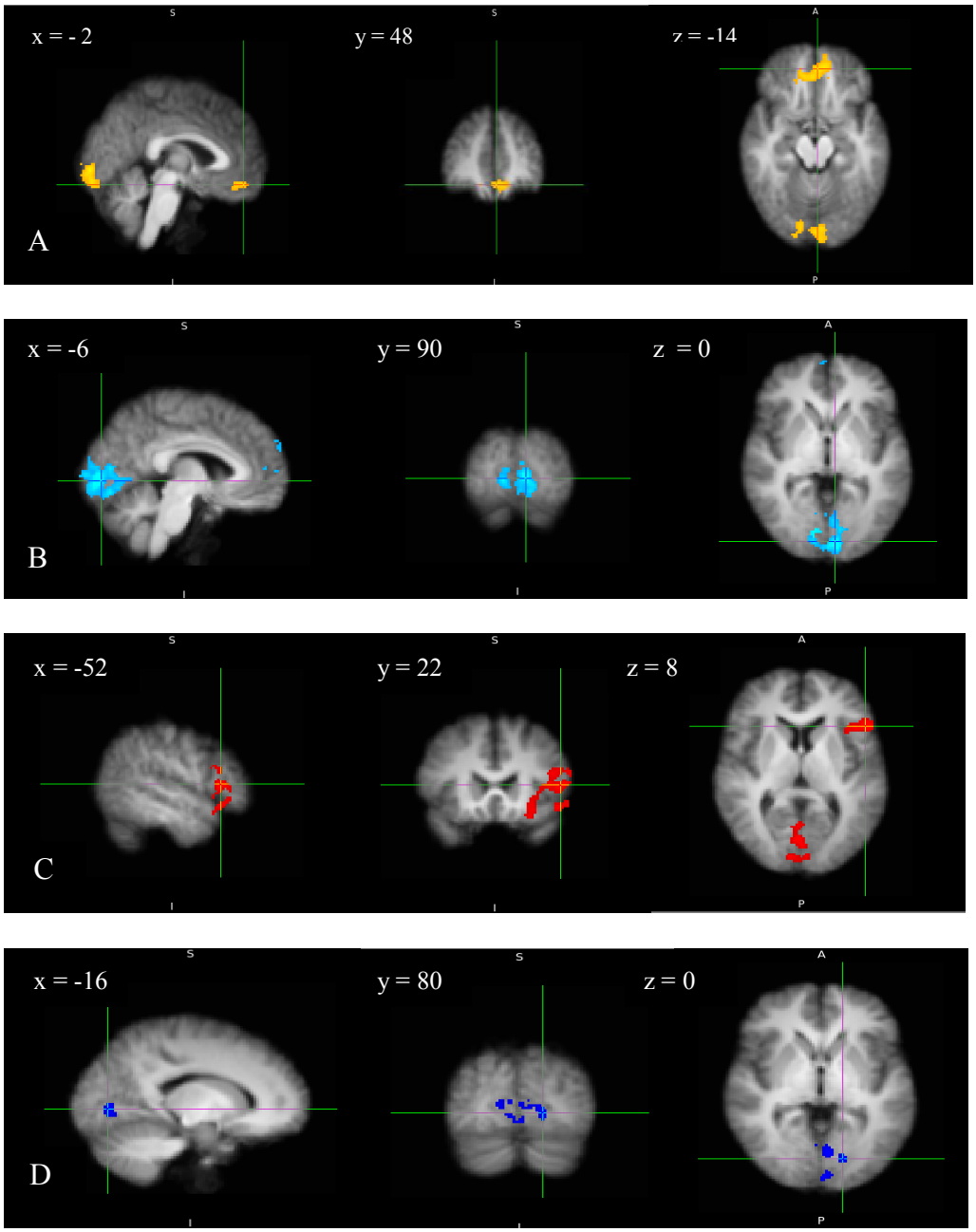


Figure 1. Brain regions in contrasts of interest showing significant activation in healthy controls: (A) aggression > neutral: three significant clusters, L occipital pole (-8, -100, -4), L frontal medial cortex (cross set at: -2, 48, -14), L inferior frontal gyrus (-58, 12, 20); (B) positive > neutral: two significant clusters, L/R intracalcarine cortex (cross set at: -6, -90, 0), R frontal pole (10, 72, 12); (C) sexual > neutral: three significant clusters, L inferior frontal gyrus (cross set at: -52, 22, 8), R/L intracalcarine cortex (4, -78, 8), R paracingulate gyrus (-10, 38, 28); (D) color > neutral: two significant clusters, L lingual gyrus (cross set at: -16, 80, 0), L infrior frontal gyrus (-48, 32, 20).

colour figure

Figure 2

Significant Clusters for Contrast Sexual - Neutral Words: Controls > Sex offenders

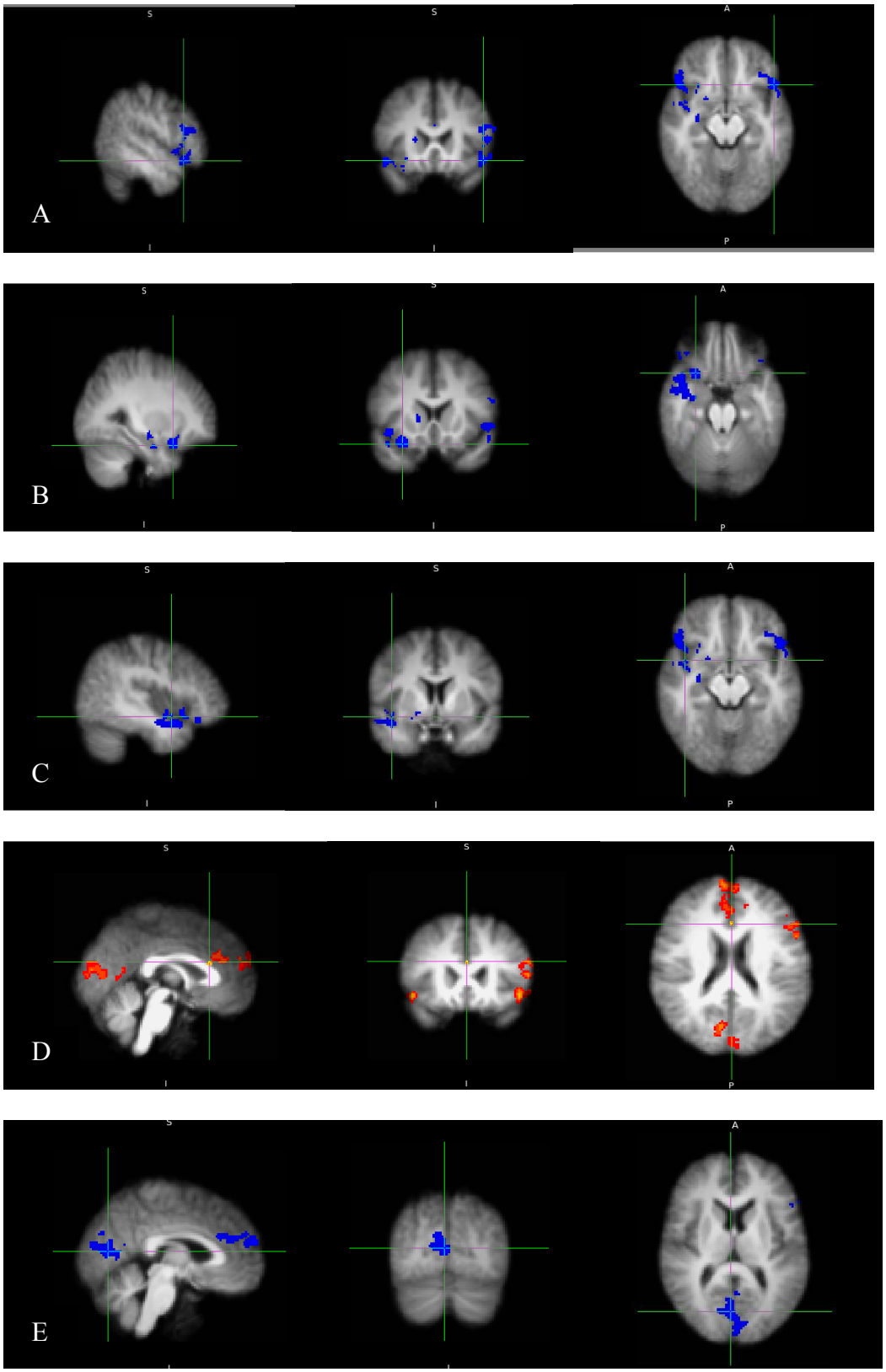


Figure 2. Brain regions showing significant activation in the comparison sexual > neutral: controls > sexual Abusers. (A) L frontal orbital cortex (-50, 22, 12), extending to L inferior frontal gyrus and L middle frontal gyrus; (B) R frontal orbital cortex (30, 16, -20); (C) R planum polare (44, 6, -12), extending to R temporal pole and superior temporal gyrus; (D) L/R cingulate gyrus, anterior division (0, 24, 22); (E) intracalcarine cortex (4, -76, 12).

colour figure

Figure 3

Significant Clusters for Contrasts of Interest: Controls > Sex offenders

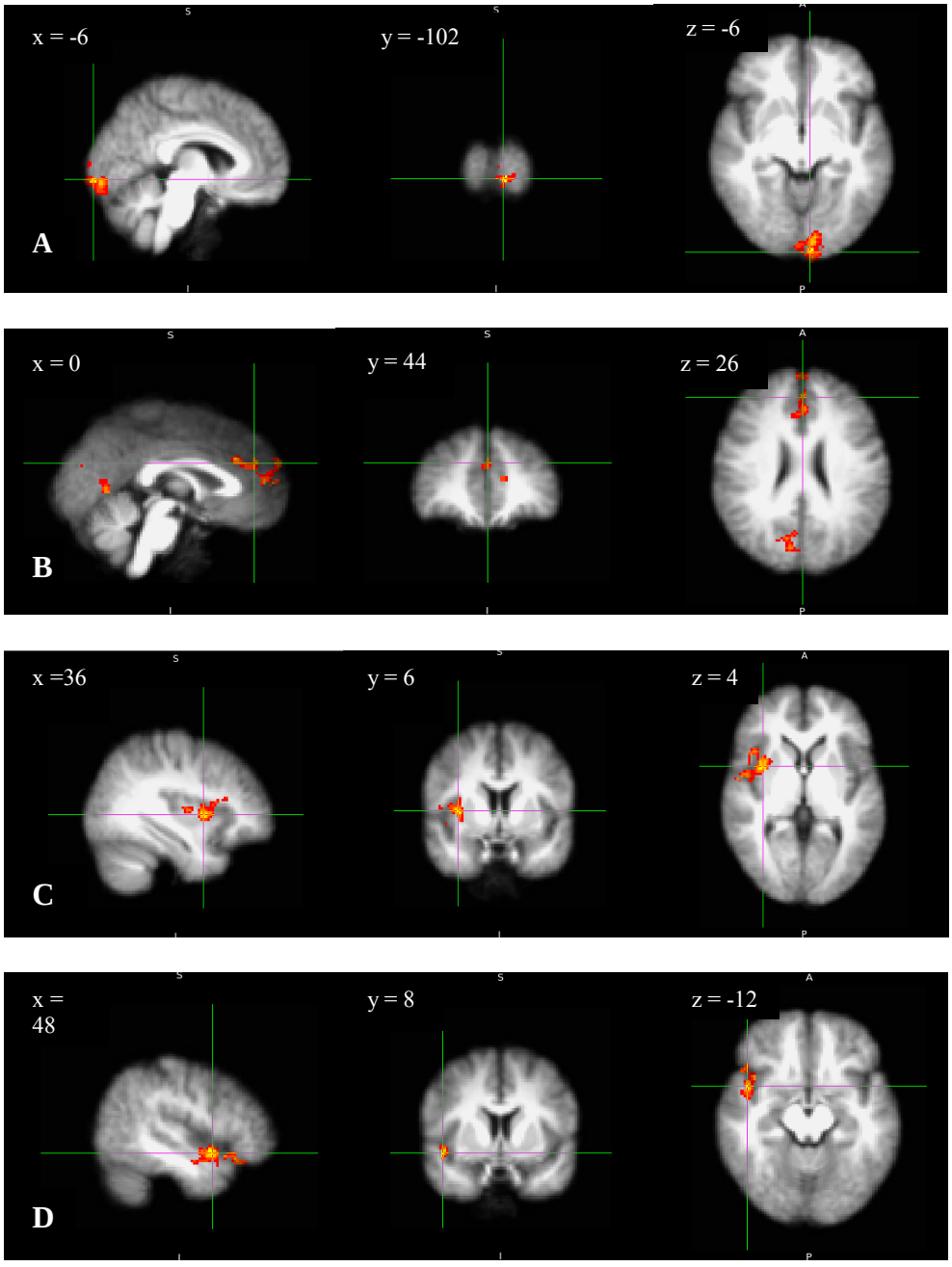


Figure 3. Brain regions in contrasts of interest showing significant activation in healthy controls > sexual abusers. (A) aggression > neutral: one significant cluster, L occipital pole (cross set at: -6, -102, -6); (B) positive > neutral: two significant clusters, R paracingulate gyrus (cross set at: 0, 44, 26), R lingual gyrus (8, -68, 4); (C) negative > neutral: one significant cluster, R insular cortex (cross set at: 36, 6, 4); (D) color > neutral: one significant cluster, R temporal pole (cross set at: 48, 8, -12).

Chapter 7

Further Development of fMRI Baseline Activation Patterns from an Emotional Stroop Task Paradigm for Use in work with Sexual Abusers

The study in Chapter 6 established baseline brain activation patterns for non-offending control subjects to the Smith and Waterman (2004) word stimulus set, more specifically for the processing of sexual material. When baseline activation patterns were compared to a small group of sexual abusers, differential patterns of brain activation were evident in areas that have been linked with cognition and emotion. The purpose of the present study was to establish an additional baseline of brain activation patterns for non-offending controls to the newly-derived word stimulus set that can later be compared with groups of sexual abusers. Preliminary evidence of differences in activation patterns to the new word stimulus set is provided in a case study comparison of activation patterns between non-offending controls and an exhibitionist.

Introduction

Neuropsychological findings from Cantor et al. (2004, 2005a, 2005b, 2006) have found that paedophilic, and non-paedophilic, men may differ in brain function and in brain structure. For example, convicted paedophilic men have been found to have lower IQs (Cantor et al., 2004, 2005a), poorer visuospatial and verbal memory scores (Cantor et al., 2004), higher rates of non-right-handedness (Cantor et al., 2004, 2005b), and elevated rates of poor school performance, or having required placement in special education programmes (Cantor et al., 2006) than non-paedophilic men. Additionally, Blanchard et al. (2002, 2003) reported that paedophiles have elevated rates of having suffered childhood head injuries resulting in unconsciousness.

Cohen et al.'s (2002) Dual Dysfunction neuroanatomic theory of paedophilia suggests that such men exhibit dysfunctions in both temporal and frontal regions. For example, observed difficulties with behavioural inhibition on tests of executive function in various groups of sexual offenders (Dolan, Millington, & Park., 2002; Langevin, Lang, Wortzman, Frenzel, & Wright, 1989; Ponseti et al., 2001; Stone & Thompson, 2001; Valliant et al., 2000) have been linked with frontal cortex dysfunction (Graber, Hartmann, Coffman, Huey, & Golden, 1982). Additionally, associations between temporal lobe epilepsy and paraphilias (Kolárský, Freund, Machek, & Polák, 1967) and between temporal lobe lesions and hypersexuality as exhibited in Klüver-Bucy Syndrome (Lilly, Cummings, Benson, & Frankel, 1983) have lead temporal-limbic theorists to suggest that temporal lobe structures either regulate sexual behaviour (Hucker et al., 1986), or that they play a role in behavioural disinhibition (Graber et al., 1982).

More recently, Cantor et al. (2008) conducted a structural MRI study examining volumes of white matter, grey matter and cerebrospinal fluid (CSF) between a group of paedophiles and nonsexual offenders. Cantor et al. (2008) found that paedophilic men have significantly lower white matter volumes than offending (non-paedophilic) men in areas of the temporal and parietal lobes, extending to the superior fronto-occipital fasciculus and the right arcuate fasciculus. Cantor et al. (2008) also found a significant negative relationship between Phallometric Pedophilia Index scores and white matter volumes. A Phallometric Index Score is a single score that summarizes a patients' level of interest in children (see Cantor et al. 2008, pp. 170 for calculation). The authors then purport that these areas of the brain may operate as a network for recognizing sexually relevant stimuli. Additionally, Cantor et al. (2008) associate the lower volumes of white matter displayed in these areas by paedophiles as indication that they experience partial disconnection within that network.

Previous fMRI studies using the emotional Stroop paradigm have identified several differential brain activation patterns involved during the processing of disorder-specific stimuli relevant to the disorders in examination (Ersche et al., 2010; Malhi et al., 2005; Shin et al., 2001; van den Heuvel et al., 2005). For example, stimulant-dependent individuals, as compared to healthy controls, exhibit increased activation in the: (1) prefrontal cortex (PFC); (2) precentral cortex (PCC); (3) parietal cortex; (4) inferior temporal cortex; (5) left thalamus; (6) caudate nucleus; (7) ACC; and (8) posterior cingulate, during the presentation of drug-word stimuli (Ersche et al., 2010). Wingenfeld et al. (2009) have reported that patients with Borderline Personality Disorder (DSM-IV-TR, APA, 2000) fail to display a fronto-limbic pattern of activation when presented with generalised and individualised negative emotion

words whereas healthy controls have displayed a pattern of increased activation in the ACC, right frontal lobe areas, left medial temporal gyrus, right hippocampus and left cuneus to the same word stimuli. Additionally, abnormal activation patterns in the inferior frontal gyrus (IFG) have been reported in disorder-related Stroop studies using patient samples with diagnoses of post-traumatic stress disorder, obsessive-compulsive disorder, bipolar disorder, panic disorders and stimulant-dependence (Ersche et al., 2010; Malhi et al., 2005; Shin et al., 2001; van den Heuvel et al., 2005).

However, cognition and emotion have been described in the literature as complexly intertwined, with emotion potentially having an influence on the process of attention. Unfortunately, this has lead to difficulties in forming a clear distinction between differential processing systems involved in cognition and emotion. Spinella, White, Frank, and Schiraldi (2006) have suggested that pre-frontal systems connect reciprocally with limbic (emotion) and other sub-cortical structures.

Contributing to the difficulties in being able to differentiate between the two processing systems is that cognitive and emotional tasks have displayed areas of common activation (Corbetta & Shulman, 2002). This point is made more apparent when we review literature regarding the structures involved in the processing of the traditional colour-word Stroop task.

For example, the following brain structures are thought to be responsible for resolving cognitive interference during incongruent colour-word stimulus presentation compared to congruent stimuli: (1) the anterior cingulate cortex (ACC) (Leung et al., 2000; Pardo et al., 1990; van Veen & Carter, 2005); (2) the insula (Leung et al., 2000); (3) the middle frontal regions (prefrontal cortex; PFC) (Leung et al., 2000; van Veen & Carter, 2005); (4) parietal regions (Leung et al., 2000; van Veen & Carter, 2005); (5) mid-temporal regions (Leung et al., 2000); (6) the inferior frontal gyrus

(IFG) (Bench et al., 1993; Leung et al., 2000; Taylor et al., 1994); and (7) the basal ganglia (Bench et al., 1993; Taylor et al., 1994). The brain regions listed above are thought to be part of the brain's attentional mechanisms due to the nature of the colour-word Stroop task and the requirement of participants to selectively attend to colour information while suppressing word information (Compton et al., 2003; Corbetta & Shulman, 2002; van Veen, & Carter, 2005).

Additionally, adequate frontal lobe functioning is thought to be necessary in order to overcome response inhibition and interference resolution, required to successfully complete the colour-word Stroop task (Adleman et al., 2002). This is because lesion studies have found that patients with lateral prefrontal and left hemisphere frontal lobe lesions commit more errors on the Stroop task than normal controls (Perret, 1974; Vendrell et al., 1995).

It is clear from the combination of findings from the colour-word Stroop task and the emotional Stroop task that the structures involved in the processing of incongruent colour-word information are often the same ones that are involved in the processing of emotional word stimuli.

Differences in mean RT and emotional Stroop bias scores have consistently been found between offenders and controls (including the studies within this thesis).The study in Chapter 6 examined whether this was also true for the processing of offence-related and sexual word information examining brain activation patterns of non-offending controls compared to a small group of sexual abusers using this type of paradigm. fMRI was used in the study in hopes of being able to pinpoint specific mechanisms that may be responsible for the key differences found between offender and control groups. Baseline brain activation patterns for non-offending control subjects in the processing of sexual material were identified. Overall, the results from

the contrasts of interest (i.e., aggression – neutral; sexual – neutral; etc.) displayed significant clusters of activation in the aggression, positive and sexual contrasts for control subjects in the occipital lobe, PCC, PFC, and temporal lobe.

A key finding when the controls were compared to sexual abusers was that the sexual abusers in the study exhibited decreased levels of activation (and reduced inhibition) in areas of the brain that have been linked with both cognitive and affective functions. These results then provided evidence that sexual abusers do in fact process emotional and sexual information differently than non offenders.

Study Purpose

The purpose of the current study was to expand on the available literature concerning the processing of sexual interest information and to use fMRI in an attempt to identify particular structures that are responsible for processing offence-related word stimuli specific to deviant sexual interests. The present study used the results from Chapter 6 as a baseline to indicate areas of activation that should be present during the processing of emotional and sexually-relevant stimuli in non-offending control subjects. This baseline was employed for comparison purposes to the patterns of activation exhibited for the new stimulus set. The activation patterns established in non-offending controls to the new word stimuli were then compared in a case study of an exhibitionist. The objective was to add to knowledge of baseline areas of activation present in healthy control participants, for these offence-related stimuli, in order to allow for more comparisons to be made with sexual abuser samples in the future. Additionally, in establishing solid baseline areas of brain activation for these word stimuli we can gain a clearer understanding of why differences in the processing of these stimuli are present between controls and sexual abusers.

Therefore, the aims of the study were:

- To determine whether similar patterns of increased activation are evident in the non-offending controls between the two word stimulus sets;
- To establish baseline areas of brain activation for healthy control subjects to the set of newly-derived stimuli;
- To compare the brain activation patterns of the control subjects in a case study to the brain activation patterns of an exhibitionist;
- To create an understanding of the cognitive mechanisms involved in the processing of normal attitudes towards sexual interest information;
- To use the knowledge gained from this research to help guide future work and research with sexual abusers.

Method

Participants

The ten control subjects that took part in the study from Chapter 6 also completed the experiments in this study. Unfortunately, data from two of the sexual abusers that took part in the study in Chapter 6 was compromised leaving data for one participant from this group. Therefore, the one unpaid sexual abuser participant (convicted of exhibitionism) was used as a case study for the study in this chapter. All participants gave written informed consent in accordance with the ethical procedures of the Birmingham University Imaging Centre (BUIC) in Birmingham, UK (available at: http://www.buic.bham.ac.uk.html). All participants had normal or corrected-to-normal vision.

Apparatus/Materials

Computerized versions of the emotional Stroop task were presented to participants using version 2.0 of E-Prime software (Psychology Software Tools, Inc.). Participants lay in a 3T Philips Achieva MRI scanner and viewed emotional Stroop stimuli via a tilted mirror on the eight-channel SENSE head coil (P-reduction=2). The colour-identification response latencies for each trial were detected and recorded through two five-button serial response boxes (Psychology Software Tools, Inc.) with two coloured buttons on each response box (green, red, and blue, white) identified as responses for the purpose of this study.

A T1-weighted high resolution anatomical scan (1 mm^3) was acquired during each session. Functional data was acquired using a T2* - weighted echoplanar imaging sequence (with 32 ascending slices, a repetition time of 2000 ms, a time to echo of 35 ms, a flip angle of 85˚ and a resolution of 2.5 mm^3, restricting the field of

view to the lower portion of the cortex) to measure blood oxygenation level-dependent (BOLD) signal.

Procedures

Participants completed a colour congruent/incongruent word set as well as the stimulus set derived for the purposes of this thesis as outlined in Chapter 2. All word stimuli were presented in a partially-blocked format for analysis purposes. Therefore, a word category block consisted of five words from the specified word category (i.e., EPD, MEPD, SA, MSA, PD, MPD) and the order of word category block presentation was counterbalanced. Prior to each word a fixation 'x' appeared on the screen for 500ms. Each word was presented on the screen for 2500ms, with participants providing their ink colour response as quickly as possible within that time frame. After every six blocks of word presentation a six second break was given where participants viewed a blank screen. For the colour-word stimulus set, a six second break was not provided as there were only 50 word trials in total, therefore a break was deemed not necessary. No opportunity to correct errors made on the Stroop task was provided.

Data Analysis

Emotional Stroop bias scores were calculated by subtracting the mean RT of matched words from the mean scores of experimental words. This resulted in three bias scores to be compared across the three groups: EPD Stroop, SA Stroop and PD Stroop. Due to the small number of participants and the case study nature of the analysis, the non-parametric version of the independent samples t-test (Mann-Whitney) was run on the mean RT's and Stroop bias scores between the controls and the case study.

FMRI data processing was again carried out using FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following lower-level preprocessing was applied to each participant's data: motion correction using MCFLIRT (absolute mean displacements per scan averaged across participants was 0.26) (Jenkinson, Bannister, Brady & Smith, 2002); slice-timing correction; non-brain removal using BET (Smith, 2002); spatial smoothing (5mm full width at half maximum Gaussian kernel); grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; and highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=45.0s). ICA-based exploratory data analysis was carried out using MELODIC (Beckmann & Smith, 2004) in order to investigate the possible presence of unexpected artifacts or activation. No components were deleted. Each participant's functional data set was registered with their anatomical image (T1) and then transformed into Montreal Neurological Institute (MNI) space using FLIRT (Jenkinson, et al., 2002; Jenkinson & Smith, 2001).

Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). Contrasts of interest included comparing activation during matched trials to the activation during experimental word category trials (i.e., EDP – MEPD, SA – MSA, PD – MPD). Additionally, a cross cope comparison was made between the emotional words from the Smith and Waterman stimulus set and the emotional words from the newly-derived stimulus set (i.e. emotion > non-emotion from one stimulus set compared to emotion > non-emotion from the other stimulus set).

Regressors describing participant motion were included as regressors of no interest. Higher-level analysis for each participant and group analysis was carried out

using FLAME (FMRIB's Local Analysis of Mixed Effects) stage 1 with automatic outlier detection (Beckmann, Jenkinson & Smith, 2003; Woolrich, 2008; Woolrich et al., 2004). Z (Gaussianised T/F) statistic images were thresholded using clusters determined by $Z > 2.3$ and $Z > 1.8$ and a (corrected) cluster significance threshold of $P=0.05$ (Worsley, 2001).

Establishing Baseline Activations

Uncorrected images of the incongruent - congruent contrast of interest were used to verify that the control subjects from this study were displaying similar patterns of brain activation to the control subjects reported elsewhere in the fMRI literature using the colour-word Stroop task. The rationale for using a colour-word stimulus set in addition to the emotional word stimuli was to further validate the activation patterns observed in this study. For example, in ensuring that the control subjects in this study exhibited similar patterns of increased activation as previously reported in the literature for healthy controls, we can be more confident about the reliability of the comparisons made across the groups in this study.

Three separate sources from the functional MRI data from Chapter 6 were used as a baseline for areas of activation present during the presentation of emotional and offence-related word stimuli for non-offending control subjects: (1) the significant areas of increased activation reported for the contrasts of interest (i.e., aggression-neutral, positive-neutral and sexual-neutral); (2) the offence-specific sexual - neutral contrast of interest findings averaged out across the controls; and (3) activation areas observed for all emotional words in the Smith and Waterman (2004) stimulus set (i.e., positive, negative, aggressive and sexual) compared to the activation

patterns for all non-emotional word stimuli (i.e., neutral and incongruent colour words).

Results

No significant differences were found when comparing the mean RT's and Stroop bias scores. This is not surprising given the small number of participants. The mean RT's and emotional Stroop bias scores are displayed in Table 1 for the controls and the case study. As can be seen from Table 1, the exhibitionist was consistently slower to complete the task across the different word categories. Similarly to the adult sexual abusers from Chapter 3, the exhibitionist displayed a bias towards the emotional/personality descriptors. Unlike the sexual abusers in Chapter 3, the exhibitionist displayed a large bias towards the physical descriptor word stimuli, whereas the sexual action words did not elicit an effect. Across the word categories in this study the exhibitionist displayed Stroop bias effects that were contrary to those displayed by the non-offending controls.

Table 1

Mean RT's and emotional Stroop bias scores for controls and case study

Word type	Group	RT Ms (SD)	Stroop Bias
EPD	Controls	760.75 (156.92)	-17.01 (55.17)
	Case study	999.66	23.92
MEPD	Controls	777.75 (147.61)	
	Case study	975.73	
SA	Controls	812.70 (159.76)	34.60 (76.60)
	Case study	896.67	-78.96
MSA	Controls	778.11 (149.10)	
	Case study	975.63	
PD	Controls	781.53 (156.19)	1.38 (60.01)
	Case study	906.97	75.65
MPD	Controls	780.15 (156.19)	
	Case study	831.31	

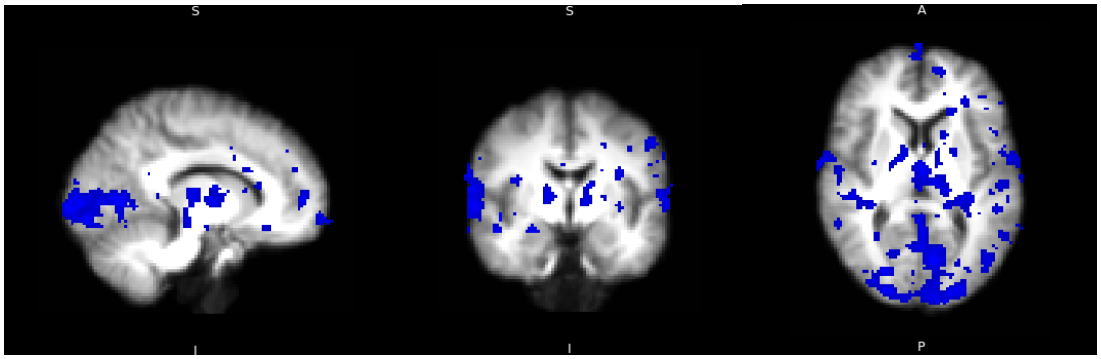
Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD); all experimental words (EXP). 'M' represents matched categories.

Baseline Imaging Results

The uncorrected images from the incongruent - congruent colour contrast of interest for non-offending controls displayed activation in the following areas: (1) ACC; (2) posterior cingulate cortex; (3) left and right thalamus; (4) left putamen; (5) left paracingulate gyrus; (6) left IFG; (7) fontal pole; (8) occipital lobe; (9) superior

temporal gyrus; (10) insular cortex; (11) parietal; and (12) temporal lobes (see figure 1A). Verification of the activation areas present during the incongruent - congruent contrast then confirmed that the control subjects in this experiment display similar areas of activation as have been recorded in previous studies using the colour-word Stroop task (Bench et al., 1993; Leung et al., 2000; Pardo et al., 1990; Taylor et al., 1994; van Veen & Carter, 2005).

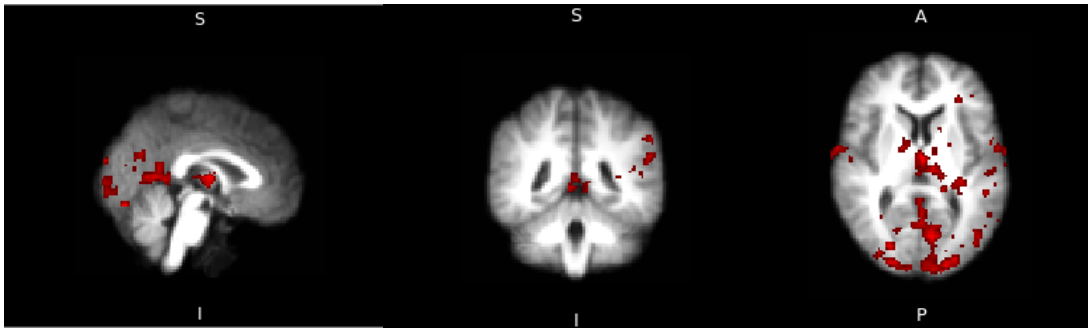
Figure 1A. *Uncorrected Areas of Activation for Non-Offending Controls: Incongruent – Congruent Contrast of Interest*



Note: Figure 1A Brain regions in incongruent > congruent contrast of interest showing uncorrected activation in healthy controls (i.e., clusters are not thresholded to meet a z-score).

Further analysis of the incongruent – congruent contrast of interest thresholded at clusters determined by $Z > 1.8$ ($p = 0.05$) revealed significant areas of activation for the control subjects in the occipital lobe, temporal lobe, thalamus and IFG (see figure 1B).

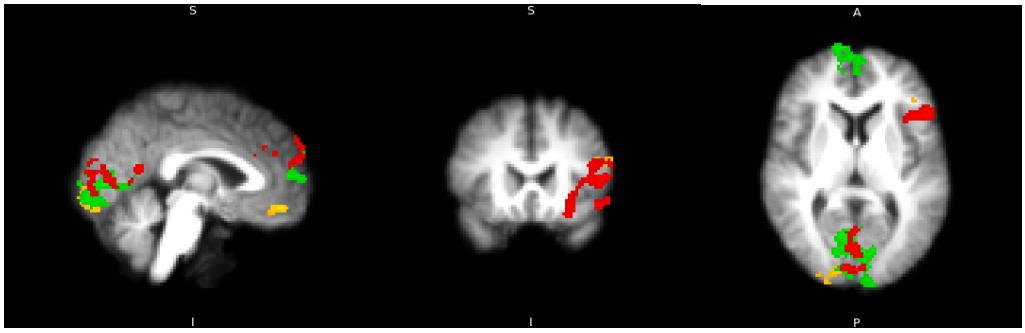
Figure 1B. *Significant Clusters for Contrast Incongruent – Congruent: Non-Offending Controls (threshold Z > 1.8, p=.05)*



Note. Figure 1B Brain regions in incongruent > congruent contrasts of interest showing significant activation in healthy controls: (1) R/L occipital pole (26, -96, 2), extending to L lingual gyrus (-8, -72, 4); (2) L postcentral gyrus, extending to R/L thalamus (4, -20, 6; 0, -10, 8) and L parahippocampal gyrus (-24, -40, 0); (3) L inferior frontal gyrus (-50, 14, 20), extending to L frontal pole (-36, 44, 24) and L superior temporal gyrus (-66, -6, -2); and (4) R temporal pole (52, 16, -10), extending to R postcentral gyrus (66, -6, 14) and R precentral gyrus (68, 2, 18).

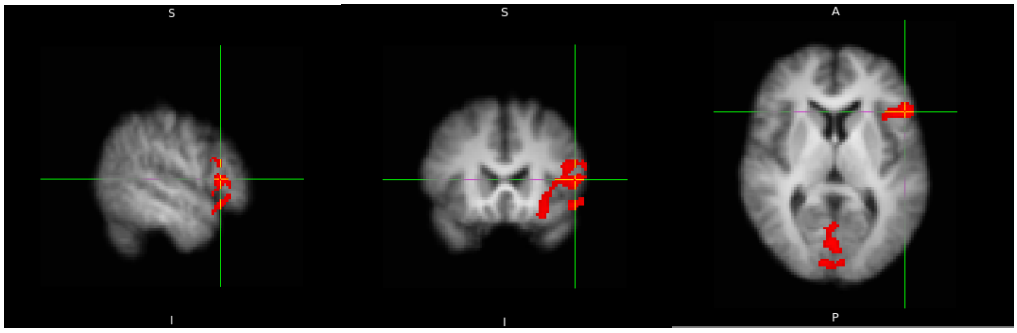
Baseline images for the emotional and offence-specific observed increases in brain activation from the contrasts of interest in Chapter 6 are presented in Figure 2A. The significant results between the word categories are displayed overlapping each other to illustrate where the similar activations were apparent between the word categories. Control subjects exhibited increased activation the prefrontal cortex (PFC), precentral cortex (PCC), inferior frontal gyrus (IFG), anterior cingulate cortex (ACC) and temporal and parietal lobes (*x*, *y*, *z* coordinates provided in Chapter 6). The offence-specific (i.e., sexual - neutral) imaging data is displayed in Figure 2B.

Figure 2A. *Overlapping Areas of Activation for Contrasts of Interest in the Smith and Waterman Word Stimulus Set: Non-Offending Controls*



Note. Figure 2A Brain regions in contrasts of interest showing significant activation in healthy controls: (yellow) aggression > neutral; (green) positive > neutral; (red) sexual > neutral.

Figure 2B. *Significant Areas of Activation for Sexual – Neutral Contrast of Interest: Non-Offending Controls*

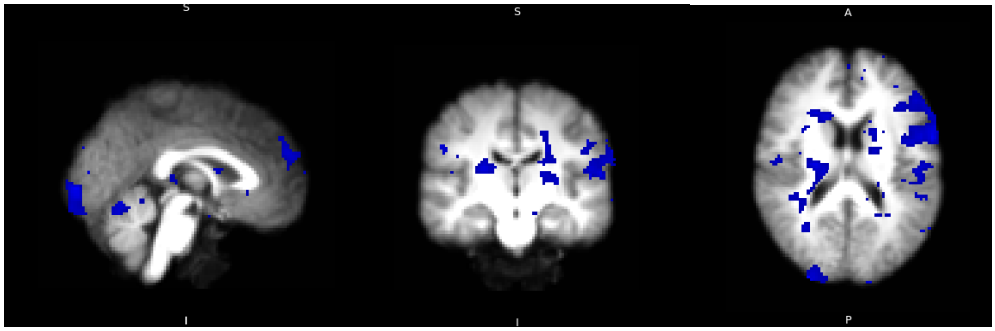


Note. Figure 2B brain regions in Sexual > Neutral contrast of interest showing significant activation in healthy controls: three significant clusters, (1) L inferior frontal gyrus (cross set at: -52, 22, 8); (2) R/L intracalcarine cortex (4, -78, 8); (3) L frontal pole (-6, 62, 32), extending to R paracingulate gyrus (-10, 38, 28).

When the activations for all of the Smith and Waterman emotional words were compared to the activations for the non-emotional words uncorrected images displayed activation in the following areas: (1) left precentral gyrus; (2) left IFG; (3) middle frontal gyrus (bilateral); (4) left frontal orbital cortex; (5) left amygdala; (6)

left hippocampus; (7) left insular cortex; (8) bilateral ACC; (9) supramarginal gyrus (bilateral) anterior division; (10) right parietal operculum cortex; (11) occipital lobe; and (12) left temporal pole. See Figure 2C for the uncorrected data for the contrast emotional words - nonemotional words from the Smith and Waterman stimulus set averaged out for controls.

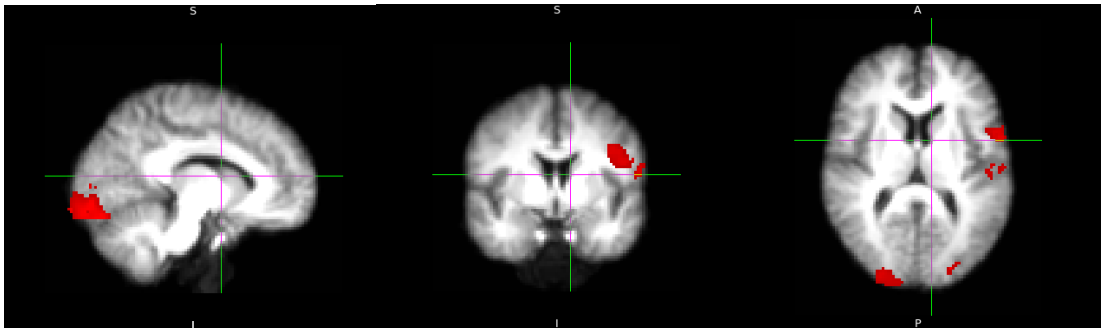
Figure 2C. *Uncorrected Areas of Activation for Non-Offending Controls: Emotion > Non-Emotion, Smith and Waterman Word Categories*



Note. Figure 2C Brain regions in emotional > non-emotional contrast of interest showing uncorrected activation in healthy controls (i.e., clusters are not thresholded to meet a z-score).

Finally, further analysis was completed between the activations present for all of the emotional words and the non-emotional word categories in the Smith and Waterman (2004) word stimulus set. From this comparison 3 significant clusters resulted for the healthy control participants (see Figure 2D).

Figure 2D. *Significant Clusters for Contrast Emotional – Non-Emotional: Non-Offending Controls (threshold Z > 2.3, p=.05)*

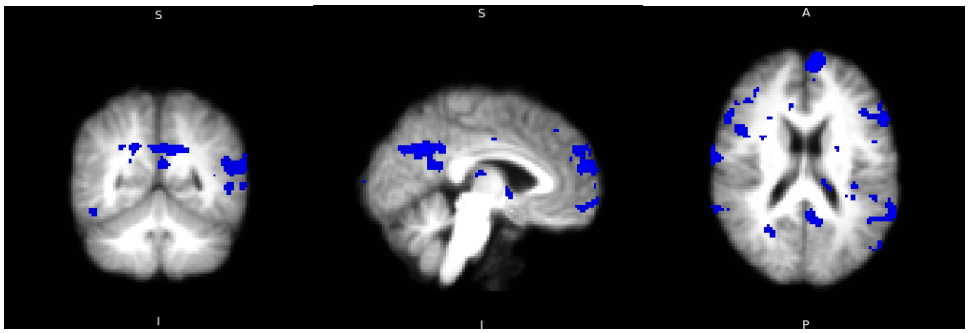


Note. Figure 2D Brain regions in emotional > non-emotional contrast of interest showing significant activation in healthy controls: (1) left occipital pole (-36, -96, -12); (2) left precentral gyrus (-62, 6, 14); and (3) the left supramarginal gyrus (-40, -32, 32), extending to the left postcentral gyrus (-42, -34, 42) and the left central opercular cortex (-54, -18, 12; -62, -18, 14).

Development of Baseline Activations for the New Word Stimuli

Viewing of the uncorrected images shows activation during both emotional and matched words to be in the: (1) precentral gyrus (bilateral); (2) left IFG; (3) left superior frontal gyrus; (4) left frontal pole; (5) middle frontal gyrus (bilateral); (6) left paracingulate gyrus; (7) left supramarginal gyrus; (8) left lateral occipital cortex; (9) left posterior cingulate gyrus and precuneous cortex; (10) thalamus (bilateral); and (11) occipital lobe. Uncorrected images of the emotion > non-emotion contrast of interest overall for the non-offending controls reveals added activations in these areas, however these activations were not significant at Z > 1.8 once the data had been corrected (see Figure 3A).

Figure 3A. *Uncorrected Areas of Activation for Non-Offending Controls: Emotion > Non-Emotion word activation, Newly-Derived Word Stimuli*

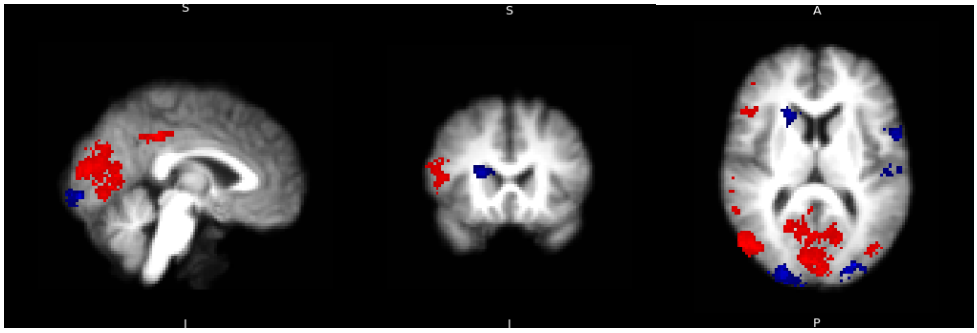


Note. Figure 3A Brain regions in emotional > non-emotional contrast of interest for the newly-derived word stimuli showing uncorrected activation in healthy controls (i.e., clusters are not thresholded to meet a z-score).

The results from the analysis of the contrasts of interest (i.e., EPD-MEPD, SA-MSA, and PD-MPD) averaged out for the control subjects did not yield any significant areas of activation. Additionally, when all of the emotional words (i.e., EPD, SA, PD) were compared to the matched words, no significant areas of activation were evident.

Uncorrected imaging data (see Figure 3B) reveals different areas of activation for the emotional words in the current study compared to the activation exhibited by controls for the emotional words from the Smith and Waterman word list. Analysis of emotional words > non-emotional words compared between the newly derived set and the baseline yielded several significant differences in areas of activation (see Table 2).

Figure 3B. *Areas of activation for Contrast Emotion > Non-Emotion: Smith and Waterman Stimulus Set (blue) and Newly-Derived Word Stimuli (red)*



Note. Figure 3B Brain regions in emotional > non-emotional contrasts of interest showing uncorrected activation in healthy controls (i.e., clusters are not thresholded to meet a z-score): Red (Newly-Derived > Smith and Waterman); Blue (Smith and Waterman > Newly-Derived).

Table 2

Areas of Significant Activation for the Emotional > Non-Emotional Contrast of Interest between the Newly-Derived Word Stimuli and Smith and Waterman Stimuli.

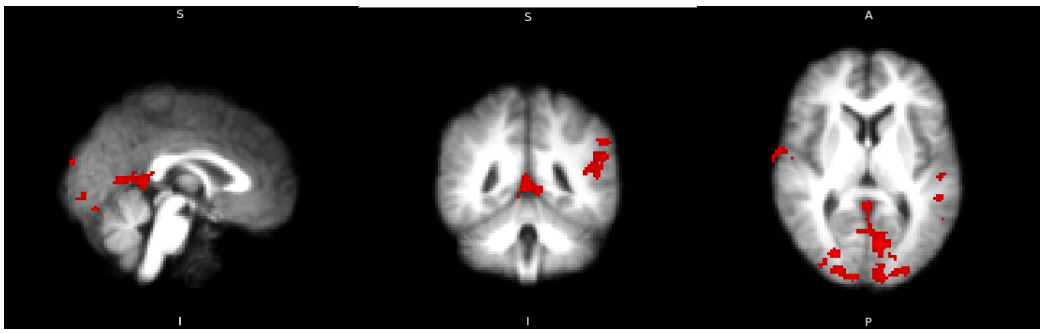
Contrast	Structure	Location	Z-Score	Voxels
Newly-Derived > Smith & Waterman	R cuneal cortex	4, -76. 32	4.08	5798
	L lateral occipital cortex, superior division	-40, -88, 28	3.43	495
	R precuneous cortex (cingulate gyrus)	4, -44, 40 (4, -36, 40)	4.16	437
	R inferior frontal gyrus (R frontal pole)	52, 34, 2 (52, 38, -6)	3.58	387
	R superior temporal gyrus (R middle temporal gyrus, insular cortex, planum polare)	64, -26, 6 (56, -16, -10; 42, -10, -6 48, -10, -6)	3.31	289
	R frontal pole	22, 68, -4	3.42	206
Smith & Waterman > Newly-Derived	L occipital pole	-36, -96, -10	4.81	4278
	L precentral gyrus	-60, 6, 16	3.48	562
	L supramarginal gyrus (L postcentral gyrus)	-40, -32, 32 (42, -34, 42)	3.5	537

Note. x, y, z coordinates of significant activations extending from the peak areas of activation are listed in brackets.

Case Study Comparisons

In the comparison between the non-offending controls and the exhibitionist for the colour incongruent – congruent contrast of interest four significant areas of peak activation were evident for the control subjects and not for the sexual abuser (see figure 4A).

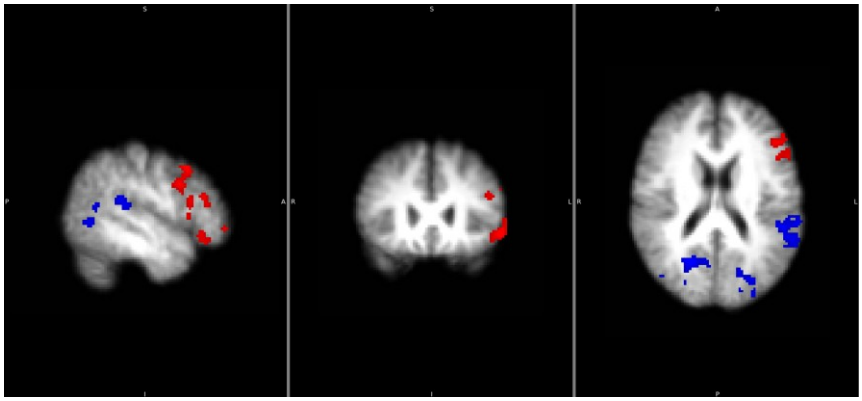
Figure 4A. *Significant Clusters for Contrast Incongruent – Congruent: Non-Offending Controls > Exhibitionist (threshold Z > 1.8, p=.05)*



Note. Figure 4A Brain regions showing significant activation in the comparison incongruent > congruent: non-offending controls > exhibitionist: (1) L lingual gyrus (-8, -72, 4), extending to L occipital pole; (2) R occipital pole (30, -102, -12); (3) R superior temporal gyrus (66, -4, 0), extending to R postcentral gyrus, R temporal pole and R precentral gyrus; and (4) L postcentral gyrus (-58, -24, 42), extending to the L parietal operculum cortex, L postcentral gyrus and L supramarginal gyrus (anterior and posterior divisions).

Finally, in the comparisons for the contrast of interest for the newly-derived word stimuli, the non-offending controls exhibited significant areas of increased activation compared to the exhibitionist. Increased activation was evident for the controls in the ACT-MACT contrast of interest and the PHY-MPHY contrast of interest (see Figure 4B).

Figure 4B. *Significant Clusters for Contrast SA – MSA (blue) and PD – MPD (red): Non-Offending Controls > Exhibitionist*



Note. Figure 4B Brain regions in contrasts of interest showing significant activation in non-offending controls > exhibitionist. (Blue) SA > MSA: two significant clusters: R supracalcarine cortex (22, -60, 20), extending bilaterally to the precuneous cortex (16, -66, 34; -10, -68, 30) and the L middle temporal gyrus (-60, -52, -8); (Red) PD > MPD: one significant cluster: L frontal orbital cortex (-56, 26, -8) extending to the L middle frontal gyrus (-50, 16, 46; -56, 14, 38), L precentral gyrus (-50, 10, 32) and L inferior frontal gyrus (-48, 28, 22; -58, 20, 0).

Discussion of Results

Control subjects in this study showed similar areas of activation in the incongruent – congruent colour Stroop contrast of interest as healthy controls have exhibited in previous research using the colour-word Stroop (Bench et al., 1993; Leung et al., 2000; Pardo et al., 1990; Taylor et al., 1994; van Veen & Carter, 2005). Verification of these activations was conducted to ensure that additional comparisons could be carried out from a reliable baseline. Additionally, the baseline activations evident for the Smith and Waterman stimuli were consistent with previous findings from fMRI and emotional Stroop task activation patterns (Ersche et al., 2010; Malhi et al., 2005;

Shin et al., 2001; van den Heuvel et al., 2005). Therefore, the non-offending controls in this study are responding in a manner consistent with previous findings, and the choice to use these activations as a baseline for further comparisons to be made were appropriate.

No significant differences were evident for the contrasts of interest from the newly-derived word categories for the non-offending controls. These results could suggest that the new word categories are noisier in nature than the ones used in the previous study. However, it is also possible that significant differences were not evident for the non-offending controls between the contrasts because the word stimuli were not of a salient nature to this group. For example, these word stimuli were developed in order to reflect the sexual interests of sexual abusers. Therefore, one might expect that no significant differences in brain activations would be evident between the target words and matched words for non-sexual abusers. The results would support this argument when we consider that similar areas of activation were evident for the controls for both the target words and matched words in this stimulus set.

Additionally, the lack of significant differences in activation between all emotional words and non-emotional words would suggest that the structures activated were involved in the successful completion of the task itself (i.e., cognitive function), rather than on overcoming interference in information-processing caused by emotional content. Importantly, fewer areas and lower volumes of activation were evident for the overall emotion contrast in the key areas that have been implicated in overcoming colour and emotional Stroop interference such as the ACC, temporal and parietal regions. Therefore it would appear that the structures involved in processing the 'emotional' words were simply processing the words on a cognitive attentional

level, because the same structures were involved in the processing of the matched words as well.

However, significant differences were evident for the non-offending controls in the emotion > non-emotion contrasts across the word stimulus sets. This means that there were differences exhibited by the same participants for what would be considered the emotional words from the Smith and Waterman word list and the emotional words in the new stimulus set. Although the contrasts did share common areas of activation, this finding is further evidence that the target words between the stimulus sets are different in context.

Case Study Comparison

For the colour-word Stroop task four areas of significant activation were present for the non-offending controls over the sexual abuser. The areas of significant activation were in the occipital, temporal, precentral and parietal areas of the brain. These structures have been implicated in resolving cognitive interference during incongruent colour-word stimulus presentation compared to congruent stimuli (Leung et al., 2000; van Veen & Carter, 2005). Importantly, this would indicate that the sex offender showed decreased levels of activation (and reduced inhibition) in these areas. Therefore, it would appear that the exhibitionist processes cognitive information differently from the controls.

The behavioural data for the newly-derived word stimuli reveals Stroop bias patterns for the controls that are similar to those experienced by the non-offending adult controls in Chapter 3. For example, the controls yielded a negative Stroop bias for the emotional/personality descriptors, a small bias for physical descriptors, and a positive bias for the sexual action words. Interestingly, the exhibitionist showed a

response bias towards the emotional/personality descriptors (as did the sexual abusers in Chapter 3) and physical descriptors. However, a negative Stroop bias was observed for the sexual action word stimuli.

Although the non-offending controls did not display significant differences in activation for the contrasts of interest for the stimulus set in this study, when compared to the sexual abuser, increased activation for the sexual action and physical descriptor word stimuli was evident. Not only is this important because it indicates a pattern of deactivation for the sexual abuser, but also because of the nature of the word categories and the structures where these differences were evident. For the sexual action words the exhibitionist displayed reduced inhibition in the supracalcarine cortex and middle temporal areas. For the physical descriptors the exhibitionist experienced reduced inhibition in the frontal orbital cortex, middle frontal gyrus, precentral gyrus and inferior frontal gyrus. Therefore, the sexual abuser in this study displayed results that correspond with Cohen et al's (2002) dual dysfunction theory, whereby dysfunctions in both temporal and frontal regions have been observed in paedophilic offenders and have lead theorists to suggest play a role in behavioural disinhibition and regulating sexual behaviour for various groups of sexual offenders (Graber et al., 1982; Hucker et al., 1986).

Conclusions

The non-offenders from this study did not exhibit significant differences in brain activation between the target words (i.e., emotional/personality descriptors, sexual actions and physical descriptors) and the matched words for the new stimulus set. This result suggests that the stimuli were not salient for this group and that the activations that were observed functioned to process the stimuli on a cognitive level,

rather than having an affective function. The sexual abuser in this study displayed reduced inhibition in temporal and frontal/prefrontal areas that have previously been linked with behavioural disinhibition and the regulation of sexual behaviour (Dolan, Millington, & Park., 2002; Graber et al., 1982; Hucker et al., 1986; Langevin, Lang, Wortzman, Frenzel, & Wright, 1989; Lilly, Cummings, Benson, & Frankel, 1983; Ponseti et al., 2001; Stone & Thompson, 2001; Valliant et al., 2000). These findings would suggest that, not only does the sexual abuser display dysfunction in cognitive processing, but also displays dysfunction in the processing of sexual interest information.

CHAPTER 8

Discussion

Thesis Aims

The overall aims of this thesis were to examine differences in information-processing of sexual material between groups of offenders and non-offenders, and to determine whether the emotional Stroop task is a reliable tool to be used in the assessment of deviant sexual interest. In addition, it was an aim of this thesis to examine the brain structures that are involved in the processing of emotional and sexual information.

Summary of Findings

Chapter 1, to be published in the Journal of Sexual Aggression, reviewed the emotional Stroop literature and encouraged researchers to consider cognitive techniques, specifically the emotional Stroop task, to assess implicit attitudes related to sexual interest. Here it was suggested that researchers and clinicians also consider the factors that may influence Stroop results before conducting an experiment using this paradigm, and the importance of putting careful consideration into the development of word stimulus sets was highlighted.

The aim of Chapter 2 was to outline the general methodology used throughout the studies discussed in the thesis and to describe the development of a new set of emotional Stroop word stimuli to be used with sexual abusers. The words were derived by individuals that work with sexual abusers, and sexual abusers' descriptions of who would be most at risk from them. Importance was placed on the need to empirically derive word stimuli that were less general in sexual interest content and more specific to the interests of sexual abusers in order to better reflect the motives,

thoughts or feelings experienced by them. It was suggested that by concentrating on the development of more appropriate word stimuli, the improvement of this measure could enable the Stroop task to adequately assess sexual interest preferences of sexual abusers and be able to differentiate between offender groups.

Chapter 3 examined the differences in mean RTs and emotional Stroop bias scores between adult sexual abusers, adult offending controls and adult non-offending controls for the Smith and Waterman (2004) and newly-derived stimulus set. The study found that overall sexual abusers took significantly longer than non-offending controls to respond to the emotional Stroop task, and that the offenders in general were slow in completing the task. Adult sexual abusers displayed a response bias to the sexual word stimuli and to the emotional/personality words that were specified to their sexual interests. These results provided further evidence that the emotional Stroop task can be a useful tool to use with sexual abusers to measure sexual interest, particularly when work on the development of appropriate word stimuli is a focus of the research.

Chapter 4 explored the utility of the emotional Stroop task with adolescent sexual abusers, adolescent non-offending controls and adolescent non-offending controls using the same emotional Stroop word sets that were used in the adult study. It was found that adolescent sexual abusers displayed a processing bias toward general sexual word content (i.e., sexual words from the Smith and Waterman stimulus set), however, the study lacked significant results and the adolescent groups were typically unresponsive to the emotional Stroop tasks. It was not surprising that the adolescents responded differently to this task compared to adults when we consider the differential response patters of adolescents on other measures such as risk assessments, PPG, and tests of executive function, and the limited amount of research

available to inform us on the specific differences between adolescent and adult sexual abusers (Beckett, 2009; Grant, 2009). Chapter 5 examined the differences between the results in Chapter 3 and Chapter 4 and provided evidence of the similarities and differences in response bias to sexual word stimuli between the groups. The formal comparison of these separate groups has confirmed that sexual abusers respond quite differently to the task than non-offending control subjects. In addition, there was evidence that adolescent groups as a whole differ quite significantly in the processing of emotional and sexual word stimuli compared to adult samples for these word stimulus sets.

Chapter 6 aimed at creating a better understanding of the cognitive processes of sexual offenders using neuroimaging. fMRI was used to examine differences in brain activation patterns between sexual abusers, and non-offending controls, in processing sexual and emotional material during the presentation of the Smith and Waterman (2004) word stimulus set. Baseline activation patterns for this word stimulus set were established for control subjects and then compared to the sexual abusers. Areas of peak activation for the non-offending controls were evident for the aggressive, positive and sexual contrasts of interest in areas that had previously been linked in neuroimaging to cognition and emotion. When the results from the controls were compared to activation patterns of the sexual abusers, the sexual abusers exhibited a pattern of deactivation in the structures that were established in the baseline. The most notable finding for this comparison was the presence of five significant peak clusters of activation exhibited by the controls and not the sexual abusers for the sexual > neutral contrast in frontal and temporal regions, including the inferior frontal gyrus and parts of the ACC.

Finally, Chapter 7 established an additional set of baseline brain activation patterns for non-offending controls to the new word stimuli that were more specific to the sexual interests of sexual abusers. This new baseline for controls was then compared to a case study of an exhibitionist. For the control subjects no significant differences were evident in the contrasts of interest from the newly-derived word categories (i.e., activations for the target words and matched words elicited similar activations). When the non-offending controls were compared with the exhibitionist, the exhibitionist displayed a pattern of deactivation in temporal and frontal/prefrontal areas of the brain for the sexual action and physical descriptor words, areas of the brain where various types of sexual abusers have been reported to display dysfunction (Cohen et al., 2002). The exhibitionist also showed a pattern of reduced inhibition for the colour-word Stroop task compared to the non-offending controls.

Limitations

Although the findings of this thesis contribute to our understanding of the cognitive and affective processing of sexual abusers, the limitations of the research should be noted. First, a clear limitation of Chapters 6 and 7 is the small sample of sexual abusers used in the studies. Another limitation is the heterogeneity of the sexual abuser groups overall. This restricted the analysis that could be conducted in terms of examining the subgroups of the samples. These limitations could not be resolved due to difficulties with participant recruitment. Access to information confirming the presence of learning disabilities, other cognitive disorders, and risk level would have been advantageous to include in the analysis as well. For example, deficiencies with reading ability or attention could have contributed to the minimal Stroop effects elicited by the adolescents, and risk level could have been examined for a relationship

with Stroop bias. Though, access to this information or confirmations regarding the presence of these variables was not accessible. Also, there are natural limitations in conducting this research in forensic settings. For example, appropriate measures should be taken to reduce the extent of noise and number of distractions that tend to accompany prison or treatment settings when conducting a study involving the emotional Stroop task. When deciding on the methods that will be used to record reaction times, the environmental setting where these studies are conducted should be a consideration. In addition, when choosing additional measures to include in the study the potential that time constraints could contribute should not be forgotten.

Implications of the Emotional Stroop

Previous studies have found that sexual abusers display a processing bias towards sexual word stimuli (Price & Hanson, 2007; Smith & Waterman, 2004). It has been suggested that the word stimuli previously used to assess the sexual interests of sexual abusers on this task are of a general sexual nature and that more should be derived that would reflect the sexual interests of sexual abusers more exclusively (Price & Hanson, 2007; Smith, 2009). The results from this thesis provide further evidence to support this view and this has important implications for the assessment of sexual interest.

Therefore, the findings in this thesis provide a contribution to our understanding of the cognitive processes of sexual abusers, and how they process emotional and sexual information. The emotional Stroop task is an information-processing task that measures the strength of representations associated with the words on a cognitive and affective level (Smith, 2009), therefore, the emotional Stroop task may have practical applications in work with sexual abusers.

For example, in Chapter 3, differences in information-processing were present between the sexual abusers and other groups. Not only did the adult sexual abusers demonstrate a processing bias for the sexual stimuli using the Smith and Waterman word stimulus set, but they displayed significant differences in response bias to the word stimuli (i.e., emotional/personality descriptors) that were specified to reflect the sexual interests of sexual abusers. Although the results were not significant between the sexual abusers and other offenders, the differential patterns in response bias across the word categories for the adult groups were interesting. More importantly, the clear differences in response patterns between the two word stimulus sets across the groups support the notion that the Stroop task may be able to measure more individualised sexual interest. Chapter 2 was an attempt at developing word stimuli that were more specific to those interests, and the response patterns between the adult groups were distinct across the different word categories in the new stimulus set. For example, the sexual abusers did not display a response bias to physical descriptors of individuals; violent offenders did not show a bias for sexual action words; and non-offending controls did not experience bias towards the emotional/personality descriptors of individuals.

The Stroop task and Adolescents

There has been little empirical evidence to support the use of self-report measures, risk assessments and PPG to measure the sexual interests of adolescent sexual abusers (Abel et al., 2004; Beckett, 2006; Grant, 2006; Seto, Lalumière, & Blanchard, 2000). Regarding the assessment of the sexual interests of adolescent sexual abusers, there was no evidence to suggest that the emotional Stroop task (using word stimuli) was an adequate measure to use for this purpose. However, the study highlighted the

importance of additional considerations that need to be taken into account in assessment efforts for adolescent samples (i.e., developmental factors). It was suggested that another variant of the emotional Stroop task might be a more appropriate approach to take when conducting future research with adolescent offending samples. For example, a picture-Stroop task (because it does not incorporate word stimuli).

The results from Chapter 5 can contribute to improving our understanding of similarities and differences that are present between adolescent and adult sexual abusers with regards to the assessment of sexual interest. This comparison provided evidence that the two groups respond quite differently to the same assessment procedure, and that it is important not to generalise assessment methodologies across adolescent and adult sexual offender groups, but to treat them differently in the assessment of sexual interests. For example, the response patterns of the adults would suggest that this tool can be used to assess sexual abusers' general bias towards sexual stimuli, and to more specific sexual interests of this subgroup. In contrast, the response patterns of the adolescent samples to the themed emotional Stroop task would suggest that another variant of the task may be more appropriate to measure specific sexual interests of this age cohort. However, a consistency found in the response patterns of both adolescent and adult sexual abusers was that they took longer to respond to the word stimuli. In fact, this finding could be generalised across the offender groups. This could suggest that there is the presence of an interaction between the effects of the stimulus salience and possible neurological impairments (i.e., impulsivity) with offender groups (Smith, 2009). Importantly, if we attribute this general slowing effect to cognitive deficiencies, this does not mean that we need to discount the Stroop bias observed between the groups. This is because Stroop bias is a

reflection of further increases in RTs and interruptions in information-processing caused by the content of the varying word categories.

The study in Chapter 6 was the first attempt made to link the behavioural findings from the emotional Stroop task with functional imaging data for sexual abuser samples. The results from Chapter 6 suggest that rather than heightened emotion for sexual or aggressive stimuli, the sexual abusers show decreased responsivity to affective stimuli. The activations that were exhibited by the controls for the new stimulus set in Chapter 7 indicate that they were processing the words on a cognitive attentional level, rather than experiencing interference in information-processing caused by additional emotional word content. This is explained by the lack of differential activations between the emotional and matched word categories. Importantly, the temporal and frontal areas of the brain where the sexual abusers displayed patterns of reduced inhibition are areas that have been suggested to play a role in behavioural disinhibition and regulating sexual behaviour (Graber et al., 1982; Hucker et al., 1986).

Future Directions

Future research should consider the limitations that were evident in the studies presented in this thesis when planning to conduct research using the emotional Stroop task. For work with adult samples, further efforts to improve the word stimulus sets would be a valuable contribution. Additionally, attempts at increasing the sample sizes of the subgroups of sexual abusers would help in strengthening the findings and determining whether the general slowing in response times and Stroop bias effects can be generalised towards sexual abusers as one group, or whether we should continue our efforts to develop a series of word lists specific to sub-groups of offenders.

Future efforts in using this tool to assess sexual interest with adolescent samples should not be abandoned. Instead, research attempts should be sure to include the ability to measure additional developmental factors or learning disabilities that may have an effect on Stroop results using this sample. An exploration of several variants of the Stroop task may be required to find an appropriate fit to measure sexual interest of individuals in this age cohort.

Finally, it would certainly be of great value to increase the sample size in work that uses fMRI and the emotional Stroop, to aid in forming links between the behavioural data and the mechanisms responsible for processing the word stimuli. The results combined from Chapter 6 and 7 suggest that sexual abusers process sexual, emotional and cognitive information differently.

Conclusions

Due to the relatively recent involvement that the emotional Stroop task has had within forensic research, the results from this study provide a contribution to our understanding of the information-processing of sexual material, and the cognitive processes that may be responsible for the results that we have observed. Clearly, the emotional Stroop task elicits differential responses for separate groups of individuals. The key seems to be to derive word stimuli that are specific to the construct in question. An overall aim of the thesis was to derive word stimuli that would be able to differentiate between offender types, specific to sexual interest. Although this aim was not met, the results would suggest that the Smith and Waterman (2004) stimuli are tapping into sexual interest on a more general level and that the new word stimuli appear to be generating more varied response patterns specific to individual interests. This study, therefore, brought us one step closer to deriving words that are reflective

of sexual interests specific to sexual abuser populations. Given the importance attributed to sexual interests in the assessment and treatment of sexual abusers, further development of this information-processing tool to be able to measure the implicit nature of deviant sexual interests would be a valuable contribution to theory and practice in work with sexual abusers.

Appendix A

Emotional Stroop Words by Category (Smith & Waterman, 2004)

Neutral	Aggression	Positive	Negative	Sexual	Colour
Door	Rage	Devotion	Abandoned	Rape	Purple
Group	Anger	Affection	Abused	Lust	Grey
Chair	Tear	Admire	Afraid	Slap	Flesh
Telephone	Assault	Euphoric	Aggressive	Victim	Pink
Dog	Kick	Fond	Agony	Slut	Orange
Coat	Shout	Grateful	Angry	Whore	Scarlet
Sofa	Punch	Tolerant	Arrogant	Fuck	Maroon
Bag	Hate	Affectionate	Bad	Bitch	Crimson
Diary	Argue	Amused	Bitchy	Child	White
Newspaper	Temper	Love	Crazy	Woman	Black
Eat	Fight	Joy	Cruelty	Force	Cyan
Oven	Kill	Proud	Detest	Penetrate	Brown
Floor	Punish	Fond	Envy	Control	Tin
Shopping	Annoyed	Funny	Fear	Schoolgirl	Bronze
Umbrella	Guilt	Glad	Frightened	Power	Mauve
Windy	Scream	Comfortable	Fury	Dominance	Gold
Radio	Crush	Beloved	Glum	Force	Silver
Painting	Slash	Calm	Greed	Oral	Jade
Milk	Smash	Peace	Hateful	Abuse	Topaz
School	Cut	Daring	Spiteful	Incest	Emerald
Ball	Wound	Cheerful	Suspicious	Molest	Purple
Pencil	Injure	Warm	Rage	Man	Grey
Cigarette	Threaten	Protective	Pain	Grope	Flesh
Football	Knife	Hope	Lose	Prostitute	Pink
Shoe	Wound	Lively	Misery	Anal	Orange

Appendix B

Emotional Stroop Words (Newly Derived)

Emotional/ Personality Descriptors (EPD)		Matched EPD		Sexual Actions (SA)	Matched SA	Physical Descriptors (PD)	Matched PD
Angelic	Vulnerable	Festive	Profitable	Cuddle	Darken	Attractive	Occasional
Bright	Caring	Annual	Glassy	Fondle	Invent	Blonde	Costly
Bubbly	Cute	Chalky	Hazy	Kiss	Chip	Bum	Ink
Childish	Devious	Tropical	Spatial	Licking	Shuffle	Fit	Ill
Friendly	Lively	Electric	Linear	Playing	Measure	Little	People
Innocent	Loving	Accurate	Sturdy	Pretending	Collecting	Petite	Breezy
Lonely	Promiscuous	Plastic	Translucent	Stroking	Rambling	Sexy	Oval
Mature	Pure	Hungry	Busy	Sucking	Packing	Short	Level
Nice	Seductive	Airy	Forgetful	Teaching	Speaking	Slim	Loud
Outgoing	Sensual	Seasonal	Prickly	Tempt	Brake	Small	Right
Pleasant	Simple	Concrete	Likely	Tickling	Knitting	Tall	Slow
Polite	Sneaky	Cloudy	Catchy	Touching	Printing	Thin	Deep
Scared	Special	Speedy	Winding	Caress	Paddle	Gorgeous	Speckled
Shy	Sweet	Icy	Solid	Wanking	Raining		
Trusting	Flirtatious	Spacious	Descriptive	Wrestling	Gardening		
				Tease	Cough		

Appendix C

RT and Emotional Stroop Bias Scores Medians in Milliseconds for the Smith and Waterman (2004) Word Stimulus Set: Adult Samples

Word Category	Adult Group	Median RT's (ms)	Median Stroop Bias Scores, ms
Neutral	Adult sexual abusers	749.28	
	Adult offending controls	681.40	
	Adult non-offending controls	593.88	
Colour	Adult sexual abusers	771.96	45.63
	Adult offending controls	730.40	27.32
	Adult non-offending controls	604.06	17.28
Positive	Adult sexual abusers	764.44	29.54
	Adult offending controls	692.56	13.73
	Adult non-offending controls	593.86	-15.93
Negative	Adult sexual abusers	793.25	48.80
	Adult offending controls	681.68	17.60
	Adult non-offending controls	593.56	7.14
Aggression	Adult sexual abusers	785.68	19.05
	Adult offending controls	727.36	-7.28
	Adult non-offending controls	596.37	-2.24
Sexual	Adult sexual abusers	782.36	48.98
	Adult offending controls	715.26	33.86
	Adult non-offending controls	595.31	1.32

Appendix D

RT and Emotional Stroop Bias Score Medians in Milliseconds for the Newly-Derived Word Stimulus Set: Adult Samples

Word Category	Adult Group	Median RT, ms	Median Stroop Bias Score, ms
EPD	Adult sexual abusers	749.17	-1.37
	Adult offending controls	640.63	19.77
	Adult non-offending controls	570.33	-20.49
MEPD	Adult sexual abusers	715.62	
	Adult offending controls	657.50	
	Adult non-offending controls	570.17	
SA	Adult sexual abusers	743.80	21.60
	Adult offending controls	661.56	8.50
	Adult non-offending controls	583.27	5.69
MSA	Adult sexual abusers	715.19	
	Adult offending controls	679.13	
	Adult non-offending controls	570.81	
PD	Adult sexual abusers	742.31	16.64
	Adult offending controls	713.54	6.36
	Adult non-offending controls	573.82	15.23
MPD	Adult sexual abusers	782.77	
	Adult offending controls	660.08	
	Adult non-offending controls	582.25	

Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories.

Appendix E

RT and Emotional Stroop Bias Scores Medians in Milliseconds for the Smith and Waterman (2004) Word Stimulus Set: Adolescent Samples

Word Category	Adolescent Group	Median RT, ms	Median Stroop Bias scores, ms
Neutral	Adolescent sexual abusers	798.01	
	Adolescent offending controls	692.43	
	Adolescent non-offending controls	670.24	
Colour	Adolescent sexual abusers	790.79	-9.29
	Adolescent offending controls	699.07	39.90
	Adolescent non-offending controls	642.83	10.29
Positive	Adolescent sexual abusers	797.22	-14.17
	Adolescent offending controls	669.57	-13.93
	Adolescent non-offending controls	648.26	-0.53
Negative	Adolescent sexual abusers	785.08	-6.19
	Adolescent offending controls	663.31	14.90
	Adolescent non-offending controls	646.38	1.52
Aggression	Adolescent sexual abusers	794.13	1.44
	Adolescent offending controls	677.35	-18.25
	Adolescent non-offending controls	631.83	-26.60
Sexual	Adolescent sexual abusers	879.39	38.33
	Adolescent offending controls	652.53	1.49
	Adolescent non-offending controls	632.55	-11.95

Appendix F

RT and Emotional Stroop Bias Score Medians in Milliseconds for the Newly-Derived Word Stimulus Set: Adolescent Sample

Word Category	Adolescent Group	Median RT, ms	Median Stroop Bias Score, ms
EPD	Adolescent sexual abusers	753.44	-2.68
	Adolescent offending controls	642.79	-6.23
	Adolescent non-offending controls	663.58	-0.07
MEPD	Adolescent sexual abusers	784.85	
	Adolescent offending controls	654.43	
	Adolescent non-offending controls	635.73	
SA	Adolescent sexual abusers	734.25	-7.81
	Adolescent offending controls	683.75	-7.06
	Adolescent non-offending controls	641.69	4.38
MSA	Adolescent sexual abusers	739.64	
	Adolescent offending controls	661.06	
	Adolescent non-offending controls	643.88	
PD	Adolescent sexual abusers	781.28	11.52
	Adolescent offending controls	668.90	3.99
	Adolescent non-offending controls	675.23	4.69
MPD	Adolescent sexual abusers	717.56	
	Adolescent offending controls	661.50	
	Adolescent non-offending controls	641.50	

Note. Abbreviations represent the following: emotional/personality descriptors (EPD); sexual actions (SA); physical descriptors (PD). 'M' represents matched categories.

References

Abel, G. G., Gore, D. K., Holland, C. L., Camp, N., Becker, I. V., & Rathner, J. (1989). The measurement of the cognitive distortions of child molesters. *Annals of Sex Research, 2,* 135-152.

Abel, G. G., Jordan, A., Hand, C. G., Holland, L. A., & Phipps, A. (2001). Classification models of child molesters utilizing the Abel Assessment for sexual interest. *Child Abuse and Neglect, 25,* 703-718.

Abel, G. G., Jordan, A., Rouleau, J. L., Emerick, R., Barboza-Whitehead, S., & Osborn, C. (2004). Use of visual reaction time to assess male adolescents who molest children. *Sexual Abuse: A Journal of Research and Treatment, 16,* 255-265.

Adleman, N. E., Menon, V., Blasey, C. M., White, C. D., Warsofsky, I. S., Glover, G. H. et al. (2002). A developmental fMRI study of the Stroop color-word task. *NeuroImage, 16,* 61-75.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR* (4th ed., text rev.). Washington, DC:Author.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV* (4th ed.). Washington, DC:Author.

Andrade, J. T., Vincent, G. M., & Saleh, F. M. (2006). Juvenile sex offenders: A complex population. *Journal of Forensic Sciences, 51*, 163-167.

Ashley, V., & Swick, D. (2009). Consequences of emotional stimuli: Age differences on pure and mixed blocks of the emotional Stroop. *Behavioral and Brain Functions, 5,* 14-25.

Augustine, J. R. (1996). Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Research Reviews, 22*, 229-244.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283-316.

Banich, M. T., Crowley, T. J., Thompson, L. L., Jacobson, B. L., Liu, X., Raymond, K. M. et al. (2007). Brain activation during the Stroop task in adolescents with severe substance and conduct problems: A pilot study. *Drug and Alcohol Dependence, 90,* 175-182.

Banich, M. T., Milham, M. P., Atchley, R., Cohen, N. J., Webb, A., Wszalek, T., …Magin, R. (2000). fMRI studies of Stroop tasks reveal unique roles of anterior and posterior brain systems in attentional selection. *Journal of Cognitive Neuroscience, 12,* 988-1000.

Banich, M. T., Milham, M. P., Jacobson, B. L., Webb, A., Wszalek, T., Cohen, N. J., & Kramer, A. F. (2001). Attentional selection and the processing of task-irrelevant information: Insights from fMRI examinations of the Stroop task. In C. Casanova & M. Ptito (Eds.), *Vision: From neuroses to cognition* (Progress in Brain Research, Vol. 134, pp. 459-470). Amsterdam: Elsevier.

Barbaree, H. E. (1990). Stimulus control of sexual arousal: Its role in sexual assault. InW. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 115–138). New York: Plenum Press.

Barbaree, H. E., & Marshall, W. L. (2006). An introduction to the juvenile sex offender: Terms, concepts, and definitions. In H. E. Barbaree & W. L. Marshall (Eds.), *The Juvenile Sex Offender* (pp. 1-18). New Tork: The Guilford Press.

Barlow, D. H. (1977). Assessment of sexual behavior, in A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of Behavioral Assessment.* New York: John Wiley & Sons, Ltd.

Becker, J. V., Hunter, J. A., Goodwin, D., Kaplan, M. S., & Martinez, D. (1992). Test-retest reliability of audiotaped phallometric stimuli with adolescent sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 5,* 45-51.

Becker, J. V., Kaplan, M. S., Cunningham-Rathner, J., & Kavoussi, R. J. (1986). Characteristics of adolescent incest sexual perpetrators: Preliminary findings. *Journal of Family Violence, 1,* 85-97.

Beckett, R. (2006). Risk Prediction, decision making and evaluation of adolescent sexual abusers. In M. Erooga & H. Masson (Eds.), *Children and Young People Who Sexually Abuse Others. Challenges and Responses (2nd ed)* (pp. 215-233). London: Routledge.

Beckett, R. (1999). Evaluation of adolescent sexual abusers. In M. Erooga & H. Masson (Eds.), *Children and Young People Who Sexually Abuse Others. Challenges and Responses* (pp. 204-224). London: Routledge.

Beckett, R.C. (1987). *Beliefs About Children Scale (BACS)*. Oxford: Oxford Forensic Service.

Beckett, R. C., Beech, A. R., Fisher, D., & Fordham, A. S. (1994). *Community-Based Treatment For Sex Offenders: An Evaluation of Seven Treatment Programmes*, London: Home Office

Beckett, R. C., Fisher, D., Mann, R. E., & Thornton, D. (1996). Relapse prevention questionnaire and interview. In H. Eldridge (Ed.), *Therapists guide for maintaining change: Relapse prevention manual for adult male perpetrators of child sexual abuse.* Thousand Oaks, CA: Sage.

Beckmann, C., Jenkinson, M., & Smith., S. M. (2003). General multi-level linear modelling for group analysis in FMRI. *NeuroImage, 20*, 1052-1063.

Beckmann, C. F. & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging, 23*, 137-152.

Beech, A. R. (1998). A psychometric typology of child abusers. *International Journal of Offender Therapy and Comparative Criminology, 42*, 319-339.

Beech, A. R., Kalmus, E., Tipper, S. P., Baudouin, J.-Y., Humphreys, G. W., & Flak, V. E. (2008). Children induce an enhanced attentional blink in child molesters. *Psychological Assessment, 20,* 397-402.

Beech, A. R., Oliver, C., Fisher, D., & Beckett, R. (2006). STEP 4: The Sex Offender Treatment Programme in Prison: Addressing the Offending Behaviour of Rapists and Sexual Murderers. Birmingham: available at www.hmprisonservice.gov.uk/assets/documents/100012DBStep_4_SOTP_report_2005.pdf

Bench, C. J., Frith, C. D., Grasby, P. M., Friston, K. J., Paulesu, E., Frackowiak, R. S .J., & Dolan, R . J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia, 31,* 907-922.

Bielak, A. A. M., Mansueti, L., Strauss, E., & Dixon, R. A. (2006). Performance on the Hayling and Brixton tests in older adults: Norms and correlates. *Archives of Clinical Neuropsychology, 21,* 141-149.

Blanchard, R., Christensen, B. K., Strong, S. M., Cantor, J. M., Kuban, M. E., Klassen, P., …Blak, T. (2002). Retrospective self-reports of childhood accidents causing unconsciousness in phallometrically diagnosed pedophiles. *Archives of Sexual Behavior, 31,* 511-526.

Blanchard, R., Kuban, M. E., Klassen, P. E., Dickey, R., Christensen, B. K., Cantor, J. M., & Blak, T. (2003). Self-reported head injuries before and after age 13 in pedophilic and non-pedophilic men referred for clinical assessment. *Archives of Sexual Behavior, 32,* 573-581.

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in the anterior cingulate cortex. *Nature, 402,* 179-181.

Boonstra, A. M., Oosterlaan, J., Sergeant, J. A., & Buitelaar, J. K. (2005). Executive functioning in adult ADHD: a meta-analytic review. *Psychological Medicine, 35,* 1097-1108.

Bremner, J. D., Vermetten, E., Vythilingam, M., Afzal, N., Schmahl, C., Elzinga, B., & Charney, D. S. (2004). Neural correlates of the classic color emotional Stroop in women with abuse-related post-traumatic stress disorder. *Biological Psychiatry, 55,* 612-620.

Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests*. London: Harcourt Assessment.

Burt, J. S. (2002). Why do non-color words interfere with color naming? *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1019-1038.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215-222.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging-validation study with functional MRI. *Human Brain Mapping, 6,* 270-282.

Buss, A. H., & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63*, 452-459.

Cantor, J. M., Blanchard, R., Christensen, B. K., Dickey, R., Klassen, P. E., Beckstead, A. L., …Kuban, M. E. (2004). Intelligence, memory, and handedness in pedophilia. *Neuropsychology, 18,* 3-14.

Cantor, J. M., Blanchard, R., Robichaud, L. K., & Christensen, B. K. (2005a). Quantitative reanalysis of aggregate data on IQ in sexual offenders. *Psychological Bulletin, 131,* 555-568.

Cantor, J. M., Klassen, P. E., Dickey, R., Christensen, B. K., Kuban, M. E., Blak, T., …Blanchard, R. (2005b). Handedness in pedophilia and hebephilia. *Archives of Sexual Behavior, 34,* 447-459.

Cantor, J. M., Kuban, M. E., Blak, T., Klassen, P. E., Dickey, R., & Blanchard, R. (2006). Grade failure and special education placement in sexual offenders' educational histories. *Archives of Sexual Behavior, 35,* 743-751.

Cantor, J. M., Noor, K., Christensen, B. K., Zipursky, R. B., Barbaree, H. E., Dickey, R., …Blanchard, R. (2008). Cerebral white matter deficiencies in pedophilic men. *Journal of Psychiatric Research, 42,* 167-183.

Carroll, A., Hemingway, F., Bower, J., Ashman, A., Houghton, S., & Durkin, K. (2006). Impulsivity in juvenile delinquency: Differences among early-onset, late onset, and non-offenders. *Journal of Youth and Adolescence, 35,* 519-529.

Carter, C., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280,* 747-749.

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An H215O study of Stroop task performance. *Neuroimage, 2,* 264-272.

Cassiday, K. L., McNally, R. J., & Zeitlin, S. B. (1992). Cognitive processing of trauma cues in rape victims with post-traumatic stress. *Cognitive Therapy and Research, 16,* 283-295.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind, 11,* 63-65.

Christiansen, H., & Oades, R. D. (2009). Negative priming within a Stroop task in children and adolescents with attention-deficit hyperactivity disorder, their siblings, and independent controls. *Journal of Attention Disorders, 13,* 497-504.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332-361.

Cohen, L. J., Nikiforov, K., Gans, S., Poznansky, O., McGeoch, P., Weaver, C., et al. (2002). Heterosexual male perpetrators of childhood sexual abuse: A preliminary neuropsychiatric model. *Psychiatric Quarterly, 73*, 313–336.

Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test on childhood, adulthood and aging. *Journal of Genetic Psychology, 100,* 47-53.

Compton, R. J., Banich, M. T., Mohanty, A., Milhem, M., Herrington, J., Miller, G.A., …Heller, W. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional Stroop tasks. *Cognitive, Affective and Behavioral Neuroscience, 3,* 81-96.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus driven attention in the brain. *Nature Reviews, 3,* 201-215.

Cox, W. M., Fadardi, J. S., & Pathos, E. M. (2006). The Addiction-Stroop test: Theoretical considerations and procedural recommendations. *Psychological Bulletin, 132,* 443-476.

Craig, L., Browne, K. D., & Beech, A. R. (2008). *Assessing risk in sex offender: A practitioner's guide*. Chichester, UK: Wiley.

Dalgleish, T. (2005). Putting some feeling into it – the conceptual and empirical relationships between the classic and emotional Stroop tasks: A commentary on Algom, Chajut and Lev (2004). *Journal of Experimental Psychology: General, 134*, 585-591.

Dalgleish, T. (1995). Performance on the emotional Stroop task in groups of anxious, expert, and control subjects: A comparison of computer and card presentation. *Cognition & Emotion, 9,* 341-362.

Dolan, M., Millington, J., & Park I. (2002). Personality and neuropsychological function in violent, sexual, and arson offenders. *Medicine, Science, and the Law, 42*, 34–43.

Dolan, M. C., & Rennie, C. E. (2007). Is juvenile psychopathy associated with low anxiety and fear in conduct-disordered male offenders? *Journal of Anxiety Disorders, 21,* 1028-1038.

Dubner, A. E., & Motta, R. W. (1999). Sexually and physically abused foster care children and posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 67,* 367-373.

Dunbar, K. N., & MacLeod, C. M. (1984). A horse race of a different color: Stroop interference patterns with transformed words. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 622-639.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale (BPVS)*. Hampshire, UK: Pardy and Son.

Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interaction between emotion and cognition. *Cognition and Emotion, 12,* 353-385.

Egloff, B., & Schmukle, S. C. (2004). Gender differences in implicit and explicit anxiety measures. *Personality and Individual Differences, 36*, 1807-1815.

Egner, T., & Hirsch, J. (2005). The neural correlates and functional integration of cognitive control in a Stroop task. *NeuroImage, 24,* 539-547.

Erooga, M., & Masson, H. (2006). Children and young people with sexually harmful or abusive behaviours: Underpinning knowledge, principles, approaches and service provision. In M. Erooga & H. Masson (Eds.), *Children and Young People Who Sexually Abuse Others. Challenges and Responses (2nd ed)* (pp. 215-233). London: Routledge.

Ersche, K. D., Bullmore, E. T., Craig, K. J., Shabbir, S. S., Abott, S., Müller, U., …Robbins, T. W. (2010). Influence of compulsivity of drug abuse on dopaminergic modulation of attentional bias in stimulant dependence. *Archives of General Psychiatry, 67*, 632-644.

Field, A. (2005). Discovering Statistics Using SPSS (2nd ed.). London: Sage.

Foa, E. B., Feske, U., Murdock, T. B., Kozak, M. J., & McCarthy, P. R. (1991). Processing of threat-related information in rape victims. *Journal of Abnormal Psychology, 100*, 156-162.

Flak, V. (2009). *Evaluation of sexual interest in child sex offenders using a computerised measure*. Unpublished PhD Thesis, University of Birmingham, Birmingham, United Kingdom.

Flak, V., Beech, A. R., & Fisher, D. (2006). Forensic assessment of deviant sexual interests: The current position. *Issues in Forensic Psychology,* 6, 70-83.

Foa, E. B., Feske, U., Murdock, T. B., Kozak, M. J., & McCarthy, P. R. (1991). Processing of threat-related information in rape victims. *Journal of Abnormal Psychology, 100,* 156-162.

Forth, A. E., & Hare, R. D. (1989). The contingent negative variation in psychopaths. *Psychophysiology, 26,* 676-682.

Fraisse, P. (1969). Why is naming longer than reading? *Acta Psychologica, 30,* 96-103.

Freeman-Longo, R. E., Bird, S., Stevenson, W. F., & Fiske, J. A. (1995). *1994 Nationwide survey of treatment programs & models: Serving abuse reactive children and adolescent & adult sexual offenders.* Orwell, VT: Safer Society Press.

Gallagher-Duffy, J., Mackay, S., Duffy, J., Sullivan-Thomas, M., & Peterson-Badali, M. (2009). The pictorial fire Stroop: A measure of processing bias for fire-related stimuli. *Journal of Abnormal Child Psychology, 37*, 1165-1176.

Gawronski, B., & Bodenhousen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132,* 692-731.

Geer, J. H., Estupinan, L.A., & Manguno-Mire, G. M. (2000). Empathy, social skills and relevant cognitive processes in rapists and child molesters. *Aggression and Violent Behavior, 5,* 99-126.

Gerhold, C. K., Browne, K. D., & Beckett, R. (2007). Predicting recidivism in adolescent sexual offenders. *Aggression and Violent Behavior, 12*, 427-438.

Geurtz, H. M., Verte, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder in autism? *Journal of Child Psychology and Psychiatry, 45,* 836-854.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 875 -894.

Golden, C. J. (1978). *Stroop Colour* and*Word Test: A manual for clinical* and *experimental uses*. Stoelting Company, Illinois.

Graber, B., Hartmann, K., Coffman, J. A., Huey, C.J., & Golden, C.J. (1982). Brain damage among mentally disordered sex offenders. *Journal of Forensic Sciences, 27*, 125–34.

Grant, H. (2006). Assessment issues in relation to young people who have sexually abusive behaviour. In M. Erooga & H. Masson (Eds.), *Children and Young People Who Sexually Abuse Others. Challenges and Responses (2nd ed)* (pp. 67-76). London: Routledge.

Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121,* 480-506.

Gray, N. S., Brown, A. S., MacCulloch, M. J., Smith, J. & Snowden, R. J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology, 114,* 304-8.

Gress, C. L. Z., & Laws, D. R. (2009). Cognitive modelling of sexual arousal and interest: Choice reaction time measures. In D. Thornton & D. R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 177-201). Chichester, UK: Wiley.

Gross, T. M., Jarvik, M. E., & Rosenblatt, M. R. (1993). Nicotine abstinence produces content-specific Stroop interference. *Psychopharmacology, 110,* 333-336.

Hafer, C. L. (2000). Do innocent victims threaten the Belief in a Just World? Evidence from a modified Stroop task. *Journal of Personality and Social Psychology, 79,* 165-173.

Hanson, R. K. & Bussiere, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. [References]. *Journal of Consulting and Clinical Psychology, 66,* 348-362.

Hanson, R. K., Harris, A. J. R., Scott, T.-L., & Helmus, L. (2007). *Assessing the Risk of Sexual Offenders on Community Supervision: The Dynamic Supervision Project.* Public Safety Canada, Ottawa.

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of the Static-2002.* (User Report No. 2003-01). Ottawa, ON: Solicitor General Canada.

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial risk scales. *Law and Human Behavior, 24,* 119-136.

Hare, R. D. (1978). Electrodermal and cardiovascular correlates of psychopathy. In R. D.Hare & D. Schalling (Eds.), *Psychopathic behaviour: Approaches to research* (pp. 107-143). Chichester, UK: Wiley.

Harpur, T. J., & Hare, R. D. (1990). Psychopathy and attention. In J.Enns (Ed.), *The development of attention: Research and theory* (pp. 429-444). Amsterdam: North-Holland.

Heckers, S., Weiss, A. P., Deckersbach, T., Goff, D., Morecraft, R. J., & Bush, G. (2004). Anterior cingulate cortex activation during cognitive interference in Schizophrenia. *American Journal of Psychiatry, 161,* 707-715.

Herba, C. M., Tranah, T., Rubia, K., & Yule, W. (2010). Conduct problems in adolescence: Three domains of inhibition and effect of gender. *Developmental Neuropsychology, 30*, 659-695.

Hiatt, K. D., Schmitt, W. A., & Newman, J. P. (2004). Stroop tasks reveal abnormal selective attention among psychopathic offenders. *Neuropsychology, 18,* 50-59.

Hill, A. B., & Knowles, T. H. (1991). Depression and the emotional Stroop task. *Personality and Individual Differences, 12,* 481-485.

Holle, C., Neely, J. H., & Heimberg, R. G. (1997). The effects of blocked versus random presentation and semantic relatedness of stimulus words in response to a modified Stroop task among social phobics. *Cognitive Therapy and Research, 6,* 681-697.

Holmes, A. J., & Pizzagalli, D. A. (2007). Task feedback effects on conflict monitoring and executive control: Relationship to subclinical measures of depression. *Emotion, 7,* 68-76.

Houx, P. J., Jolles, J., & Vreeling, F. W. (1993). Stroop interference: Aging effects assessed with the Stroop color-word test. *Experimental Aging Research, 19,* 209-224.

Hucker, S., Langevin, R., Wortzman, G., Bain, J., Handy, L., Chambers, J., et al. (1986). Neuropsychological impairment in pedophiles. *Canadian Journal of Behavioral Science, 18*, 440–8.

Hunter, J. A., Figueredo, A. J., Malamuth, N. M., & Becker, J. V. (2003). Juvenile sex offenders: Toward the development of a typology. *Sexual Abuse: A Journal of Research and Treatment*, *15,* 27–48.

Jenkinson, M., Bannister, P., Brady, M. & Smith, S. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage, 17*, 825-841.

Jenkinson, M., & Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis, 5*, 143-156.

Johnson, W., Bouchard, T. J., Segal, N. L., Keyes, M., & Samuels, J. (2003). The Stroop Color-word test: Genetic and environmental influences; Reading, mental ability and personality correlates. *Journal of Educational Psychology, 95,* 58-65.

Johnson, M. H., & Hasher, L. (1987). Human learning and memory. *Annual Review of Psychology, 38,* 631-668.

Jorgenson, C., Davis, J., Opella, J., & Angerstein, G. (1981). Hemispheric asymmetry in the processing of Stroop stimuli: An examination of age and gender differences. *International Journal of Neuroscience, 13,* 113-116.

Jorgenson, C., Davis, J., Opella, J., & Angerstein, G. (1980). Hemispheric-asymmetry in the processing of Stroop-stimuli: An examination of gender, hand preference, and language differences. *International Journal of Neuroscience, 11,* 165-169.

Kalmus, E., & Beech, A. R. (2005). Forensic assessment of sexual interest: A review. *Aggression and Violent Behavior, 10*, 193-217.

Kerns, J. G., Cohen, J. D., Macdonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023-1026.

Kessinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective and Behavioral Neuroscience, 6,* 110-126.

Kieley, J. M., & Hartley, A. A. (1997). Age-related equivalence of identity suppression in the Stroop color-word task. *Psychology and Aging, 12,* 22-29.

Klein, G. S. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology, 77,* 576-588.

Klieger, D. M., & Cordner, M. D. (1990). The Stroop task as a measure of construct accessibility in depression. *Personality and Individual Differences, 11,* 19-27.

Kolárský, A., Freund, K., Machek, J., Polák, O. (1967). Male sexual deviations: association with early temporal lobe damage. *Archives of General Psychiatry, 17*, 735–43.

Kosson, D. S. (1998). Divided visual attention in psychopathic and nonpsychopathic offenders. *Personality and Individual Differences, 95,* 252-256.

Kosson, D. S. (1996). Psychopathy and dual-task performance under focusing conditions. *Journal of Abnormal Psychology, 105,* 391-400.

Koven, N. S., Heller, W., Banich, M. T., & Miller, G. A. (2003). Relationships of distinct affective dimensions to performance on an emotional Stroop task. *Cognitive Therapy and Research, 27,* 671-680.

Kramer, D. A., & Goldman, M. S. (2003). Using a modified Stroop task to implicitly discern the cognitive organization of alcohol expectancy. *Journal of Abnormal Psychology, 112,* 171-175.

Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Laming, D. (1979). Autocorrelation of choice-reaction times. *Acta Psychologica, 43,* 381-412.

Lane, R. D., Chua, P. M. L., & Dolan, R. J. (1999). Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia*, *37*, 989-997.

Lane, R. D., Fink, G. R., Chau, P. M-L., & Dolan, R. J. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport, 8,* 3969-3972.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., & Schwartz, G. E., (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia, 35,* 1437-1444.

Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B., & Nangia, V. (1998). Emotional arousal and activation of the visual cortex : An fMRI analysis. *Psychophysiology, 35,* 199-210.

Langevin, R., Lang, R. A., Wortzman, G., Frenzel, R. R., & Wright P. (1989). An examination of brain damage and dysfunction in genital exhibitionists. *Annals of Sex Research, 2*, 77–87.

Långström, N. (2002). Long-term follow-up of criminal recidivism in young sex offenders: Temporal patterns and risk factors. *Psychology, Crime & Law, 8,* 41-58.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6,* 62-72.

Laws, D. R. (2009). Penile plethysmography: Strengths, limitations, innovations. In D. Thornton & D. R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 7-29). Chichester, UK: Wiley.

Laws, D. R. (2003). Penile plethysmography: Will we ever get it right? In T. Ward, D. R. Laws & S. M. Hudson (Eds.), *Sexual Deviance: Issues and Controversies* (pp. 82-102). Thousand Oaks, CA: Sage.

Laws, D. R., & Gress, C. L. Z. (2004). Seeing things differently: the viewing time alternative to penile plethysmography. *Legal and Criminological Psychology, 9,* 183-196.

Leung, H., Skudlarski, P., Gatenby, J. C., Peterson, B. S., & Gore, J. C. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cerebral Cortex, 10,* 552-560.

Lilly, R., Cummings, J. L., Benson, D. F., Frankel, M. (1983). The human Klüver-Bucy syndrome. *Neurology, 33*, 1141–1145.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12,* 523-553.

Logan, A. C. & Goetsch, V. L. (1993). Attention to external threat cues in anxiety states. *Clinical Psychology Review, 13,* 541-559.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163-203.

MacLeod, C., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14,* 126-135.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4,* 383-391.

Malhi, G. S., Lagopoulos, J., Sachdev, P. S., Ivanovski, B., & Shnier, R. (2005). An emotional Stroop functional MRI study of euthymic bipolar disorder. *Bipolar Disorder, 7* (suppl 5), 58-69.

Marshall, W. L., Anderson, D., & Fernandez, Y. (1999). *Cognitive Behavioural Treatment of Sexual Offenders*. John Wiley & Sons, Inc., New York.

Marshall, W. L., & Fernandez, Y. M. (2003). Sexual preferences: Are they useful in the assessment and treatment of sexual offenders? *Aggression and Violent Behavior, 8*, 131-143.

Mathews, A. M., Mogg, K., Kentish, J., & Eysenck, M. (1995). Effect of psychological treatment on cognitive bias in generalized anxiety disorder. *Behavioural Research and Therapy, 33,* 293-303.

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the Emotional Stroop task. *Journal of Experimental Psychology, 21,* 1595-1607.

McNally, R. J. (1998). Information processing abnormalities in anxiety disorders: Implications for cognitive neuroscience. *Cognition and Emotion, 12,* 479-495.

McNally, R. J., Amir, N., Louro, C. E., Lukach, B. M., Riemann, B. C., & Calamari, J. E. (1994). Cognitive processing of idiographic emotional information in panic disorder. *Behavioural Research and Therapy, 32,* 119-122.

McNally, R. J., Clancy, S. A., Schacter, D. L., & Pitman, R. K. (2000). Cognitive processing of trauma cues in adults reporting repressed, recovered, or continuous memories of childhood sexual abuse. *Journal of Abnormal Psychology, 109,* 355-359.

McNally, R. J., Metzger, L. J., Lasko, N. B., Clancy, S. A., & Pitman, R. K. (1998). Directed forgetting of trauma cues in adult survivors of childhood sexual abuse with and without posttraumatic stress disorder. *Journal of Abnormal Psychology, 107,* 596-601.

McNally, R. J., Riemann, B. C., & Kim, E. (1990). Selective processing of threat cues in panic disorder. *Behavioural Research and Therapy, 28,* 407-412.

McNally, R. J., Riemann, B. C., Louro, C. E., Lukach, B. M., & Kim, E. (1992). Cognitive processing of emotional information in panic disorder. *Behavioural Research and Therapy, 30,* 143-149.

Milhem, M. P., Erikson, K. L., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., & Cohen, N. J. (2002). Attentional control in the aging brain: Insights from an fMRI study of the Stroop task. *Brain and Cognition, 49,* 277-296.

Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 43,* 907-912.

Moffitt, T. E. (1990). The neuropsychology of delinquency: A critical review of theory and research. *Crime and Justice: An Annual Review of Research, 12*, 99-169.

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. *Journal of Abnormal Psychology, 102,* 304-311.

Moradi, A. R., Taghavi, M. R., Neshat Doost, H. T., Yule, W., & Dagleish, T. (1999). Performance of children and adolescents with PTSD on Stroop the colour-naming task. *Psychological Medicine, 29,* 415-419.

Morton, J. S., & Chambers, S. M. (1973). Selective attention to words and colours. *Quarterly Journal of Experimental Psychology, 25,* 387-397.

Mutter, S. A., Naylor, J. C., & Patterson, E. R. (2005). The effects of age and task context on Stroop task performance. *Memory and Cognition, 33,* 514-530.

Nunes, K. L., Firestone, P., & Baldwin, M. W. (2007). Indirect assessment of cognitions of child sexual abusers with the implicit association test. *Criminal Justice and Behaviour, 34,* 454-475.

O'Donohue, W., & Letourneau, E. (1992). The psychometric properties of the penile tumescence assessment of child molesters. *Journal of Psychopathology and Behavioural Assessment, 14,* 123-174.

Ó Ciardha, C., & Gormley, M. (2009). Developing clinically useful implicit tasks of sexual interest: What are they measuring? In C. Ó Ciardha (Chair), *Indirect Measures of Assessment: Clinical Applications and Theoretical Implications.* Symposium conducted at the 28th annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Dallas, TX.

Ó Ciardha, C., & Gormley, M. (2009b). Comparing two implicit cognitive measures of sexual interest: A pictoral modified Stroop task and the implicit associations test. In D. Thornton & D.R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 177-201). Chichester, UK: Wiley.

Oliver, B. (2007). Three steps to reducing child molestation by adolescents. *Child Abuse and Neglect, 37,* 683-689.

Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. E. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences USA, 87,* 256-259.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behavior. *Neuropsychologia, 12,* 323-330.

Petty, R. E., & Briñol, P. (2006).A meta-cognitive approach to "implicit" and "explicit" evaluations: Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin, 132,* 740-744.

Petty, R. E., Fazio, R. H., & Brinol, P. (2008). *Attitudes: Insights from the New Implicit Measures*. London, U.K.: Psychology Press.

Ponseti, J., Bosinski, H. A., Wolff, S., Peller, M., Jansen, O., Mehdorn, H. M., et al. (2006). A functional endophenotype for sexual orientation in humans. *NeuroImage*, *33*, 825–33.

Ponseti, J., Vaih-Koch, S. R., Bosinski, H. A. G. (2001, July). *Sex offenders versus non-sex offenders: differences in impulsivity and attention-deficit/hyperactivity disorder?* Poster presentation at the 28th Annual Meeting of the International Academy of Sex Research, Montréal, Québec, Canada.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R.L. Solso (Ed.), *Information-processing and cognition: The Loyola Symposium* (pp. 55-85). Hillsdale; NJ: Erlbaum.

Prentky, R. A., & Righthand, S. C. (2003). *Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II): Manual*. Available online at: www.csom.org.

Price, S. A., Beech, A. R., & Allen, H. A. (2010, October). *Stroop interference effects as measured through fMRI analysis amongst sexual abusers and controls*. Poster session presented at the annual Association for the Treatment of Sexual Abusers (ATSA) conference, Phoenix, AZ.

Price, S. A., & Hanson, R. K. (2007). A modified Stroop task with sexual offenders: Replication of a study. *Journal of Sexual Aggression, 13,* 203-216.

Rabbitt, P. M. A. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology, 71,* 264-272.

Rasmussen, L. A. (1999). Factors related to recidivism among juvenile sexual offenders. *Sexual Abuse: Journal of Research and Treatment, 11,* 69-86.

Salo, R., Henik, A., & Robertson, L. C. (2001). Interpreting Stroop interference: An analysis of differences between task versions. *Neuropsychology, 15,* 462-471.

Schiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II. Perceptual learning, automatic attending and general theory. *Psychological Review, 84,* 127-190.

Schiller, P. H. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology, 72,* 105-108.

Segal, Z. V., Truchon, C., Horowitz, L. M., Gemar, M., & Guirguis, M. (1995). A priming methodology for studying self-representation in major depressive disorder. *Journal of Abnormal Psychology, 104,* 205-213.

Seto, M. C., Lalumière, M. L., & Blanchard, R. (2000). The discriminative validity of a phallometric test for pedophilic interests among adolescent sex offenders against children. *Psychological Assessment, 12,* 319-327.

Shin, L. M., Whalen, P .J., Pitman, R. K., Bush, G., Macklin, M. L., Lasko, N. B., …Rauch, S. L. (2001). An fMRI study of anterior cingulate function in posttraumatic stress disorder. *Biological Psychiatry, 50*, 932-942.

Smith, P. (2009). Assessing sexual interest with the emotional Stroop test. In D.Thornton & D. R. Laws (Eds.), Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders (pp. 159-176). Chichester, UK: Wiley.

Smith, P., & Waterman, M. (2005). Sex differences in processing aggression words using the Emotional Stroop task. *Aggressive Behavior, 31,* 271-282.

Smith, P., & Waterman, M. (2004). Procession bias for sexual material: The emotional Stroop and sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 16,* 163-171.

Smith, P., & Waterman, M. (2003). Processing bias for aggression words in forensic and nonforensic samples. *Cognition and Emotion, 17,* 681-701.

Smith, S. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*, 143-155.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 461-479.

Spinella, M., White, J., Frank, M. L., & Schiraldi, J. (2006). Evidence of orbitofrontal dysfunction in sex offenders. *International Journal of Forensic Psychology, 1,* 62-68.

Stone, M. H., Thompson, E. H., (2001). Executive function impairment in sexual offenders. *Journal of Individual Psychology, 57*, 51–9.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Taylor, S. F., Kornblum, S. Minoshima, S., Oliver, L. M. & Koeppe, R. A. (1994). Changes in medial cortical blood flow with a stimulus response compatibility task. *Neuropsychologia, 32,* 249-255.

Teasdale, J. D., Howard, R. J., Cox, S. G., Ha, Y., Brammer, M. J., Williams, S. C. R., & Checkley, S. A. (1999). Functional MRI study of cognitive generation of affect. *American Journal of Psychiatry, 156,* 209-215.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of the left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, 94,* 14792-14797.

Thornton, D., & Laws, D. R. (2009). Introduction. In D. Thornton & D. R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 177-201). Chichester, UK: Wiley.

Valliant, P. M., Gauthier, T., Pottier, D., Kosmyna, R. (2000). Moral reasoning, interpersonal skills, and cognition of rapists, child molesters, and incest offenders. *Psychological Reports, 86*, 67–75.

Van den Heuvel, O. A., Veltman, D. J., Groenewegen, H. J., Witter, M. P., Merkelbach, J., Cath, D. C.,…Dyck, R. (2005). Disorder-specific neuro-anatomical correlates of attentional bias in obsessive-compulsive disorder, panic disorder and hypochondriasis. *Archives of General Psychiatry, 62*, 922-933.

Van Mourik, R. Oosterlaan, J., & Sergeant, J. A. (2005). The Stroop revisited: A meta-analysis of interference control in AD/HD. *Journal of Child Psychology and Psychiatry, 46,* 150-165.

Van Veen, V., & Carter, C. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage, 27,* 497-504.

Vendrell, P., Junque, C., Pujol, J., Jurado, M. A., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia, 33,* 341-352.

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging, 13,* 120-126.

Ward, T., Hudson, S. M., Johnston, L., & Marshall, W. L. (1997). Cognitive distortions in sex offenders: An integrative review. *Clinical Psychology Review, 17,* 479-507.

Waters, A. J., & Feyerabend, C. (2000). Determinants and effects of attentional bias in smokers. *Psychology of Addictive Behaviors, 14,* 111-120.

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach and avoidance-related social information. *Journal of Personality and Social Psychology, 78,* 1024-1037.

Wertz, J. A., & Sayette, M. A. (2001). Effects of smoking opportunity on attentional bias in smokers. *Psychology of Addictive Behaviors, 15,* 268-271.

West, R., & Bell, M. A. (1997). Stroop color-word interference and electroencephalogram activation: Evidence for age-related decline of the anterior attentional system. *Neuropsychology, 11,* 421-427.

Whalen, P. J., Bush, G., McNally, R. J., Wilhelms, S., McInerney, S. C., Jenike, M. A. et al. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry, 44,* 1219-1228.

Wigenfeld, K., Rullkoefter, N., Mensebach, C., Beblo, T., Mertens, M., Kreisel, S., …Woermann, F. (2009). Neural correlates of the individual emotional Stroop in borderline personality disorder. *Psychoneuroendocrinology, 34,* 571-586.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3-24.

Wong, S., Olver, M. E., Nicholaichuk, T. P., & Gordon, A. (2004). *The Violence Risk Scale: Sexual Offender Version (VRS:SO)*. Regional Psychiatric Centre and University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Woolrich, M. W. (2008). Robust group analysis using outlier inference. *NeuroImage 41*, 286-301.

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multi-level linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage 21*, 1732-1747.

Woolrich, M. W., Ripley, B. D., Brady J. M., & Smith S. M. (2001). Temporal autocorrelation in univariate linear modelling of fMRI data. *NeuroImage 14*, 1370-1386.

Worling, J. R. (2002). Assessing risk of sexual assault recidivism with adolescent sex offenders. In M.C.Calder (Ed.), *Young people who sexually abuse: Building the evidence base for your practice* (pp. 353-365). Lyme Regis, U.K.: Russell House Publishing.

Worsley, K. J. (2001). Statistical analysis of activation images. In P. Jezzard, P.M. Matthews & S.M. Smith (Eds.), *Functional MRI: An Introduction to Methods (Chapter 14)*. Oxford: Oxford University Press.

Wurm, L. H., Labouvie-Vief, G., Aycock, J., Rebucal, K. H., & Koch, H. E. (2004). Performance in auditory and visual emotional Stroop tasks: A comparison of older and younger adults. *Psychology and Aging, 19,* 523-535.

Yates, P. M., Goguen, B. C., Nicholaichuk, T. P., Williams, S. M., & Long, C. A., (2000). *National Sex Offender Programs (Volume II: Moderate Intensity)*, Correctional Service of Canada, Ottawa, ON.