

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). 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 modeled on adult sex offender treatment programs. 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 behavior 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 behavior similar to that represented in the vignette (Becker et al., 1992). In addition, Seto, Lalumière and Blanchard (2000) have 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), a lack of empirical studies validating the measure for use with adolescents, and matters regarding informed consent (Grant, 2006).

Additional efforts to measure sexual interests of adolescent sexual abusers have been conducted by Abel and colleagues (2004) incorporating viewing time as measured by the Abel Assessment for Sexual InterestTM (AASI). The AASI combines self-reported ratings of sexual interest, viewing time 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 behaviors 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). Despite some evidence to support its use in measuring sexual interest, the validity of this measure has been questioned largely due to the fact that the raw data is managed and controlled by Abel Screen Inc., meaning that true independent research is not possible with this measure, and because there is a lack of published psychometric data on the AASI (see Sachsenmaier & Gress, 2009 for a review).

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 there was an existing accurate measure of sexual interest that could be used with this subgroup. One measure that has been used in the adult literature in attempts to tap into the psychological nature of deviant sexual interests with offenders is the emotional Stroop task (Price, 2011; Price & Hanson, 2007; Smith & Waterman, 2004). The emotional Stroop 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). Importantly, the emotional Stroop task is easy to administer, resilient to faking, non-invasive for participants, cost-effective and requires little training (Price et al., 2012; Smith, 2009), therefore many of the limitations listed above of existing measures are addressed with this measure. 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 (Dagleish, 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 a psychological construct such as deviant sexual interest.

The emotional Stroop task has been extensively tested in adult clinical samples, however, the research using the emotional Stroop task with adult and adolescent forensic samples is limited (Moradi et al., 1999; Price & Hanson, 2007). 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 (Smith & Waterman, 2004), and more specifically that could be able to differentiate between subgroups of sexual abusers (Price, 2011; 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. To date, the studies that have used this task to assess deviant sexual interests in the adult offender literature have found information-processing biases to sexual word stimuli for sexual abusers compared to non-offenders (see Price, Beech, Mitchell & Humphreys, 2012 for a review).

Specific to adolescents, in the general literature on the traditional color-word Stroop task (Stroop, 1935), 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 color-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, Carroll et al. (2006) found that their offender groups did not display deficits in inhibitory control as measured by the color-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. Carroll 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 behaviors and impulsivity to poor performance on the color-word Stroop task. It is important to note that Carroll 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 computerized versions (Salo, Henik, & Robertson, 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 participants 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 color-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 participants. Adolescents that exhibited fire setting behaviors 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 behavior. 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 categorization of antisocial behavior between the two offender groups.

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 aims of this study were: (1) to examine the response patterns of adolescent sexual abusers, adolescent offending controls and adolescent non-offending controls to sexual interest word stimuli; (2) to examine whether participant characteristics such as age, executive function and beliefs about children influence Stroop results with this sample; (3) 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; and (4) to contribute to research regarding the measurement of sexual interests of adolescent sexual abusers.

Method

Participants

Three groups of adolescent male participants took part in this study: (1) sexual abusers (n=24); (2) offending controls (n=21); and (3) non-offending controls (n=21). Twenty-two of the sexual abusers were recruited from an independent organization that works with young people who display inappropriate sexual behaviors. 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) organization in the West Midlands, UK. The non-offending adolescent controls were recruited from a school in the West Midlands and through recruitment of individuals attending Open Days at a West Midlands University.

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.

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 in this study completed two emotional Stroop tasks, the British Picture Vocabulary Scale-II (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).

BPVS-II. Emotional Stroop interference effects are typically not experienced by those that lack basic reading skills because they are not distracted by word meaning (Schiller, 1966), therefore the BPVS-II 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. The BPVS is intended for use with participants between the ages of 3:00 and 15:08 as a measure of verbal ability or verbal intelligence. Most of the participants in the current study 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.

Hayling and Brixton tests of executive function. The Hayling Sentence Completion test is a measure of response initiation, response suppression 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, with the outcome measure being total number of errors across 55 trials (Burgess & Shallice, 1997). These tasks share similar cognitive processing demands to the Stroop task (Perret, 1974 cited in Burgess & Shallice, 1996; De Zubicaray, Zelaya, Andrew, Williams & Bullmore, 2000), therefore it was of interest to examine

the relationship between these tests of executive function and the emotional Stroop in this study. In addition, these tests were used to confirm that participants did not suffer from impaired executive functioning that would disrupt their ability to complete the Stroop tasks successfully.

Beliefs About Children Scale. A shortened version of the Beliefs about Children Scale (BACS; Beckett, 1987) was used to test whether the participants' beliefs about children influenced Stroop results. The BACS yields 2 subscale scores: Cognitive Distortions (CD) and Emotional Congruence with children (EC), scored on a 4-point scale. The CD scale consists of 15 items that assess beliefs about children and their sexuality, such as "Children know a lot about sex" and "Children know more about sex than adults do". Beech (1998) found the test-retest reliability to be .77 and the scale has been reported by Thornton (as cited in Beech, 1998) to have high internal reliability ($\alpha = .90$).

The EC scale is also scored through 15 items that are intended to measure the understanding of what the individual believes to be the thoughts, feelings, and interests of children. Beech (1998) reported the test-retest reliability as .63. Examples of questions that measure level of emotional congruence with children include "I have loved a child at first sight" and "I know when children are interested in me". Higher scores would indicate higher levels of self-reported emotional congruence with children (Harkins, Flak, Beech & Woodhams, 2012).

The emotional Stroop task. 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 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). Button-press response recordings were used due to the noisy

environments the research was often carried out in (i.e., prisons, group treatment settings), and in order to maintain consistency across the groups.

Two emotional Stroop word stimulus sets were used: (1) Smith and Waterman (2004) word stimulus set; and (2) Price (2011) word stimulus set (see Appendix). These word sets were chosen because they have been used in the adult literature and we were interested in examining whether adolescents demonstrated similar responses. Price (2011) empirically derived the new word stimuli due to suggestions made in the adult literature that existing word lists were generally sexual in nature and may not represent a true reflection of the motives, thoughts or feelings experienced by sexual offenders (Price & Hanson, 2007). These new word stimuli (see Price (2011) for full description) were derived from the responses of workers and sexual abusers 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? 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).

Analysis

Eight emotional Stroop bias scores (from the two word stimulus sets) were compared across the groups and 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): positive, negative, color, aggression and sexual, emotional/ personality descriptors (EPD), sexual actions (SA) and physical descriptors (PD). For the Price (2011) word set, an additional emotional Stroop bias score for all of the experimental words (EXP) combined were also explored.

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 RTs of two standard deviations below the mean RTs for all of word categories indicating an unsuccessful attempt at completing the emotional Stroop task. Data from one non-offending control subject on the Stroop task using the Smith and Waterman stimuli was lost, resulting in a sample size of 20 for this group.

All available data was explored to test whether the assumptions of parametric testing were met. Attempts made to transform the data that did not meet the assumptions of parametric testing did not correct the data enough to be normally distributed despite the assumption of homogeneity of variance being maintained with the transformed data. Therefore, when necessary, the non-parametric equivalents of tests were used to explore whether significant differences were present between the groups (see Price 2011 for full description of the analysis).

All mean RTs and emotional Stroop bias scores were entered into separate one-way ANOVAs with participant group as the between-groups variable (Kruskal-Wallis test as non-parametric equivalent). Post-hoc analyses were conducted using a Bonferroni correction to control for the Type I error rate (Mann-Whitney tests with a Bonferroni correction, critical value of .025 for non-parametric equivalents). Effect sizes for the one-way ANOVAs were calculated using omega (ω) and r (non-parametric):

Two-way mixed ANOVAs with type of emotional Stroop bias score as the repeated measures variable and participant group as the between groups variable were run. 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.

Finally, analysis of covariance (ANCOVA) was carried out on the dependent variables (i.e., mean RTs 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 and on the BACS subscales of cognitive distortions and emotional congruence with children. 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 ANCOVAs were interpreted with caution because there is no available non-parametric equivalent of ANCOVA (see Price, 2011 for more detailed description of data analysis).

Procedures

Participants completed four practice blocks (25 words per block) of neutral word stimuli in order to learn the color-mappings of the response buttons prior to the presentation of target words. 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 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 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.

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 1). 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, indicating lower levels of vocabulary understanding than the non-offenders. No significant differences were evident on these variables between the offending controls and the other groups.

Insert Table 1 approximately here.

Smith and Waterman Stimulus Set

Results from the one-way ANOVAs to test for significant differences in the mean RTs for the Smith and Waterman (2004) stimulus set are displayed in Table 2. Significant differences were found between the groups on the mean RTs 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 color-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 color-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 color-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 color RTs, $H(2) = 12.37$, $p < .05$, where adolescent sexual abusers took significantly longer to color-name than the non-offending controls ($U = 93.00$, $r = .52$).

Insert Table 2 approximately here

Figure 1 displays the patterns in emotional Stroop bias effects experienced by the different groups. Results from the one-way ANOVAs to test for significant differences in emotional Stroop bias scores 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.

Insert Figure 1 approximately here.

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. 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 revealed that the main effect of type of Stroop bias was due to the fact that, overall, the color 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). 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.

Insert Figure 2 approximately here.

Finally, the results from the analysis of covariance from this word stimulus set did not yield any significant relationships between the mean RTs and Stroop bias effects and the covariates: age, BACS (CD), BACS (EC), Hayling and Brixton tests of executive function and BPVS raw scores, indicating that these variables did not in fact act as confounding variables.

Price (2011) Word Stimulus Set

Mean RTs and emotional Stroop bias scores are displayed in Table 3 for the Price (2011) word stimulus set. One-way ANOVAs yielded significant differences for mean RTs between the adolescent sexual abusers and the non-offending adolescent controls for the matched sexual action (MSA), $F(2, 62) = 5.65, p < .05$ words.

Insert Table 3 approximately here.

Nonparametric testing yielded a significant difference for mean emotional personality descriptor (EPD) RTs, $H(2) = 11.92$, 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 matched emotional/personality descriptors (MEPD; $U = 127.00, r = .42$); sexual actions

(SA; $U = 135.00$, $r = .40$); physical descriptors (PD; $U = 134.00$, $r = .40$); and matched physical descriptor (MPD; $U = 153.00$, $r = .34$) mean RTs.

Overall, significant differences were found between the groups for the experimental RTs, $H(2) = 10.75$, $p < .05$ and the matched RTs, $H(2) = 9.89$, $p < .05$. 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 RTs ($U = 114.00$, $r = .47$), and mean matched RTs ($U = 119.00$, $r = .45$).

No significant differences were yielded for the emotional Stroop bias scores (i.e., Stroop effects). 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.

Insert Figure 3 approximately here.

The two-way mixed ANOVA yielded 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.

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 RTs or Stroop bias effects from the Price (2011) 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 unsystematic. 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 RTs 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 color-name than the two control groups. This finding is consistent with the findings from the adult literature (Price, 2011; Price & Hanson, 2007; Smith & Waterman, 2004). The reasons why the sexual abusers (and offender groups more generally) consistently respond more slowly to the word stimuli in this task are unknown. 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 RTs 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 behavioral 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 color stimuli, indicating the ability to experience Stroop interference effects that are consistent with what we would expect of color-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 Price (2011) word set were not significant here. However, it is interesting to note that none of the adolescent groups experienced Stroop interference for the emotional/personality descriptors when these word stimuli displayed the greatest effect in Price (2011) between the adult sexual abusers and the non-offending controls. The finding from this

study with adolescent groups 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 that have been tested using these word stimuli have also displayed Stroop interference for the sexual action and physical descriptor word categories (Price, 2011). 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 Price (2011) word stimulus set was very small across all of the adolescent groups. This would suggest that, overall, the word stimuli are not uniform enough when analyzed together to be able to discriminate between groups of individuals.

Limitations and Directions for Future Research

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 colors 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, 2006). A solution to this problem may be to employ a variant of the Stroop task that uses picture stimuli rather than word stimuli. For example, the Pictorial Stroop task has been used recently to assess sexual interest between heterosexual and homosexual adult males (Bourke & Gormley, 2012) and further comparisons have been made between these groups and sexual offenders (Ó Ciardha & Gormley, 2012). Though the task has elicited mixed results in samples of adult males to differentiate between age categories of sexual interest in these studies, it may be a more appropriate methodology to use with adolescent groups 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.

Conclusions

The overall aims of this study were to determine whether the emotional Stroop task is a reliable tool to be used in the assessment of deviant sexual interest for adolescent groups; and to explore its utility using the same word sets that have been used in the adult literature. 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 patterns 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, 2006; Grant, 2006). 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).

References

- 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.
doi: 10.1177/107906320401600306
- 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. doi: 10.1111/j.1556-4029.2005.00010.x
- Barbaree, H. E. (1990). Stimulus control of sexual arousal: Its role in sexual assault. In W. 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 York: The Guilford Press.
- 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. C. (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. C. (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., 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.
- Beech, A. R. (1998). A psychometric typology of child abusers. *International Journal of Offender Therapy and Comparative Criminology, 42*, 319-339.
doi: 10.1177/0306624X9804200405
- 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.
- 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.
doi:10.1017/S003329170500499X

- Bourke, A. B., & Gormley, M. J. (2012). Comparing a pictorial Stroop task to viewing time measures of sexual interest. *Sexual Abuse: A Journal of Research and Treatment, 24*, 479-500. doi: 10.1177/1079063212438922
- Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton Tests*. London: Harcourt Assessment.
- Burgess, P.W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34*, 263-273. doi: 10.1016/0028-3932(95)00104-2
- 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. doi: 10.1007/s10964-006-9053-6
- 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.
- De Zubicaray, G.I., Zelaya, F.O., Andrew, C., Williams, S.C.R., & Bullmore, E.T. (2000). Cerebral regions associated with verbal response initiation, suppression and strategy use. *Neuropsychologia, 38*, 1292-1304. doi: 10.1016/S0028-3932(00)00026-9
- 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.

- Erooga, M., & Masson, H. (2006). Children and young people with sexually harmful or abusive behaviors: 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.
- 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. doi: 10.1007/s10802-009-9339-8
- Gerhold, C. K., Browne, K. D., & Beckett, R. (2007). Predicting recidivism in adolescent sexual offenders. *Aggression and Violent Behavior*, *12*, 427-438. doi:10.1016/j.avb.2006.10.004
- 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. doi: 10.1111/j.1469-7610.2004.00276.x
- Golden, C. J. (1978). *Stroop Colour and Word Test: A manual for clinical and experimental uses*. Stoelting Company, Illinois.
- Grant, H. (2006). Assessment issues in relation to young people who have sexually abusive behavior. In M. Erooga & H. Masson (Eds.), *Children and Young People Who Sexually Abuse Others. Challenges and Responses (2nd ed)* (pp. 67-76). London: Routledge.

- 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.
- Harkins, L., Flak, V. E., Beech, A. R., & Woodhams, J. (2012). Evaluation of a community-based sex offender treatment program using a Good Lives Model approach. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 1-25. doi: 10.1177/1079063211429469
- 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. doi:10.1207/s15326942dn3002_2
- 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. doi: 10.1023/A:1020663723593
- Klein, G. S. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology*, *77*, 576-588.
- 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. doi: 10.1177/107906328900200105
- Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion*, *6*, 62-72.
- 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.
- 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 color-naming task. *Psychological Medicine*, *29*, 415-419.
- 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. doi: 10.3758/BF03193068
- Ó Ciardha, C., & Gormley, M. (2012). Using a pictorial-modified Stroop task to explore the sexual interests of sexual offenders against children. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 175-197. doi:10.1177/1079063211407079
- Oliver, B. (2007). Three steps to reducing child molestation by adolescents. *Child Abuse and Neglect*, *31*, 683-689. doi:10.1016/j.chiabu.2006.09.008
- Perret, E. (1974). The left frontal lobe in man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia*, *12*, 323-330.
- 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.
- Price, S.A. (2011). Cognitive assessment of deviant sexual interest using the emotional Stroop and fMRI (PhD thesis, The University of Birmingham, Birmingham, United Kingdom). Retrieved from <http://etheses.bham.ac.uk/1541>.

- Price, S.A., Beech, A.R., Mitchell, I.J., & Humphreys, G.W. (2012). The promises and perils of the emotional Stroop task: A general review and considerations for use with forensic samples. *The Journal of Sexual Aggression*, 18, 253-268. doi: 10.1080/13552600.2010.545149
- Price, S.A., & Hanson, R.K. (2007). A modified Stroop task with sexual offenders: Replication of a study. *Journal of Sexual Aggression*, 13 (3), 203-216.
- Sachsenmaier, S. J., & Gress, C. L.Z. (2009). The Abel Assessment for Sexual Interests – 2: A critical review. In D. Thornton & D. R. Laws (Eds.), *Cognitive Approaches to the Assessment of Sexual Interest in Sexual Offenders* (pp. 177-201). Chichester, UK: Wiley.
- Salo, R., Henik, A., & Robertson, L. C. (2001). Interpreting Stroop interference: An analysis of differences between task versions. *Neuropsychology*, 15, 462-471.
- Schiller, P. H. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology*, 72, 105-108.
- 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. doi: 10.1037/1040-3590.12.3.319
- 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. (2004). Procession bias for sexual material: The emotional Stroop and sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 16, 163-171. doi: 10.1023/B:SEBU.0000023064.54397.d6

- 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.
- 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 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. doi: 10.1111/j.1469-7610.2004.00345.x
- 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.

Table 1

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

Variable	Non-Offending Controls	Offending Controls	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)
(Price, 2011)	4.71 (3.62)	7.35 (6.85)	6.04 (3.76)

Table 2

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

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

Table 3

Mean RT and Emotional Stroop Bias Scores in Milliseconds (SD) for the Price (2011) Word

Stimulus Set

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

Figure 1. Mean Emotional Stroop Bias Scores: Smith and Waterman Stimulus (2004) Set

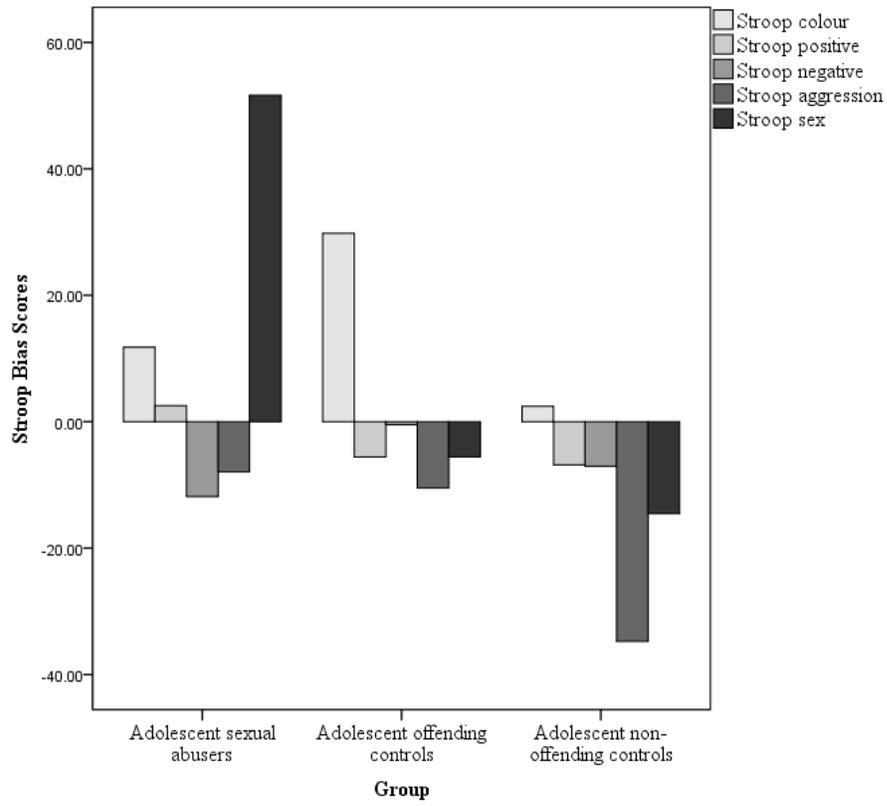


Figure 2. Means for the Main Effects of Type of Stroop Bias and Participant Group: Smith and Waterman (2004) Stimuli

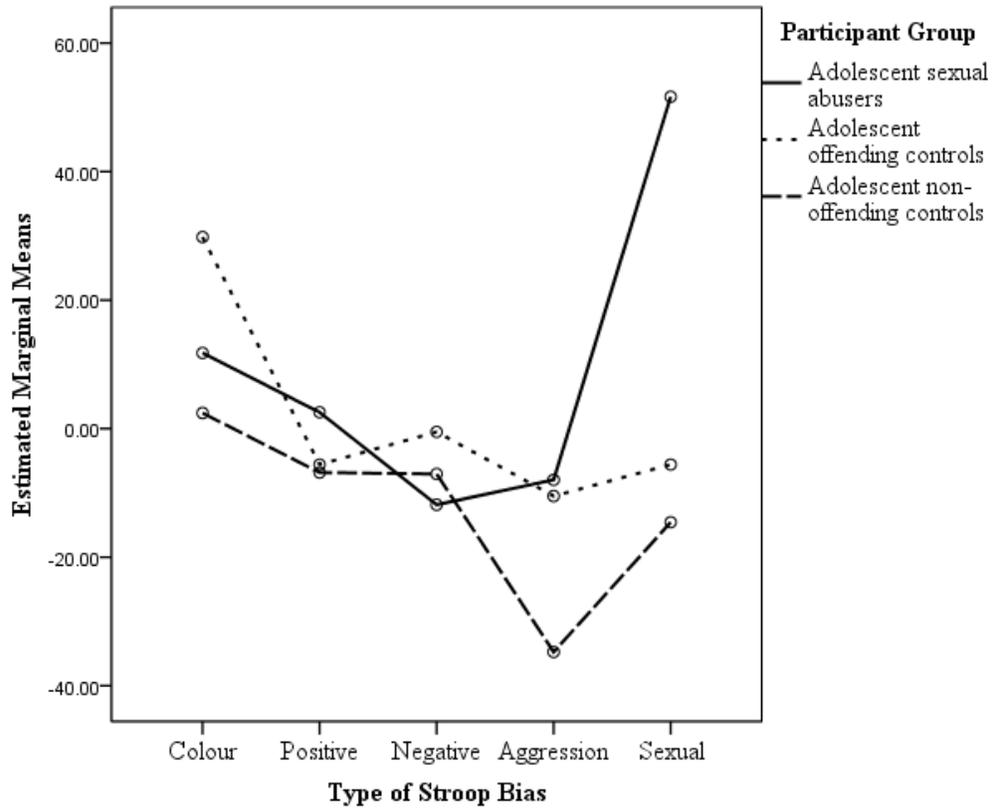
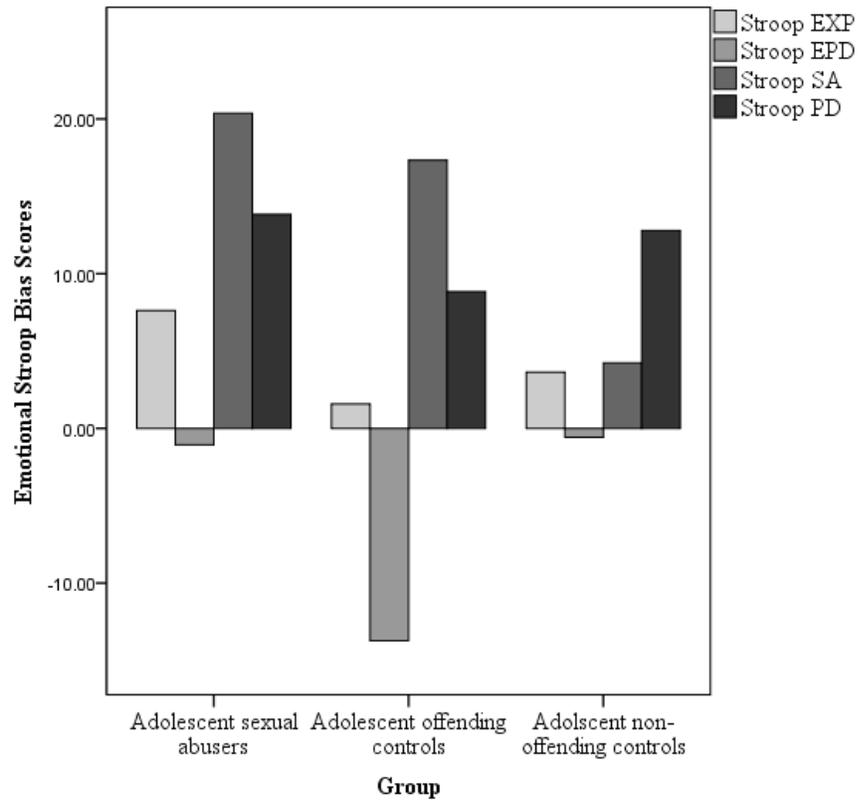


Figure 3. *Emotional Stroop Bias Scores: Price (2011) Word Stimuli*



Appendix

Word Stimuli (Price, 2011)

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