

An Indirect Measure of Sexual Interest: Using the Implicit Association Test with  
Child Molesters

A thesis submitted to  
the Faculty of Graduate Studies and Research  
in Partial Fulfillment of the requirements for the degree  
Masters of Arts  
Psychology

by

Kelly M. Babchishin

Department of Psychology

Carleton University

September, 2009

©2009 Kelly M. Babchishin



Library and Archives  
Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file Votre référence*  
ISBN: 978-0-494-60297-3  
*Our file Notre référence*  
ISBN: 978-0-494-60297-3

**NOTICE:**

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

**AVIS:**

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

  
**Canada**

### **Abstract**

Sexual interest in children has been identified as an important cause of child sexual abuse. Preliminary studies suggest that the Implicit Association Test (IAT) is a promising tool to measure sexual interest in children. In the current study, 34 child molesters were compared to 21 non-sex offenders on the Sexual Attraction to Children IAT (SAC-IAT), which was developed for the current study, as well as other measures of sexual interest. The SAC-IAT was not able to distinguish between groups; however, it was related to other measures of sexual interest that could distinguish between groups. Cumulative meta-analysis found a moderate effect size for the ability of IAT measures of sexual interest in children to distinguish between child molesters and non-molesters. Findings from the present study are an important step towards evaluating the validity of an IAT measure of sexual interest in children.

### **Acknowledgements**

First and foremost I would like to express my sincere appreciation to Kevin Nunes for his expertise, guidance, and encouragement throughout this project. I appreciate your insightful feedback and the perspectives you provided. I would also like to thank Adelle Forth and Ralph Serin for their guidance and constructive feedback.

I am grateful to the friendly and cheerful group of fellow students who have helped shape this experience as nothing less than memorable. Special thanks to Nicolas Kessous for helping me collect the data and for his encouragement throughout this project. Thank you to Leslie Helmus and Chantal Hermann for their encouragement, support, and kind words. You have been fantastic colleagues, and truly great friends.

I would also like to thank the Association for the Treatment of Sexual Abusers for helping to fund this project. Many thanks to Carolyn Bourgeois, Anita Cumbleton, Yolanda Fernandez, Carmen Gress, Cindy Hudson, Daryl Kroner, Richard Laws, Petrina Lemieux, Jan Looman, Wadgy Loza, Greg Maillet, Wendy McMahon, Jeremy Mills, Ed Peacock, Geris Serran, Bill Walker, Jennifer Walsh, as well as the support and security staff of the institutions. Your support was critical to the success of this project.

I am also grateful to my family and friends for their love and encouragement, and I especially would like to thank Lyzon Babchishin and Nicholas Boucher. Nick you have enabled me to pursue my interest and I truly appreciate your unwavering support, encouragement, and love. Lyzon, thank you for your encouragement and insight, as well as coaxing me to take time off when I missed too many of our weekend get-togethers.

## Table of Contents

Abstract.....	ii
Acknowledgements.....	iii
Theories of Child Sexual Abuse .....	1
Finkelhor’s (1984) Four-Factor Model.....	1
Hall and Hirschman’s (1992) quadripartite model of sexual aggression against children. ....	2
Ward and Keenan’s (1999) Five Implicit Theories .....	2
Ward and Siegert’s (2002) Pathway Model.....	3
Sexual Interest in Children.....	5
Measurement of Sexual Interest in Children .....	6
Physiological Measures. ....	6
Penile Plethysmography.....	6
Polygraph Testing. ....	8
Self-report Measures.....	8
Viewing Time Measures.....	10
The Implicit Association Test (IAT).....	13
The Construct Validity of IAT Measures. ....	16
The IAT Adapted to Assess Sexual Attraction to Children.....	18
Meta-Analysis.....	26
Present Study .....	27
Hypotheses.....	28
Primary Hypotheses.....	28

Secondary Hypotheses.....	29
Method.....	29
Participants.....	29
Measures.....	32
Sexual Attraction to Children IAT (SAC-IAT).....	32
Sexual Interest Profiling System (SIPS).....	33
Viewing Time.....	34
Viewing Time Attractiveness Ratings.....	35
Penile Plethysmography (PPG).....	36
Rapid Risk Assessment for Sex Offence Recidivism (RRASOR).....	36
Static-99.....	37
Stable-2000.....	37
Stable-2007.....	38
Statistical Information on Recidivism (SIR).....	39
The Screening Scale for Pedophilic Interests (SSPI).....	40
Balanced Inventory of Desirable Responding (BIDR).....	40
Clarke Vocabulary Scale (CVS).....	41
Background Questionnaire (BQ).....	41
Materials and Apparatus.....	41
Procedure.....	42
Interviews.....	42
Stimuli Ratings.....	43
Pilot Testing.....	43

Study Proper.....	44
Data Treatment.....	45
IAT Scoring Algorithm.....	45
Viewing Time. ....	46
SIPS.....	47
PPG. ....	47
Results.....	47
Data Screening.....	47
Internal Consistency.....	48
Primary Hypotheses.....	50
Descriptive Statistics.....	50
Group Differences on the SAC-IAT.....	54
Cumulative Meta-Analysis.....	54
Group differences in Viewing Time. ....	55
Group differences on Viewing Time Attractiveness Rating.....	58
Group differences on the SIPS.....	61
Alternative Explanations.....	63
Secondary Hypotheses.....	68
Convergent Validity.....	68
Measures of Sexual Interest in Children and Risk of Recidivism. ....	71
Discussion.....	73
References.....	84
Appendices.....	106

Appendix A: Issues Affecting Internal Validity and Reliability of IAT Measures .....	106
Appendix B: SAC-IAT Stimuli .....	110
Appendix C: Balanced Inventory of Desirable Responding.....	111
Appendix D: Clarke Vocabulary Scale.....	113
Appendix E: Background Questionnaire .....	115
Appendix F: Consent Form for the Preliminary Interview .....	116
Appendix G: Preliminary Interview Questions.....	117
Appendix H: Consent Form for Rating Questionnaire .....	118
Appendix I: Rating Questionnaire with Responses Included .....	119
Appendix J: Consent Form for the Pilot Study and Study Proper .....	130
Appendix K: Order of Presentation of the Measures.....	130
Appendix L: Debriefing Script .....	132
Appendix M: A Brief Description of the D Algorithm.....	133
Appendix N: Internal Consistency Analysis for the IAT.....	134

**List of Tables**

Table 1: Factors Theorized to Enable Child Sexual Abuse .....	4
Table 2: Statistical Power of Six Studies Examining IAT Measures of Sexual Interest in Children.....	26
Table 3: Description of Participants .....	31
Table 4: Blocks Presented in the SAC-IAT.....	33
Table 5: Internal Consistency of Self-Report Measures .....	49
Table 6: Means, Standard Deviations, Cohen's <i>ds</i> and AUCs for the Measures of Sexual Interest in Children .....	53
Table 7: Means, Standard Deviation, and Cohen's <i>ds</i> for SIPS Subscales and Composite Scores.....	62
Table 8: Correlations between Possible Confounding Variables and the Measures of Sexual Interest in Children.....	63
Table 9: Relationship between Measures of Sexual Interest and Variables Specific to Child Molesters.....	64
Table 10: Relationship between SAC-IAT Male and SAC-IAT Female and Other Measures of Sexual Interest .....	67
Table 11: Convergent Validity of the Measures of Sexual Interest in Children.....	69
Table 12: Relationship between the SAC-IAT and Measures of Sexual Interest Computed by Averaging the Genders.....	70
Table 13: Relationship between Measures of Sexual Interest in Children and Risk of Sexual Recidivism .....	72

**List of Figures**

Figure 1. An Example of an IAT Adapted to Assess Sexual Interest in Children.....	15
Figure 2. Receiver Operator Curves for Measures of Sexual Interests.....	52
Figure 3. Average Viewing Time for each Tanner Scale .....	56
Figure 4. Viewing Time for Female Pictures.....	57
Figure 5. Viewing Time for Male Pictures .....	58
Figure 6. Average Sexual Attractiveness Ratings.....	59
Figure 7. Sexual Attractiveness Ratings for Female Pictures.....	60
Figure 8. Sexual Attractiveness Ratings for Male Pictures .....	60

## An Indirect Measure of Sexual Interest: Using the Implicit Association Test with Child Molesters

Sexual abuse of children is considered a significant societal concern and has far-reaching ramifications on the healthy development of its victims. Individuals who have been victimized have increased likelihood of depression, suicide, sexual promiscuity, and academic difficulties (Beitchman et al., 1992; Browne & Finkelhor, 1986; Paolucci, Genuis, & Violato, 2001). Although sexual abuse is difficult to quantify (Goldman & Padayachi, 2000), meta-analytic findings suggest that approximately 22.3% of women and 8.5% of men in North America report being sexually abused in childhood (Gorey & Leslie, 1997; child defined as a person under the age of 18). In Canada, approximately 9,300 cases of child sexual abuse (274 per 100,000) were reported to police in 2003 (Canadian Centre for Justice Statistic, 2005; child defined as a person under the age of 18). The divergence between self-reported and official rates of child sexual abuse can be attributed to the underreporting of sexual assaults and the use of plea bargains (e.g., sexual assault charges lowered to lesser convictions, such as assault), which result in official rates underestimating child sexual abuse (Lisak & Miller, 2002; Terry, 2006). Considering the harm caused by the perpetrator, accurate assessment and management of sexual offenders are fundamental in reducing child sexual abuse. Greater insight into the underlying constructs of child sexual abuse is necessary to successfully reach these goals.

### **Theories of Child Sexual Abuse**

*Finkelhor (1984)* integrated several explanatory models found in the literature into a *Four-Factor Model* to explain child sexual abuse (CSA). The four factors

identified are: emotional congruence, sexual arousal to children, blockage, and inhibition. Finkelhor theorized that identifying with children and finding sexual relationships with them emotionally fulfilling, viewing of children as sexual objects and engaging in or fantasizing about sexual relationships with them, an inability to find emotional fulfillment and sexual relief within normal, adult relationships, and the absence of disinhibition (e.g., holding beliefs supportive of CSA or not being dissuaded by societal norms) are central factors enabling the onset of CSA. The four factor model further explains that these factors have an additive effect: the more factors present, the more likely CSA could occur. All four factors, however, are not required to be present for child sexual abuse to occur.

*Hall and Hirschman's (1992) Quadripartite Model of Sexual Aggression Against Children* identified (1) sexual arousal to children, (2) cognitions that justify sexual aggression against children, (3) difficulty moderating negative affective state (e.g., depression, anger), and (4) personality problems (e.g., poor social skills, impulsivity) as factors linked to CSA. Similar to Finkelhor (1984), Hall and Hirschman theorized that not all factors are required to be present for CSA to occur.

*Ward and Keenan (1999)* identified *Five Implicit Theories* of CSA. They theorized that viewing children as sexual objects, viewing CSA as unlikely to harm children (i.e., nature of harm), a sense of entitlement, uncontrollability (the belief that the cause of their behaviours is due to external factors), and viewing the world as dangerous, are factors that enable the initiation and maintenance of CSA. These factors are purported to interplay to facilitate CSA.

*Ward and Siegert (2002) Pathway Model* report similar factors from the previous theories. The pathway model postulates that social skills deficits, distorted sexual scripts (e.g., believe children are appropriate sexual partners), emotional dysregulation, and cognitive distortions are central factors facilitating CSA. This model purports that a child molester would only hold one of these factors as the primary factor, with the other factors taking a secondary role and influencing the primary one. As such, Ward and Siegert identified five possible pathways to CSA, one for each four factors, and the fifth pathway represents an individual in which all four factors are equally prevalent.

It is clear that several factors are theorized to be involved in the onset of CSA. For example, all of these theories recognize that sexual arousal in children increases the likelihood that someone will commit CSA. Cognitive distortions as well as a lack of normal adult relationships are also postulated to facilitate CSA. The strong theoretical basis for these factors has led to the development of specialized assessment instruments and research examining their relationship with CSA. Table 1 presents meta-analytic results of Hanson & Morton-Bourgon (2004) and Whitaker et al. (2008).

Table 1  
Factors Theorized to Enable Child Sexual Abuse

Factor	Theoretical Basis	Initiation <sup>a</sup> (Group difference)			Maintenance <sup>b</sup> (Sexual Recidivism)				
		<i>d</i>	95% C.I.	<i>N</i>	<i>k</i>	<i>d</i>	95% C.I.	<i>N</i>	<i>k</i>
Deviant sexual interest <sup>c</sup>	Finkelhor (1984), Hall and Hirschman's (1992) Ward and Keenan (1999)	.30	.12 – .47	591	3	.32	.16 – .47	1,278	10
Cognitions supporting child sexual abuse	Finkelhor (1984), Hall and Hirschman's (1992), Ward and Keenan (1999), Ward and Siebert's (2002)	.49 <sup>d</sup>	-	-	-	.22	.05 – .38	1,617	9
Conflicts with intimate relationship	Finkelhor (1984), Ward and Siebert's (2002)	.42	.15–.70	711	10	.36	.05–.66	298	4

*Note.* <sup>a</sup> Findings presented in Whitaker et al. (2008) based on comparisons between child molesters and non-sex offenders. <sup>b</sup> Findings presented in Hanson & Morton-Bourgon (2004), average follow-up 5 to 6 years. <sup>c</sup> Deviant sexual interest was assessed using Penile Plethysmography (PPG). This is not meant to be an exhaustive list of the theories of child sexual abuse. <sup>d</sup> Confidence interval, number of participants, and number studies was not provided, but differences were reported as significant at the  $p < .5$  level. According to Cohen (1988), *d* values of .20 are considered "small", .50 "medium", and .80 "large". C.I.: confidence interval. *k*: number of studies.

## Sexual Interest in Children

There are multiple factors involved in the initiation of child sexual abuse in the aforementioned theories. Sexual interest in children is just one factor that has both strong theoretical and research basis. As can be seen in Table 1, deviant sexual interest can differentiate non-sex offenders from child molesters, suggesting that this factor plays a part in the initiation of CSA (Whitaker et al., 2008). Sexual interest in children was also predictive of sexual recidivism in child molesters ( $d = .32, p < .05$ ; Hanson & Morton-Bourgon, 2004), suggesting that it plays a role in the maintenance of CSA. Perhaps due to both practical and theoretical importance, deviant sexual preference is included in risk assessment instruments (Stable-2007, Hanson, Harris, Scott, & Helmus, 2007; SORAG, Quinsey, Harris, Rice, & Cormier, 1998) and is an important target in many sexual offender treatment programs (e.g., Correctional Service Canada; Yates, Goguen, Nicholaichuk, Williams, & Long, 2000). Moreover, the criminal justice system considers the presence or absence of sexual deviancy in decision making, such as informing parole decisions (Simon & Schouten, 1991).

Due to the importance of sexual interest in children in the initiation of child sexual abuse, psychometrically sound measurement of this construct is essential for the effective management of child molesters. Namely, more accurate measurement of sexual interest in children will be associated with more accurate decisions and more effective treatment programs.

Better measurement of sexual interest in children can also help improve our understanding of its role in child sexual abuse. Indeed, more reliable and valid measurement of this construct can allow for psychometrically sound findings, thereby

advancing our understanding of the role sexual interest plays in the initiation of child sexual abuse. Current measures by which our understanding of sexual interest in children relies on have certain limitations, as well as strengths. A multi-modal measurement approach is therefore required to improve our knowledge on how sexual interest in children influences the initiation of child sexual abuse.

### **Measurement of Sexual Interest in Children**

#### **Physiological measures.**

*Penile Plethysmography.* Currently, penile plethysmography (PPG) is one of the most commonly used methods to assess sexual interest in children in Canada (Launay, 1999; Seto, 2004). PPG involves the physiological measurement of penile tumescence (erection) in response to various stimuli, such as visual or audio descriptions of deviant and non-deviant sexual activity. Sexual attraction is usually inferred from the amount of arousal in response to deviant stimuli (e.g., sexual abuse of a child) relative to non-deviant stimuli, such as consensual sexual activity between adults (Lalumière & Harris, 1998; Marshall & Fernandez, 2000; Pithers & Laws, 1995).

Empirical research using PPG has found considerable variability in the extent to which sexual interest in children is present among child molesters. Intrafamilial child molesters (i.e., those with related child victims), for example, demonstrate similar patterns of arousal to non-sex offenders (Freund, Watson, & Dickey, 1991; Grossman, Cavanaugh, & Haywood, 1992). Conversely, extrafamilial child molesters (i.e., with unrelated child victims) demonstrate greater PPG-assessed sexual interest in children than intrafamilial offenders and non-sex offenders (Quinsey, Chaplin, & Carigan, 1979; Seto, Lalumière, & Kuban, 1999).

Heterogeneity in the sexual interest of intrafamilial child molesters has also been reported. For instance, men who offend against girls who are not their daughters (i.e., are biologically related but not their parent, such as niece or grand-daughter) show more physiological arousal to children than do men who offend against biological daughters or stepdaughters (Blanchard, Kuban, Blak, Cantor, Klassen & Dickey, 2006; Langevin & Watson, 1991; Rice & Harris, 2002; Seto et al., 1999).

Extrafamilial child molesters with greater numbers of victims (defined as three or more victims) are more likely to display greater arousal to children than those with fewer victims (Blanchard, Klassen, Dickey, Kuban, & Blak, 2001). Findings also suggest that mixed offenders (i.e., victims include both related and unrelated children) may possess higher sexual arousal to children compared to uniquely intrafamilial or extrafamilial child molesters (Seto et al., 1999). Similarly, biological fathers and stepfathers who also molest children outside the family tend to demonstrate greater deviant sexual interest than those with uniquely related victims (Rice & Harris, 2002).

PPG, however, has been described as intrusive, technically demanding, expensive, and susceptible to faking if no preventative procedures (e.g., asking participants to press a button when certain types of stimuli are presented) are put in place (Harris, Rice, Cormier, & Quinsey, 1998; Kalmus & Beech, 2005). Further, low response rates are a frequent problem and there is a lack of standardization of the stimuli (Marshall & Fernandez, 2000). PPG has also been found to have low test-retest reliability (e.g.,  $r = .56$ ) in a group of child molesters, rapists, and non-offenders (e.g., Wormith, 1986). Despite reservations on its validity and reliability, PPG is one of the best available measures for distinguishing child molesters from non-sex offenders (Barsetti et al., 1998;

Freund & Watson, 1991; Marshall & Eccles, 1991) and from rapists (e.g., Looman & Marshall, 2001). As mentioned previously, PPG-assessed sexual interest in children has also been found to be one of the strongest predictors of sexual recidivism ( $d = 0.32$ ; Hanson & Morton-Bourgon, 2004). PPG can also provide the advantage of not having to rely solely on the offender's word or clinical judgments (Priest & Smith, 1992).

***Polygraph testing.*** Polygraph testing is another physiological measure, routinely used in community treatment programs in the United States (70% of those surveyed in 2002), although usually in conjunction with PPG or other measures of sexual interest (McGrath, Cumming, & Burchard, 2003). Research on sexual interest in children, however, infrequently uses polygraph due to several issues regarding its reliability and validity when used on sex offenders. For example, the accuracy of polygraph lowers when asking general questions (e.g., have you ever touched a child sexually?) compared to specific questions (e.g., have you touched Lucy sexually today?), resulting in greater measurement error. Also, sex offenders holding strong cognitive distortions may not score as deceptive because of lower emotional reactions (Cumming & McGrath, 2005). Perhaps most notably, polygraph-induced disclosures have not yet been linked to sexual recidivism (Cumming & McGrath, 2005).

***Self-report measures.*** Questionnaires have been created to measure sexual interest in children, such as the Sexual Interest Profiling System (SIPS; Laws, 1986). The SIPS assesses self-reported sexual attraction to various descriptions of sexual behaviours (i.e., adult consenting sex, pedophilia, incest, rape, masochism, sadism, voyeurism, exhibitionism, and frottage). The SIPS has been found to significantly distinguish child molesters with male victims from child molesters with female victims (Laws, Hanson,

Osborn, & Greenbaum, 2000). In addition, the SIPS has been found to add predictive accuracy above PPG when classifying child molesters as boy-object or girl-object child molesters (Laws et al., 2000).

Although self-report measures offer an easier technique to collect information, a significant limitation of this method is its dependence on individual responses. Individuals may have limited access to their cognitions and, consequently, may not be able to accurately report and describe certain experiences (Fazio & Olson, 2003; Greenwald & Banaji, 1995; Nosek & Smyth, 2007; Ward, Hudson, Johnston, & Marshall, 1997). Offenders may also distort their responses to be perceived in a socially desirable manner. Inaccurate responding is especially problematic when the construct under investigation is a sensitive topic (Nosek, Greenwald, & Banaji, 2005; Poehlman, Uhlmann, Greenwald, & Banaji, 2005). Indeed, child molesters who deny committing their offence tend to report less sexual interest in children than those who admit their offence (Baldwin & Roys, 1998). Moreover, child molesters are more likely to report cognitive distortions when attached to a bogus lie detector, suggesting that some child molesters are deceitful in their responses (Gannon, Keown, and Polaschek, 2007).

Recent research has created ambiguity between the long standing assumption that measures of social desirability assess response biases and that controlling for these biases increases the validity of the particular finding. Contrary to these assumptions, studies have found that higher risk offenders have lower social desirability as measured by both subscales of the Balanced Inventory of Desirable Responding (BIDR; e.g., Mills, Loza, & Kroner, 2003). The impression management subscale of the BIDR had the strongest relationship with criminal risk (Mills & Kroner, 2006). Mills and Kroner (2003) found

that controlling for social desirability decreased the correlation between offenders' self reports and criminal risk. Importantly, the link between criminal risk and social desirability continued to be present even after removing BIDR items correlated to criminal behaviour, as measured by the Statistical Information on Recidivism (SIR; Nuffield, 1982) scale (Mills & Kroner, 2006). Studies such as these suggest that scales measuring social desirability may actually be assessing enduring personality traits or defensive responses that are linked to lower recidivism rates rather than response bias. Until these findings are better understood, it is important for researchers not to simply control for social desirability in statistical analyses, but examine its relationship with criminal risk (see Tan & Grace, 2008 for a review).

In short, the usefulness of self-report measures can be limited by their susceptibility to deception, self-presentation biases, and reliance on introspection, especially in adversarial-type settings such as pre-sentencing reports (Greenwald & Banaji, 1995; Paulhus, 1984; Ward et al., 1997). Consequently, researchers express reservations about the use of self-report measures for offenders, especially without the inclusion of other measures less susceptible to presentation bias (Horley, 2000; Marshall, Anderson, & Fernandez, 1999).

**Viewing time measures.** Viewing time measures involve the presentation of pictures of nude or semi-nude individuals representing different age groups. The time spent viewing each picture is unobtrusively recorded while the participants rate the attractiveness of the individual represented in the picture (Abel, Huffman, Warberg, & Holland, 1998). Specifically, the time interval between the presentation of the stimuli and the participant's response is measured. The time spent viewing each picture is theorized

to be associated with interest, with longer viewing time indicating greater interest (Abel et al., 1998).

Viewing time measures are significantly correlated to self-reported attractiveness ratings of the pictures presenting during the procedure as well as PPG-assessed sexual interest, thereby demonstrating convergent validity (Abel et al., 1998; Quinsey, Ketsetzis, Earls, & Karamanoukian, 1996; Stinson & Becker, 2008). Furthermore, viewing time measures accurately distinguish child molesters from non-molesters (Harris, Rice, Quinsey, and Chaplin, 1996), as well as homosexual men from heterosexual men (Quinsey et al., 1996). In addition, Gress (2005) found that viewing time possesses greater accuracy than the SIPS in distinguishing child molesters from rapists. Harris et al. (1996) compared the ability of viewing time, PPG, and a self-report measure of sexual interest in children to differentiate between child molesters and community men. Although both PPG and the viewing time measure significantly discriminated between groups, PPG demonstrated slightly greater accuracy in discriminating between groups compared to viewing time and the self-report measure. Lastly, child molesters and community men did not differ on their self-reported attractiveness ratings of the pictures presented during the viewing time measure (Harris et al., 1996). These studies support the contention that responses on self-report measures of sexual interest can be biased and that indirect measures, such as viewing time measures, allow for more accurate measurement of the construct of interest.

Even when valid measures of sexual interest in children are available, the validity of research and clinical assessments is bolstered by a multi-method approach (e.g., Kazdin, 2003). For instance, Seto, Cantor, and Blanchard (2006) found that self-

reported sexual interest, sexual history reports, and possessing child pornography all independently contributed to the prediction of PPG scores. Stinson and Becker (2008) reported that the Child Molest Scale of the Multiphasic Sex Inventory –Second edition (MSI-II, Nichols & Molinder, 2000; A self-report measure) correctly classified 91.7% ( $n = 55$ ) of 60 sex offenders (defined as compliant in the treatment program) in their respective offence type (i.e., child molesters, sex offenders against an adult, exhibitionists, and sexual assault involving sadism). These authors found that the Child Molest scale of the MSI-II and self-reported fantasies significantly predicted sex offenders convicted of child sexual assault better than either one of these measures alone. In addition, although PPG and the viewing time measure did not add significant predictive validity, removing these measures from the model reduced predictive accuracy (Stinson & Becker, 2008). Lastly, Laws et al. (2000) correctly classified 100% of girl-object child molesters ( $n = 52$ ) and 70% of boy-object child molesters ( $n = 20$ ) when using both self-reported sexual interest and PPG scores. Taken together, these studies suggest that a multi-method approach allows more accurate discrimination and that self-report measures are helpful, at least in less adversarial type settings (e.g., offenders defined as compliant in Stinson & Becker, 2008).

Sexual interest in children has mostly been studied using physiological and self-report measures, neglecting the potential role of implicit cognitions. Whereas explicit cognitions are accessible through introspection and are assessed using self-report measures, implicit cognitions are typically assessed with reaction time measures. Implicit measures avoid reliance on introspective access, thereby lessening individuals' control over their responses (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005; Nosek,

Greenwald, & Banaji, 2007). Implicit cognition consists of the underlying associations between concepts, such as the association between children and sexual attractiveness.

Both implicit and explicit cognitions influence individual behaviour. Fazio's (2001) *Motivation and Opportunity as Determinants (MODE) model* explains the processes through which attitudes influence both judgements and behaviours. The MODE model posits that attitudes influence behaviour either spontaneously or under more deliberative processes. Spontaneous processes automatically influence particular judgment or behaviour in the immediate environment. In contrast, deliberative processes involve effort and a cost-benefit analysis of the judgment or behaviour that will occur. Motivation is required for deliberative processes to arise and influence behaviour. This model is useful in explaining the difference between implicit and explicit measures; implicit measures rely on the automatic process, whereas self-report measures rely on the deliberative process. Implicit and explicit cognitions appear to be correlated yet distinct constructs (Cunningham, Preacher, & Banaji, 2001; Greenwald & Farnham, 2000; Nosek & Smyth, 2007). In short, a more complete understanding of deviant sexual interest may be possible by complementing commonly used methods of sexual interest with implicit measures. The added information could improve our understanding of the construct of sexual interest in children, and, consequently, the assessment and treatment of child molesters.

### **The Implicit Association Test (IAT)**

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is an implicit measure first used in the field of social psychology to measure sensitive topics, such as racism. IAT measures have been used to examine constructs pertinent to

the fields of social psychology (e.g., Rudman, Greenwald, Mellott, & Shwartz, 1999), neuropsychology (Phelps et al., 2000), clinical psychology (e.g., Teachman, Gregg, & Woody, 2001), personality psychology (Asendorpf, Banse, & Mücke, 2002), health psychology (e.g., Sherman, Rose, Koch, Presson, & Chassin, 2003), and consumer psychology (e.g., Maison, Greenwald, & Bruin, 2004).

The IAT measure assesses the strength of cognitive associations by comparing reaction times to different pairings of the concepts of interest. IAT measures do not require direct questioning of participants and, thus reduce the impact of conscious intention or deliberative processes on responses (Nosek, Greenwald, & Banaji, 2007). Other advantages of the IAT measure are that it is easy to administer, inexpensive, and minimally affected by attempts at faking (Kim, 2003; Nosek et al., 2007).

The IAT measure works as follows: the strength of automatic associations in memory between a concept category (e.g., *adult* and *child*) and an attribute category (e.g., *sexy* and *not sexy*) is inferred from the relative speed with which one sorts stimulus words or pictures into categories. As such, the IAT measure is essentially a categorization task. Respondents sort each word or picture into one of four categories by pressing one of two keys on a computer keyboard (e.g., key *d* or *k*). Thus, two categories are indicated by one key, whereas the other key indicates the remaining two categories. Response speed is thought to depend on the extent to which the categories that share one key are associated in one's memory. Faster responses are found for tasks relying on categories that are associated (e.g., *chocolate* and *good*) and slower responses are found for tasks that are less associated (e.g., *liver* and *good*).

The IAT is composed of trials and blocks. A trial is a presentation of a single stimulus (either a picture or a word) that the participant must sort into the respective categories, whereas a block consists of several trials (usually 20 or 40 trials). In some blocks, the attribute and concept categories are paired in a manner that could be perceived as compatible for some (e.g., *adult* and *sexy* share the same response key) or incompatible for others. The pairing is also done in the reversed manner (i.e., *child* and *sexy* sharing the same response key). Figure 1 is an example of an IAT measure adapted to assess sexual interest in children. The main task is to sort the stimuli into one of the following categories: *child*, *adult*, *sexy*, or *not sexy*.

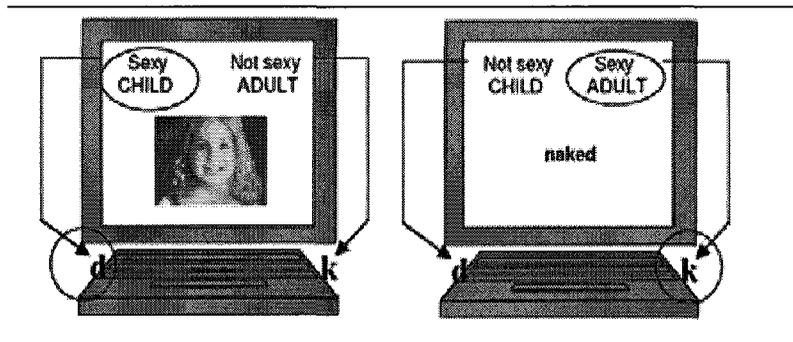


Figure 1. An example of an IAT adapted to assess sexual interest in children.

In the trial showed in the first screen of Figure 1, the picture represents a child and therefore would be sorted in the *sexy/child* categories. In the trial showed in the second screen of Figure 1, the stimulus word *naked* represents the sexy category and would therefore be sorted in the *adult/sexy* categories. For someone who is primarily sexually attracted to children, response speed should be quicker when *sexy* and *child* share the same response key (as in the first screen in Figure 1) than when *sexy* and *adult* share the same response key (as in the second screen in Figure 1). Conversely, for

someone who is primarily sexually attracted to adults, the reverse would be expected. For someone who is attracted to both adults and children equally, the response speed should be similar for both type of pairings. In short, the difference between the two attribute-concept pairings is used to assess the relative strength of the association.

**The construct validity of IAT measures.** Research has demonstrated that IAT measures can reliably predict behaviour (Houben & Wiers, 2006; Jajodia & Earleywine, 2003; Palfai & Ostafin, 2003; Teachman & Woody, 2003; Uhlmann & Swanson, 2004). For instance, IAT measures adapted to assess the evaluation of alcohol were associated with alcohol consumption and alcohol problems (Houben & Wiers, 2006; Jajodia & Earleywine, 2003; Palfai & Ostafin, 2003). Research also suggests that the IAT measure is sensitive to changes in cognitions. For example, an IAT measure adapted to assess the association between aggression and the self (i.e., aggressive self concept) found that aggressive self concept increased after playing a violent video game (Uhlmann & Swanson, 2004). Moreover, Teachman and Woody (2003) have found that scores on an IAT measure adapted to assess fear of spiders decrease during the course of a phobia treatment program.

In an attempt to assess the validity of the IAT measure, several researchers contrasted their IAT measures to the corresponding self-report measures. Nosek and Smyth (2007) reported the convergent validity between five different IAT measures and their respective self-report measures to be moderate to high, ranging from  $r = .27$  to  $r = .56$ ,  $p < .05$ . Similarly, a meta-analysis of 126 studies found a moderate correlation between the IAT and explicit self-report measures,  $r = .24$  (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005).

Hofmann et al. (2005) concluded that the relative relationship between IAT and self-report measures suggests that the measures are related; however, their relationship may be moderated by certain variables, such as social desirability. Indeed, the most prominent explanation for the sometimes low correlation between implicit and explicit measurements (e.g., Rudman et al., 1999; Nosek and Smyth, 2007) is that, unlike explicit measures, implicit measures are not biased by socially desirable responding (Gawronski, LeBel, & Peters, 2007). Research using IAT measures has found it to be less vulnerable to deliberate attempts at dissimulation as compared to self-report measures (Asendorpf et al., 2002; Greenwald & Farnham, 2000; Greenwald et al., 2003; Poehlman et al., 2005; Steffens, 2004). Although participants instructed to slow their responses in one of the IAT measure's two combined tasks to simulate desirable responding successfully produced faked scores, naïve participants (i.e., those asked to fake the IAT measure without specific instructions) did not discover this strategy (Kim, 2003; Steffens, 2004).

Motivational factors have also been found to moderate the relationship between implicit and explicit measurements (Hofmann et al., 2005). Namely, when participants are motivated, such as when accessing a socially sensitive topic, and can control their responses in the respective self-report measure, the correlation between the IAT and a self-report measure was low (Hofmann et al., 2005). In contrast, if the motivation is low, such as in an IAT measure assessing individuals' preferred colour, a higher correlation between self-reported and the IAT measure was found (Gawronski et al., 2006; Hofmann et al., 2005). In addition, Hofmann et al.'s (2005) meta-analysis found that the field of consumer research and group attitudes application (i.e., low motivation for control) had

the largest correlation between self-report and IAT measures. The lowest correlations were reported in the field of stereotype and self-esteem (i.e., high motivation for control).

If social desirability affects self-report measures but does not affect implicit measures such as the IAT measure, it would be expected that using social desirability as a moderating variable would increase the correlation between these two methods (e.g., Gawronski, LeBel, & Peters, 2007). Correspondingly, Hofmann et al. (2005) found that social desirability was a significant moderator that, when taken into account, increased the relationship between the IAT and self-report measures in a sample of non-offenders. These studies indicate that IAT measures are robust to dissimulation. Issues affecting the reliability and validity of IAT measures are summarized in Appendix A.

**The IAT adapted to assess sexual attraction to children.** The IAT measure has only recently been adapted to assess sexual interest in children (Banse, Schmidt, & Clarbour, 2009; Brown, Gray, & Snowden, 2009; Gray, Brown, MacCulloch, Smith, & Snowden, 2005; Mihailides, Devilly, & Ward, 2004; Nunes, Firestone, & Baldwin, 2007; Steffens, Yundina, & Panning, 2008).

Mihailides et al. (2004) first used the IAT measure to examine implicit cognitions theorized to have an influence on the initiation and maintenance of child sexual abuse: (1) endorsing a view of children as sexual beings, (2) uncontrollability of sexuality, and (3) sexual entitlement. With respect to the IAT measure adapted to examine the perception of children as sexual beings, the authors compared 25 men convicted of child sexual abuse, 25 men convicted of non-sexual offences, 25 male university students, and 25 female university students in terms of implicitly assessed sexual interest in children. The authors used the categories *not children* versus *children* and *sexual* versus *not sexual*. Examples

of the stimulus words for each of the categories are *cauldron* for not children, *kids* for children, *lust* for sexual, and *direction* for not sexual. Mihailides et al. found that the Children as Sexual Beings IAT was able to distinguish between groups. Specifically, child molesters viewed children as more sexual than non-sex offenders ( $d = 0.63$ ), male university students ( $d = 0.92$ ), and female university student ( $d = 0.97$ ). The construction of Mihailides et al.'s IAT measure, however, had several weaknesses that have been identified by Nosek et al. (2007).

Mihailides and colleagues did not use contrasting concept categories (e.g., *adult* and *child*); instead, they used unrelated categories (e.g., *not children* and *children*). Using unrelated categories does not allow for the examination of relative sexual interest. Instead, differences in reaction times between pairings of the concepts *children* and *sexy* relative to *adult* and *sexy* may produce a better examination of relative sexual interest. The use of a neutral category (e.g., *not children*) assumes that results can be interpreted as an absolute assessment of the concept of interest (e.g., the association between *children* and *sex*); however, this can only be true if the neutral contrasting category contributes no meaningful variability to the measurement (e.g., Jajodia & Earlywine, 2003). Given that unrelated categories can contribute variability to the IAT effect, Greenwald (2004) and Nosek et al. (2007) suggest against the use of unrelated categories in the construction of IAT measures.

In addition, the stimulus words used by Mihailides and colleagues to represent the *children* category were phonetically similar to those used to represent the *not children* category (e.g., *kids* and *lids*), so to decrease variance between the categories. This was also true of stimuli words representing the *sexual* and *not sexual* (e.g., *lust* and *dust*)

category. Instead, it is possible that this characteristic reduced coherency and contrast between the categories, thereby increasing errors associated with sorting difficulty and caused participants to use other sorting rules (e.g., guessing or examining font as opposed to the words themselves; Nosek et al., 2007). Therefore, the internal validity of this study may have been compromised.

Gray et al. (2005) created an IAT measure to assess the degree to which children, relative to adults, were associated with sex. The IAT measure was administered to 18 child molesters and 60 non-child molesters (including rapists and non-sex violent offenders). Gray et al. used the categories *adult* versus *child* and *sex* versus *not sex*. Examples of their stimulus words for each of the categories are *beard* for adult, *infant* for child, *breasts* for sex, and *elbow* for not sex. Gray et al. found that child molesters' response latencies were significantly slower on the trials in which *adult* and *sex* shared the same response key than on trials in which *child* and *sex* shared the same response key. In contrast, non-molesters showed the opposite pattern. This finding suggests that child molesters had a stronger association between *child* and *sex* compared to *adult* and *sex*. The magnitude of the difference between the child molesters and non-molesters was large ( $d = 0.83$ ). Moreover, the IAT measure demonstrated high predictive validity in distinguishing between child molesters and non-child molesters (Area Under the Curve = .73; a large effect based on the criteria from Rice & Harris, 2005).

Nunes and colleagues (2007) also used the IAT measure to examine cognitions related to child sexual abuse. Participants were 27 child molesters and 29 non-sex offenders. The IAT measures were adapted to assess: (1) evaluation of self, (2) evaluation of children, (3) social power of adult, (4) social power of children, (5) sexual

attractiveness of self, and (6) sexual interest in children. This review will focus on the IAT measure designed to assess sexual interest in children. Nunes and colleagues used the categories *adult* versus *child* and *sexy* versus *not sexy*. Examples of their stimulus words for each of the categories are *mature* for adult, *young* for child, *beautiful* for sexy, and *cold* for not sexy. Nunes et al. found that child molesters perceived children to be significantly more sexually attractive compared to the non-sex offenders (medium effect size,  $d = 0.66$ ), results that are congruent with those of the above mentioned studies.

Further, the Sexy Child IAT was positively associated with risk of sexual recidivism as measured by the Static-99 (Hanson & Thornton, 2000),  $r = .43, p < .05$ . Nunes and colleagues hypothesized that the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997) would be more strongly correlated to their IAT than the Static-99. Briefly, the RRASOR has been described as a better measure of sexual deviancy than the Static-99, with three of the four items (e.g., male victims, unrelated victims, prior sexual offending) described as proxies for sexual deviancy (Doren, 2004; Roberts, Doren, & Thornton, 2002). Although these items are also included in the Static-99, the non-RRASOR items of the Static-99 are correlated with Psychopathy Checklist-Revised (PCL-R; Hare, 1991) scores ( $r = .53, p < .001$ ) and hence are more representative of antisociality, whereas the RRASOR items are not correlated to the PCL-R scores ( $r = .03, p > .05$ ; Roberts et al., 2002). Despite being demonstrated as a better measure of sexual deviancy, the correlation coefficient between the RRASOR and the Sexy Child IAT was small to moderate and did not reach statistical significance,  $r = .27$ . Such a finding could be due to low power ( $n = 24$ ), which increases the chance of finding no significant differences, even if one may exist. Further exploring the relationship

between the RRASOR and an IAT measure of sexual interest in children would be informative.

It is important to note that Nunes et al. (2007) administered six different IAT measures and testing spanned two days. Experience with the IAT measure has been found to result in reduced IAT effects (Greenwald & Nosek, 2001). Although the presentation order of the six IAT measures were counterbalanced, the low number of participants did not allow for a statistical examination of whether the IAT measures administered first resulted in larger IAT effects than those administered in the second testing session. It is possible that the administration of multiple IAT measures would have resulted in an overall reduction in IAT effects, potentially reducing the magnitude of effects.

Steffens et al. (2008) constructed an IAT measure adapted to assess sexual interest to children relative to women. The IAT was tested in a community sample of sex offenders in Germany. The concept categories were *women* versus *child* and the attribute categories were *erotic* versus *not erotic*. Four stimuli represented the concept categories (*woman, women, and child, children*) and four represented the attribute categories (*erotic, arousing, not erotic, asexual*). Groups included 16 child molesters diagnosed by a psychologist as “exclusively pedophiles” and 30 sex offenders who were diagnosed as “not exclusively pedophiles”, of which 21 were rapists and 9 were child molesters. The exclusively pedophilic diagnosis was based on psychological variables (e.g., type of sexual fantasies, problems with adult relationships), rather than just offence characteristics (Melanie Steffens, personal communication, March 18, 2009). Groups were found to be significantly different, with the pedophilia group showing greater

association between children and sexual attractiveness than the non-pedophilic group,  $d = 0.53$  (95% C.I. = .08 - 1.15).

Steffens et al. (2008) was the first study to separate groups into sex offenders diagnosed as pedophilic and those diagnosed as not primarily pedophilic. Previous studies selected groups based on behaviour (i.e., men charged or convicted for child sexual abuse and those not charged or convicted for such offences). The diagnosis of pedophilia, however, was based on clinical assessment and little information as to the basis of the diagnoses or the reliability of the diagnoses was available. Importantly, offenders with sexual interest in children may have been included in the “not exclusively pedophilia group”. This may have reduced group differences and, in fact, compared to the other IAT measures of sexual interest in children, the Children-Women IAT has the smallest effect size.

Banse et al. (2009) recently developed two IAT measures (Girls-Women IAT and Boys-Men IAT) of sexual interest in children using pictures of children and adults in swimming suits. Banse et al. also administered a viewing time measure and examined the relationship between the IAT measures, viewing time, and group membership. The Girls-Women IAT (composed of only female pictures) was able to distinguish between groups ( $d = 0.71, p < .05$ ), whereas the Boys-Men IAT (composed of only male pictures) did not distinguish between groups ( $d = 0.37, p > .05$ ). Interestingly, the Girls-Women IAT was correlated to the viewing time measure ( $r = .26, p \leq .05$ ) and the Screening Scale for Pedophilic Interest (SSPI; a file based measure of sexual deviancy),  $r = .38, p \leq .05$ . The Boys-Men IAT was not correlated to any of these measures.

The lack of difference in the Boys-Men IAT can be attributed to the fact that only two men in the comparison group (2.7%,  $n = 75$ ) and 12 child molesters (31.6%,  $n = 38$ ) reported a homosexual orientation and that, on average, child molesters are more likely to display sexual interest in female rather than male children (Seto, 2008). If both categories (i.e., boys and men) are equally rated in sexual attractiveness (or lack thereof) for both groups, the IAT difference score cannot distinguish between groups. With only 38 child molesters, it is possible that the base rate of sexual interest in boys was too low for the groups to be found significantly different. In other words, the lack of difference may not be due to the psychometric properties of the Boys-Men IAT, but to the fact that the two groups were truly similar.

In addition, the IAT measures reported in Banse et al. used sexually explicit pictures of children and adults in swimming suits. In the construction of an IAT measure, it is important that each stimulus unambiguously belong to their respective categories (Nosek et al., 2007). It is plausible that the use of pictures of children and adults in swimming suits equally represents *sexy* or *not sexy* and *adult* or *children*. Using faces of children and adults may rectify or at least reduce the ambiguity between the categories. Therefore, using faces as opposed to sexually explicit pictures may allow for a better constructed IAT measure and, consequently, a more valid measure of sexual interest in children. In addition, better measurement without the use of sexually explicit stimuli is advantageous in terms of easier compliance in research as well as the ethical considerations involved in using sexually explicit materials.

Brown et al. (2009) extended the examination of the IAT measure to evaluate its ability to distinguish between child molesters with victims under the age of 12 years and

child molesters with victims between 12 and 15 years of age. This study also examined the effect of denial on the discriminative ability of the IAT measure. The authors used the categories *adult* versus *child* and *sex* versus *non-sex*. In this IAT measure, pictures were used to represent the categories *adult* and *child*. Examples of the stimuli are *kiss* for sex and *run* for non-sex. Fifty-four child molesters with victims under 12 years of age, 21 child molesters with victims between 12 and 15 years of age, and 49 non-sex offenders completed this IAT measure. Of the child molesters with victims 15 years of age or less, there were 55 deniers and 20 admitters, as determined by interviews and the Sex Offence Information Questionnaire (Hogue, 1994).

Brown et al. demonstrated that child molesters with victims under the age of 12 had a higher association between child and sex compared to both child molesters with victims between 12 and 15 years of age ( $d = 0.77$ ) and non-sex offenders ( $d = 0.92$ ). The magnitude of the differences between groups was large. Moreover, the IAT measure had high predictive validity in distinguishing between child molesters and non-sex offenders (AUC = .74). The IAT measure was also able to differentiate child molesters who denied their offences from non-sex offenders ( $d = 1.01$ ). In addition, child molesters described as deniers did not significantly differ from admitters on the IAT ( $d = 0.27$ ). Brown et al.'s findings suggest that IAT measures can accurately assess child molesters who deny their offences. This is consistent with findings in non-offender populations, in which the IAT measure is relatively unaffected by social desirability and general dissimulation (Asendorpf et al., 2002; Greenwald et al., 2003; Greenwald & Farnham, 2000; Poehlman et al., 2005; Steffens, 2004).

Mihailides et al. (2004), Nunes et al. (2007), and Steffens et al. (2008) had rather low statistical power, whereas Banse et al. (2009), Gray et al. (2005), and Brown et al. (2009) had high power as measured by G\*Power 3.0.10 (Faul, Erdfelder, Lang, & Buchner, 2007; see Table 2).

*Table 2*

Statistical Power of Six Studies Examining IAT Measures of Sexual Interest in Children

Study	$N_{\text{child molester}}$	$N_{\text{nonmolester}}$	Effect Size	Power
Banse et al. (2009)	33	33	$d = 0.71^a$	.887
Brown et al. (2009)	54	49	$d = 0.92$	.996
Gray et al. (2005)	18	60	$d = 0.83$	.862
Mihailides et al. (2004)	25	25	$d = 0.63$	.588
Nunes et al. (2007)	27	29	$d = 0.66$	.679
Steffens et al. (2008)	16	30	$d = 0.53$	.388

*Note.* A two-tailed alpha of .05 and a two-group design was used to estimate power.  $d$  represents the effect size of the respective IATs in comparisons between child molesters and non-sex offenders. <sup>a</sup>  $d$  based on the Girls-Women IAT.

**Meta-analysis.** A meta-analysis of previous studies was conducted to synthesize the findings (Banse et al., 2008; Brown et al., 2009; Gray et al., 2005; Mihailides et al., 2004; Nunes et al., 2007; Steffens et al., 2008) utilizing the procedures and formula outlined by Hanson and Broom (2005). A moderate effect size was found ( $Mean d = 0.73$ , 95% C.I. = 0.54 – 0.93) for the ability of these IAT measures of sexual interest in children to distinguish child molesters from non-sex offenders. Furthermore, findings

were consistent across the six studies. In other words, the variability between studies was not greater than that expected by chance,  $Q = 1.52, p > .05$ .

Despite the fact that the IAT measures were constructed differently and studies possessed different levels of statistical power, the fact that each found significant group differences suggests that child molesters and non-sex offenders do indeed differ in their IAT effect. Even though these studies provide an optimistic view on the usefulness of the IAT measures of sexual interest in children, additional analyses are required to fully understand the psychometric properties of these types of IAT measures. Namely, these studies examined the ability of the IAT measures of sexual interest in children to differentiate between groups of child molesters and non-sex offenders. Although this constitutes an important preliminary step, it is also necessary to gauge the convergent validity of these types of IAT measures by comparing them to other commonly used measures of sexual interest, such as PPG and self-report measures. Greater confidence that these types of IAT measures are truly assessing sexual interest could be obtained by demonstrating that they are related to other measures of sexual interest.

### **Present Study**

Sexual interest in children has been identified as an important cause of child sexual abuse (e.g., Hall & Hirshman, 1992; Ward & Keenan, 1999). As mentioned previously, even when valid measures of sexual interest in children are available, assessments and research can be strengthened by a multi-method approach. Several studies highlight the ability of the IAT measure of sexual interest in children to distinguish between child molesters and non-sex offenders. In addition, research in social psychology found that IAT measures provide independent information from self-report

measures (Greenwald & Farnham, 2000; Nosek & Smyth, 2007). As such, an IAT measure of sexual interest in children may allow for a more comprehensive assessment of sexual interest when used in collaboration with PPG and self-report measures. Research, however, is needed to fully evaluate the psychometric properties of an IAT measure adapted to assess sexual interest in children.

The main purpose of the current study was to examine whether differences existed between child molesters and non-sex offenders on the Sexual Attraction to Children IAT (SAC-IAT), which was developed for the current study. In addition, this study also examined the convergent validity of the SAC-IAT with other measures of sexual interest as well as the internal consistency of the SAC-IAT. The secondary purpose of this study was to examine differences in the SIPS and a viewing time procedure (developed for the current study) between child molesters and non-sex offenders. The latter analyses were completed in order to better explain the relationship between the SAC-IAT and the secondary measures of sexual interest. For example, if the SIPS does not distinguish between groups because all participants report no sexual interest in children, then it could explain a lack of correlation with the SAC-IAT. Lastly, for the child molester group, the relationship between the measures of sexual interest and risk of sexual recidivism was also examined.

In short, the goal of the proposed research was to carefully develop, refine, and evaluate the validity of an adaptation of an IAT measure of sexual attraction to children.

### **Hypotheses**

**Primary hypotheses.** Based on previous findings, it was hypothesized that child molesters and non-sex offenders would be significantly different in their SAC-IAT

scores. Second, it was hypothesized that group differences would exist in Viewing Time, with child molesters viewing pictures of children significantly longer than non-sex offenders. It was also expected that groups would not respond significantly differently on the SIPS. Further, it was hypothesized that the SAC-IAT would be positively correlated with Viewing Time, PPG, and the SIPS.

**Secondary hypotheses.** It was hypothesized that, among child molesters, greater scores on the SAC-IAT would be associated with greater risk for sexual recidivism, as measured by validated risk assessment instruments.

## Method

### Participants

Participants were recruited at five federal institutions in Ontario. A total of 201 offenders were called into the psychology building of the respective institutions. Of these, 46 offenders did not present themselves at their scheduled appointment in order to meet the researchers. Of those who met with the researchers ( $n = 155$ ), a total of 79 consented and completed the study. Participation was voluntary and informed consent was provided by the participant before testing. Prior to data collection, ethics approval was granted by Carleton University and Correctional Services Canada (CSC). Consent rates differed based on the gender of the researchers, with the female researcher having a higher consent rate (65%; 44 of 68) than the male researcher (41%; 34 of 84). Gender of the researcher, however, was not related to any of the measures in the present study. Of those who did not consent to the study, the majority did so after the researcher gave an overview of the study (88%; 65 of 74).

The sample consists of 21 non-sex offenders (no history of charges or convictions for sex offences or self-reported sex offences), 35 child molesters (victim under the age of 12 in the index offence), 18 sex offenders with victims 12 years of age or older (10 sex offenders with victim 18 years of age or older and 8 sex offenders with victim between 12 and 17 years of age), and 5 non-sex offenders who reported perpetrating a sexual assault of which there was no record in their official files. All participants had adequate reading ability as measured by the Clarke Vocabulary Scale (Paitich, 1977). Table 3 presents the descriptive statistics of the sample.

Table 3  
Description of Participants

	Sex Offenders			Non-Sex Offenders				
	Victim < 12	Victim 12-17	Victim 18+	No Self-Reported Sex Offences	Self-Reported Sex Offences	Sex		
	Mean (SD)	Mean (SD)	Mean (SD)	N	N	N		
Age at testing	50.24 (12.33)	43.83 (14.15)	38.73 (11.02)	10	39.74 (13.15)	21	37.38 (8.53)	4
Education Level	10.09 (2.40)	10.50 (1.51)	10.11 (1.76)	9	10.80 (1.70)	20	10.00 (1.41)	4
Clarke Vocabulary <sup>1</sup>	29.63 (5.26)	27.63 (5.71)	26.90 (7.95)	10	29.48 (4.94)	21	24.67 (6.81)	3
SIR	9.58 (9.11)	3.71 (7.50)	1.67 (8.59)	9	1.15 (11.53)	20	-0.67 (16.17)	3
BIDR	13.14 (5.96)	11.29 (5.50)	14.80 (6.02)	10	14.10 (7.64)	21	20.00 (3.74)	4
BIDR-IM	6.86 (4.04)	5.86 (3.02)	6.50 (3.31)	10	6.00 (4.74)	21	8.00 (1.83)	4
BIDR-SD	6.29 (2.77)	5.43 (3.05)	8.30 (3.50)	10	8.10 (3.53)	21	12.00 (2.83)	4
Sentence Length <sup>2</sup>	5.27 (2.71)	4.26 (0.99)	4.75 (1.89)	7	4.57 (2.48)	12	11.23 (11.78)	4
Sexual Prior Convictions <sup>3</sup>	8.85 (9.58)	4.00 (4.28)	3.70 (3.83)	10	0.00 (0.00)	21	0.00 (0.00)	4
Non-sexual Prior Violent Convictions <sup>3</sup>	1.03 (1.31)	2.25 (3.69)	3.00 (2.26)	10	3.62 (3.89)	21	8.50 (7.33)	4
Non-sexual Prior Non-violent Convictions <sup>3</sup>	1.32 (2.38)	8.13 (8.84)	8.50 (21.02)	10	19.71 (30.10)	21	11.00 (12.19)	4

Note. Numbers fluctuate due to missing data. <sup>1</sup>Clarke Vocabulary scores below 14 indicate a low vocabulary level. <sup>2</sup>Sentence length computed in years and do not include offenders with indeterminate sentences. Nine non-sex offenders, 12 child molesters with victim under 12, 2 sex offenders with victim between 12 and 17, and 3 sex offenders with victim equal to or over 18 had indeterminate sentences. <sup>3</sup>Numbers of convictions included both index and prior convictions. BIDR-IM: Balanced Inventory of Desirable Responding: Impression Management Subscale. BIDR-SD: Balanced Inventory of Desirable Responding: Social Desirability Subscale.

## Measures

**Sexual Attraction to Children IAT (SAC-IAT).** The SAC-IAT was designed to measure sexual attraction to children. This IAT was adapted such that pictures of adults' and children's faces were presented as exemplars of the *adult* and *child* categories. These faces were digitally modified pictures of children and adults (Not Real People Stimuli, Laws & Gress, 2004). Twelve exemplars represented the concept categories (*adult* and *child*), with three pictures per category (i.e., boy, girl, woman, and man). Correspondingly, six exemplars represented the attribute categories (i.e., *sexy* and *not sexy*). These exemplars are presented in Appendix B.

Table 4 presents the sequences of blocks of the SAC-IAT. The recommended numbers of trials indicated in Nosek et al. (2007) were followed. The SAC-IAT had seven trial blocks in which either the concept category (e.g., *adult* and *child*) and/or the attribute category (e.g., *sexy* and *not sexy*) were presented. Some blocks were practice blocks to accustom participants to both the stimuli and the sorting rules. The SAC-IAT had four critical blocks where the sorting of the attribute and concept category was accomplished simultaneously. The four critical blocks differed in terms of which concepts were combined (i.e., *adult* and *sexy* or *child* and *sexy*). Block 5 had 40 trials rather than 20 trials (the original method) because increasing the number of trials to 40 has been found to reduce the impact of the first combined task in generating stronger associations than the second combined task, when the compatible combination was completed first (Nosek et al., 2005; more detail regarding compatibility and its effect on the IAT measure is presented in Appendix A). Higher scores in the SAC-IAT represented greater sexual interest in children.

Table 4

## Blocks Presented in the SAC-IAT

Block	Number of Trials	Left-key response category	Right-key response category
1	20	Child	Adult
2	20	Not sexy	Sexy words
3	20	Child or Sexy	Adult or Not Sexy
4	40	Child or Not Sexy	Adult or Sexy
5	40	Adult	Child
6	20	Adult or Not Sexy	Child or Sexy
7	40	Adult or Sexy	Child or Not Sexy

*Note.* Blocks 3 and 4 are counterbalanced with blocks 6 and 7 between participants. Block 1 is counterbalanced with block 5, however, the block before the second combined task has 40 trials and the block before the first combined task has 20 trials.

**Sexual Interest Profiling System (SIPS).** The SIPS (Laws, 1986) assessed sexual interest via situational vignettes. It consisted of 19 scales and had a total of 130 items. There were 20 items for the adult consenting, pedophilia, and incest scales (each subscale had 10 items per gender). The rape, masochism, sadism, voyeurism, exhibitionism, and frottage scales had 10 items (5 items per gender). The rape, sadism, and masochism subscales only involved sexual scenarios with children. Lastly, the transvestism scale contained 10 items of male to female cross-dressing. Each vignette consisted of a description of a sexual behaviour (e.g., voyeurism item: “*You’re peering in through a bathroom window. A young boy is sitting on the toilet masturbating while he looks through a magazine*”). Participants responded by rating the attractiveness of the description on a 7-point Likert scale (ranging from *very unattractive* to *very attractive*). Scores for each scale range from 0 to 70, with higher scores indicating a higher

attractiveness rating for that scale. This measure was copyrighted and, therefore, was not included in the appendices.

The SIPS has been found to significantly differentiate between child molesters with boy versus girl victims and to contribute independently to the classification of child molesters above PPG (Laws et al., 2000). The SIPS, however, did not differentiate child molesters from rapists, correctly classifying only 3 of 19 child molesters (Gress, 2005).

**Viewing Time.** Viewing Time was constructed for the purpose of the current study. Participants were presented with a series of nude pictures via a laptop computer and asked to rate the sexual attractiveness of each picture on a 7-point Likert scale (ranging from very unattractive to very attractive). The pictures were drawn from the Not Real People Stimuli (Laws & Gress, 2004), and do not represent real people, but are instead computer generated pictures synthesizing several individuals into one. Although the pictures came from the same stimuli set used for the SAC-IAT, pictures included in the SAC-IAT were not included in Viewing Time. The inclusion of the same picture in both measures could reduce the internal validity of the study. Specifically, Viewing Time was always completed after the SAC-IAT (due to its sexual explicitness) and if some pictures were presented twice, it is possible that familiarity may have caused participants to examine the picture longer rather than interest. Consequently, time spent viewing the picture may no longer have been solely a measure of sexual interest.

Each person represented in the pictures corresponded to a specific Tanner Stage (i.e., development stage), depending on their external primary and secondary sex characteristics. The first three Tanner scales characterize children (i.e., Tanner 1 generally represents a person 0-3 years of age, Tanner 2 a person 4 to 7 years of age,

Tanner 3 represent a person 8 to 12, Tanner 4 represents a person from 13 to 15 years of age, and Tanner 5 represents a person age 16 or older). Each Tanner stage was represented by eight pictures (four for each gender), for a total of 40 pictures. Response latencies represented the time it took for participants to score the attractiveness of a particular person and to move to the next picture by pressing the space bar button of the laptop.

Time spent viewing a child picture relative to an adult picture has been found to distinguish child molesters from men from the community (Harris et al., 1996) and rapists (Gress, 2005). The internal consistency of one viewing time measure has been found to be strong, ranging from  $\alpha = .72$  to  $\alpha = .90$  depending on the category (e.g., adult woman) of the picture (Letourneau, 2002). Viewing time measures of sexual interest have been found to have strong convergent validity with other measures of sexual interest (Abel et al., 1998; Harris et al. 1996). For instance, when removing outliers, correlation between a viewing time measure and PPG ranged from  $r = .18$  for female adolescent stimuli to  $r = .48$  for male child stimuli (Letourneau, 2002). This measure, however, may be susceptible to faking once the clients learn or guess that the time spent viewing is calculated to assess sexual interest (Gress, 2005). In Banse et al. (2009), the viewing time measure was found to be correlated with the Girls-Women IAT,  $r = .26, p \leq .05$ .

**Viewing Time Attractiveness Ratings.** During the administration of the Viewing Time, participants were asked to rate the sexual attractiveness of the individual depicted in each picture on a 7-point Likert scale. Scores in each subscale range from 7 to 28, with higher scores indicating greater sexual attractiveness ratings.

**Penile Plethysmography (PPG).** As aforementioned, PPG is a physiological measure of sexual arousal and involves the presentation of various sexual behaviours, depicted either through audio, slides, or video, to infer sexual interest. PPG scores were noted during file-reviews and were available only for the child molesters. The comparison group (i.e., non-sex offenders) did not have any PPG scores available in their file because the PPG measure of sexual interest in children is only administered to offenders with a sexual offence against a child. The current study used PPG scores obtained from assessments conducted at the Millhaven Assessment Unit (MAU). MAU uses the Quinsey Child Sexual Violence Assessment. These stimuli have been found to have adequate internal consistency, ranging from  $\alpha = .74$  to  $\alpha = .94$  for intrafamilial offenders and from  $\alpha = .40$  to  $\alpha = .98$  for extrafamilial offenders (Fernandez, 2002). PPG has been found to possess strong criterion validity when comparing child molesters to non-molesters (Barbaree & Marshall, 1989; Quinsey & Chaplin, 1988) and to rapists (Looman & Marshall, 2001; Miner, West, & Day, 1995).

**Rapid Risk Assessment for Sex Offence Recidivism (RRASOR).** The RRASOR (Hanson, 1997) is a 4-item actuarial instrument designed to measure risk of sexual recidivism. Scores range from 0 to 6, with a higher score indicating greater risk of sexual recidivism. The items include: (1) prior sexual offenses, (2) any unrelated victims, (3) any male victims, and (4) offender is less than 25 years of age. The RRASOR was found to possess strong predictive accuracy for sexual recidivism,  $AUC = .71$  (Hanson, 1997). Barbaree et al. (2001) reported similar predictive validity for the RRASOR in predicting sexual recidivism ( $AUC = .76$ ,  $r = .26$ ). Similarly, Sjöstedt and Långström (2001) reported AUC of .73 for sexual and non-sexual violent recidivism. A recent meta-

analysis conducted by Hanson and Morton-Bourgon (2007) also found that the RRASOR was moderately associated with sexual recidivism (*Mean d* = 0.59).

**Static-99.** The Static-99 (Hanson & Thornton, 2000) is a 10-item actuarial instrument designed to assess risk of sexual recidivism. Items are static (i.e., not changeable through intervention) and were derived from the RRASOR and the Structured Anchored Clinical Judgment-Min (SACJ-Min; Grubin, 1998). Items include: (1) prior sex offences, (2) prior sentencing dates, (3) non-contact sex offense convictions, (4) non-sexually violent index offence convictions, (5) prior non-sexual violent convictions, (6) unrelated victims, (7) stranger victims, (8) male victims, (9) ever lived with a lover for two years, and (10) age. Scores range from 0 to 12, with higher scores indicating a higher probability of sexual recidivism.

The reliability and validity of the Static-99 have been demonstrated in numerous jurisdictions (Barbaree, Seto, Langton, & Peacock, 2001; de Vogel, de Ruiter, van Beek, & Mead, 2004; Doren, 2004; Hanson, 2001, 2006; Harris, Phenix, Hanson, & Thornton, 2003; Sjöstedt and Langström, 2001). A meta-analysis conducted by Hanson and Morton-Bourgon (2007) found that the Static-99 was associated with sexual recidivism (*Mean d* = 0.70).

**Stable-2000.** The Stable-2000 (Hanson & Harris, 2000) is an interview-based assessment design to evaluate dynamic risk factors for sexual recidivism. The six risk factors include (1) significant social influences, (2) intimacy deficits, (3) sexual self-regulation, (4) attitudes supportive of sexual assault, (5) cooperation with supervision, and (6) general self-regulation. The sexual self-regulation item refers to the presence of sexual deviance in children for child molesters and refers to the presence of sexual

deviance towards violent or coercive sex for sex offenders against adults. The coding of this item is largely based on PPG scores as well as self-reported sexual fantasies. The risk factors are drawn from previous research (e.g., Hanson, Gizzarelli & Scott, 1994; Quinsey, Coleman, Jones, & Altrows, 1997; Wilson, 1999; see review by Hanson & Morton-Bourgon, 2004, 2005) and risk assessment scales (Sex Offender Need Assessment Rating, SONAR, Hanson & Harris, 2000; STEP Deviance, Beech, Friendship, Erikson, & Hanson, 2002; and Structured Risk Assessment, SRA, Thornton, 2002). Each risk factor is evaluated using a three-point rating scale ranging from “0 – no problem,” “1 – some concern/slight problem” to “2 – present/definite concern”. Total scores range from 0 to 12 (0-4: Low; 5-8: Moderate; 9-12: High). The Stable-2000 was reported to have moderate predictive validity for sexual recidivism (AUC=.66, CI: .59 to .72), violent recidivism (AUC = .65, 95% C.I.: .59-.70) and general recidivism (AUC = .67, 95% C.I.: .63-.71; Hanson, Harris, Scott, & Helmus, 2007).

**Stable-2007.** The Stable-2007 (Hanson et al., 2007) is an interview- and file-review based instrument designed to assess dynamic (i.e., potentially changeable through intervention) risk factors for sexual recidivism. Items include: (1) significant social influences, (2) capacity for relationship stability, (3) emotional identification with children, (4) hostility toward women, (5) general social rejection, (6) lack of concern for others, (7) impulsivity, (8) poor problem solving skills, (9) negative emotionality, (10) sex drive/sex preoccupation, (11) sex as coping, (12) deviant sexual preference, (13) cooperation with supervision. The third item, emotional identification with children, is scored only for child molesters with victims aged 13 or less. The deviant sexual preference item refers to the presence of sexual deviance in children for child molesters

and refers to the presence of sexual deviance towards violent or coercive sex for sex offenders against adults. The coding of this item is largely based on PPG scores as well as self-reported sexual fantasies. Scores range from 0 to 26 for child molesters, with higher scores indicating larger risk of sexual recidivism.

The Stable 2007 has been found to have strong internal consistency,  $\alpha = .80$  (Hanson et al., 2007) and strong predictive validity for predicting sexual recidivism (AUC = .76, 95% C.I.: .69-.82) and for any recidivism, (AUC = .70, 95% C.I.: .66-.74; follow-up time of up to four years). The Stable-2007 better predicts sexual recidivism for extrafamilial child molesters (AUC = .77, 95% C.I.: .64-.90) than for intrafamilial child molesters, AUC = .58, CI: .31-.85 (Hanson et al., 2007). A Cox regression was conducted to test for the incremental validity of the Stable 2007 with the Static-99. Hanson et al. found that the Stable-2007 added a significant amount of information, beyond the Static-99, in predicting sexual recidivism,  $\beta = .06$ ,  $p = .05$ , and all type of recidivism combined,  $\beta = .07$ ,  $p < .001$  (Hanson et al., 2007).

**Statistical Information on Recidivism (SIR).** The SIR (Nuffield, 1982) is a measure of general recidivism consisting of 15 items. Score range from - 27 to + 30 and combined into 5 risk categories, ranging from very good risk to very poor risk, with lower scores indicating greater risk of general recidivism. More specifically, scores from 6 to 30 indicates a low risk of general recidivism, scores from 1 to 5 indicate a low-moderate risk, scores from -4 to 0 indicate a moderate risk, scores from -8 to -5 indicate a moderate-high risk, and, lastly, scores from -27 to -9 indicate a high risk of general recidivism. The SIR scale has been found to have good predictive validity, AUC = .74,

and is significantly correlated with general recidivism,  $r = .42, p < .001$  (Bonta, Harman, Hann, & Cormier, 1996).

**The Screening Scale for Pedophilic Interests (SSPI).** The SSPI (Seto & Lalumière, 2001) is a four-item scale constructed to assess pedophilia from information commonly found in official files. Scores range from 0 to 5, with a higher score indicating greater sexual deviancy. Items include (1) any male child victims, (2) more than one child victim, (3) any prepubescent victim (a child under the age of 12), and (4) any extrafamilial victims.

The SSPI has been found to accurately identify child molesters with pedophilic interests, as defined by PPG scores,  $AUC = .70$  (Seto & Lalumière, 2001). Seto, Harris, Rice, and Barbaree (2004) found a significant correlation between SSPI and PPG scores in child molesters,  $r = .28, p < .01$ . Similarly, Seto and Lalumière (2001) found a significant correlation between SSPI and PPG scores in child molesters,  $r = .34, p < .001$ . The strong relationship between scores in the SSPI and PPG has also been found in a group of adolescent child molesters,  $r = .46, p < .005$  (Seto, Murphy, Page, & Ennis, 2003). The SSPI has been found to be moderately correlated to an IAT measure of sexual interest in women relative to girls,  $r = .26, p \leq .05$  (Banse et al., 2009). The SSPI has also been found to be a moderate predictor of sexual recidivism ( $AUC = .62$ ) and violent recidivism ( $AUC = .67$ ), after an average follow-up of 5 years (Seto et al., 2004).

**Balanced Inventory of Desirable Responding (BIDR).** The BIDR (Paulhus, 1984) is a 40-item questionnaire that measures the tendency to respond in a socially favourable manner. It is comprised of two subscales, the Impression Management (IM; self-presentation bias to portray oneself in a favourable light) subscale and the Self-

Deception Enhancement (SDE; tendency to perceive oneself in a favourable light) subscale (Paulhus, 1984). Both subscales are comprised of 20-items and scored on a 7-point Likert scale, ranging from *not true* to *very true* (Appendix C). Half of the items are reverse scored and higher scores indicate greater social desirable responding.

The BIDR-IM has displayed strong internal consistency ( $\alpha = .84$ ,  $N = 539$ ) in forensic samples (Kroner & Weekes, 1996). Nunes et al. (2007) reported a similar internal consistency coefficient for child molesters ( $\alpha = .88$ ,  $n = 27$ ) and non-sex offenders ( $\alpha = .89$ ,  $n = 29$ ). The BIDR-SDE also demonstrates strong internal consistency ( $\alpha = .88$ ,  $n = 90$ ; Li & Bagger, 2007).

**Clarke Vocabulary Scale (CVS).** The CVS (Paitich, 1977) is a forty-item multiple-choice questionnaire designed to assess participants' reading comprehension. A series of words are presented and participants are required to select the most appropriate definition from four options (Appendix D). This measure was administered to verify that participants were able to understand the instructions and stimuli required to successfully complete the measures involved in the current study. A score less than 14 indicate low vocabulary level.

**Background Questionnaire (BQ).** The BQ was created for the purpose of the current study and involves a series of demographic questions regarding topics such as sexual orientation and past sexual offences. See Appendix E for complete list of items. This questionnaire was used to examine potential confounding variables that may impact the internal validity of this study, such as education, first language, treatment participation, and self-reported history of sexual offences against children or adults.

## **Materials and Apparatus**

The SAC-IAT and all other measures, with the exception of file-based measures (i.e., SSPI, Static-99, Stable-2007, SIR, and PPG) were administered via a laptop computer. When present, the Static-99, Stable 2000 and 2007, PPG, and SIR scores were transcribed directly from participants' criminal files. The SSPI was coded from the participants' criminal file information only if the participant had at least one victim under the age of 14.

The E-Prime 1.0 computer program (Schneider, Eschman, & Zuccolotto, 2002) was used to construct the SAC-IAT. All data analyses were performed with Version 17.0 of the Statistical Package for the Social Sciences (SPSS) for Windows.

### **Procedure**

**Interviews.** Correctly constructed IAT measures should contain stimulus words (e.g., man, naked) that unambiguously belong to their respective categories (e.g., *adult*, *sexy*; Nosek et al., 2007). To ensure proper stimuli selection, preliminary semi-structured interviews were conducted on 13 child molesters and 3 non-sex offenders to help generate words that best represent the category *sexy* and *not sexy*. Participants were informed that their answers would be confidential in order to create an environment conducive to discussion (see Appendix F). Appendix G presents the interview questions. Participants' responses were reviewed by the researchers and indicated that there was little consensus among sex offenders on the sexual attractiveness of words describing physical attributes (e.g., "fat", "skinny", "fragile") . As such, researcher focused on words that were synonyms of the word *sexy* or *not sexy* (e.g., "beautiful", "disgusting") or linked with sexual activities.

**Stimuli ratings.** The Rating Questionnaire included a series of words thought to represent the categories *sexy* and *not sexy* (selected by reviewing both the interview responses and the literature for other IAT measures adapted to assess sexual attractiveness) and pictures from the Not Real People Stimuli (Laws & Gress, 2004) in order to verify that these pictures could be consistently sorted as adults or children. The Rating Questionnaire included 33 pictures of children and adults' faces and 24 words (14 words thought to represent the category *sexy* and 10 words thought to represent *not sexy*).

All participants were informed that participation would be anonymous (Appendix H). The questionnaire was administered to a total of 45 sexual offenders currently receiving treatment at federal institutions in Ontario. Of the 45 participants, 14 reported only having convictions for sexual offences against a person 16 years of age or older (i.e., self-reported rapists), 10 participants reported only convictions for sexual offences against a person under 12 years of age (i.e., self-reported child molesters), 8 participants reported exclusively convictions for sexual offences against a person between the age of 12 and 15, and, lastly, 12 participants reported convictions for two or more of these types of sexual offences. One participant indicated no to all offence related items. Appendix I presents the rating scale with offenders' responses included. The words and pictures that were most consistently rated as belonging to a given category were included in the SAC-IAT.

**Pilot testing.** Once the SAC-IAT and the other elements (e.g., instructions, self-report questionnaires) were programmed into the E-Prime computer program, the procedure was pilot tested on 2 child molesters and 3 non-sex offenders (see Appendix K for the consent form). In addition to completing the testing procedure, participants were

also asked to give feedback on the procedure. Two of the five participants were initially confused with the combination tasks (e.g., both thought that the stimulus words or pictures had to represent both categories, for instance, that *smile* had to belong to both the category *sexy* and *adult*, as opposed to only the category *sexy*). As such, instruction screens explicitly explaining the sorting rules were created and presented before the presentation of each of the combination task in the study proper.

**Study proper.** Participant selection was as inclusive as possible. Offenders were included in the study if they had any recent sex offences (although if no other participant was available, participants with sex offenses prior to 10 years ago were included) or if they were non-sex offenders. Child molesters with victims under 12 as well as non-sex offenders were the preferred groups and hence were approached first to participate when they could be identified.

The current study was being conducted in the context of a larger study. This required the administration of three supplementary measures that were not analyzed in the present study. These measures include: the Attitude toward Child Sexual Abuse IAT (CSA-IAT; Kessous, 2009), the Sex with Children Scale (SWCH; Mann, Webster, Wakeling, & Marshall, 2007), and the Evaluation of Child Sexual Abuse Scale (ECSA; created for the purpose of the other study).

The researcher explained the study and those interested in participating signed a consent form (Appendix J). Participants completed all measures via a laptop computer in a room with the researcher. For security reasons, most prisons staff did not allow offenders to be alone in a room to complete the measures. In order to be consistent,

researchers were always in the room with the participant facing away from the computer screen so as to not be able to see how participants were responding.

The order in which the measures were presented was counterbalanced across participants to address potential order effects, with the exception of the SIPS and Viewing Time. These measures were only counterbalanced with each other and were presented only after the completion of the other measures to avoid the potential increase in salience due to their sexual explicitness. Appendix K indicates the order of presentation of the measures in greater detail.

The debriefing information was presented on the laptop; this allowed participants time to return to baseline should they have been aroused by the nude pictures or descriptions of sexual behaviour. The debriefing process lasted two minutes (Appendix L). Upon completion, participants were thanked for their participation and given an opportunity to ask questions.

File reviews were conducted for each participant after testing. Information on demographics, risk of sexual recidivism, criminal histories, and other variables of interest (e.g., PPG scores) were recorded for all participants. All risk assessment measures (i.e., Static-99, SIR, Stable 2000 and 2007) were directly transcribed from participants' criminal file as opposed to coded from file information. Only the SSPI was coded from file information.

### **Data Treatment**

**IAT scoring algorithm.** The E-Prime program recorded mistake counts and reaction times for each participant. This data were used to compute the *D* algorithm (Greenwald et al., 2003), which involves several steps explained in Appendix M. For this

study, the IAT effect ( $D$ ) represented the difference in mean reaction time when *child* and *sexy* were combined as compared to when *adult* and *sexy* were combined. Positive scores indicated participants responded more quickly when *child* and *sexy* were paired together than when *adult* and *sexy* were paired together. Thus, positive scores suggest that the participant had a stronger association between *child* and *sexy* than *adult* and *sexy*. A negative score indicated the reverse. Twenty-eight participants had at least one extreme trial (i.e., trials in which the participant took more than 10,000 milliseconds to sort the stimulus into the appropriate category). Extreme trials were dealt with using the procedures outline by Greenwald et al. (2003; see Appendix M). No participants had more than 10% of their trials with reaction times below 300 milliseconds, indicating no random or inattentive responding. As such, no corrective action was taken with regards to this issue.

**Viewing Time.** All reaction times for the viewing time measure were logarithmically transformed (i.e., log 10) to correct for positive skewness. The longest average viewing time for either female adult or male adult pictures (Tanner 5; age 16+) was subtracted from the longest average viewing time for either female child or male child pictures (Tanner 1, 2, or 3; representing ages 0-12). That is, whichever gender had the longest viewing time was used in computing the difference score. Thus, a score similar to the SAC-IAT score was created, with more positive scores representing greater interest in children relative to adults. Viewing Time that did not account for gender (that is, subtracted the average of the adult female and male pictures from the average of the longest viewed Tanner of male and female children) produced similar results. In this case, the average viewing time for adult pictures of both males and female was subtracted

from the longest average viewing time of both gender from either Tanner 1, 2, or 3 (male or female pictures).

**SIPS.** Each subscale and composite scores of the SIPS were calculated using the instructions outlined in the SIPS questionnaire (Laws, 1986). The Pedophilia Composite Score was the mean response of both the female and male pedophilia subscale and incest subscales of the SIPS. The Paraphilia Composite Score was the mean response of the voyeurism, exhibitionism, frottage, and tranvestism subscales of the SIPS. Lastly, the Violence Composite Score was the mean response of the rape, sadism, and masochism subscale of the SIPS. A differential score was also computed for the SIPS to parallel the other measures (e.g., Viewing Time, SAC-IAT). The highest rated subscale for adults (either adult heterosexuality or homosexuality scale) was subtracted from the highest rate subscales for children (either incest male, incest female, pedophilia male, or pedophilia female subscale), with positive scores suggesting greater sexual interest in children.

**PPG.** Raw PPG data for the Quinsey Child Sexual Violence Assessment was transcribed from participants' files. A total of 21 participants had PPG data available. A PPG score was then calculated by subtracting the highest adult non-violent category (either male or female) from the highest child non-coercion category (either girl or boy). Similar to the other measures of sexual interest (SAC-IAT, Viewing time, SIPS), a positive score indicated greater sexual interest in children.

## Results

### Data Screening

Prior to all analyses, variables were examined for violations of statistical assumptions. When a violation of normality occurred, data were usually corrected by

reducing outliers (defined as higher than  $\pm 2.32$ ) to the next highest or lowest score and, in extreme cases, through transformations. When any transformations were required, the analyses were conducted on both the untransformed and transformed variables. Only if different, is the transformed data presented. If both the untransformed and transformed data produced similar results, only untransformed data were presented.

### **Internal Consistency**

The internal consistency of the SAC-IAT was calculated using the method described in Nosek, Greenwald, and Banaji (2005). *D* coefficients were calculated for the combination of *child* and *sexy* (both block 3 and block 6) and for the combination of *adult* and *sexy* (both block 4 and block 7) and the two blocks were correlated. Appendix N presents detailed overview of the internal consistency analysis for an IAT measure. The internal consistency was  $r(78) = .49, p < .001$ , when including all participants, and  $r(57) = .58, p < .001$ , when only including child molesters with victims under the age of 12 and non-sex offenders.

The internal consistency alpha ( $\alpha$ ) was also computed between the two *D* coefficients to enable comparisons between the internal consistency of the SAC-IAT and the other measures presented in this study. The internal consistency of the SAC-IAT was moderate ( $\alpha = .64; N = 78$ ) when including all participants and when only including child molester with victims under 12 and non-sex offenders ( $\alpha = .73; N = 57$ ). Table 5 presents the internal consistency  $\alpha$  of the other measures used in the current study.

Table 5

## Internal Consistency of Self-Report Measures

Measure	Items	N	A
<b>BIDR</b>			
BIDR-SD	20	77	.69
BIDR-IM	20	77	.78
Clarke Vocabulary Scale	40	77	.84
<b>SIPS<sup>1</sup></b>			
Adult heterosexuality	10	74	.96
Adult homosexuality	10	74	.98
Female Pedophilia	10	74	.97
Female Incest	10	74	.97
Male Pedophilia	10	74	.95
Male Incest	10	74	.98
Voyeurism	10	74	.94
Exhibitionism	10	74	.93
Frottage	10	74	.93
Rape	10	74	.68
Masochism	10	74	.71
Transvestism	10	74	.76
<b>Viewing Time Responses</b>			
Tanner 1 Female	4	72	.97
Tanner 1 Male	4	72	.94
Tanner 2 Female	4	72	.92
Tanner 2 Male	4	72	.96
Tanner 3 Female	4	72	.89
Tanner 3 Male	4	72	.95
Tanner 4 Female	4	72	.75
Tanner 4 Male	4	72	.96
Tanner 5 Female	4	72	.88
Tanner 5 Male	4	72	.98

*Note.*  $\alpha$ : Alpha. Robinson, Shaver, and Wrightsman (1991) describe an alpha ( $\alpha$ ) of  $\geq .80$  as exemplary,  $.70 - .79$  as extensive,  $.60 - .69$  as moderate, and  $< .60$  as minimal. Alpha coefficient could not be computed for the Sadism subscale of the SIPS because of negative average covariance among items due to the low variance in item scoring (only three scores were over 1 for all items of this subscale). Selecting only offenders with victim over 18 did not change the covariance among items.

## Primary Hypotheses

**Descriptive statistics.** Analyses involving group differences were conducted on child molesters with victim under the age of 12 ( $n = 35$ ) and non-sex offenders ( $n = 21$ ). Of the child molester group, 30 had extrafamilial victims (unrelated victims under 12 years of age) and 5 had intrafamilial victims. Participants were classified as intrafamilial child molesters if the victim was related, such as a biological child, step-child (over two years), nephew, niece, first cousin, grandchild, or sibling under the age of 12 (Harris, Phenix, Hanson, & Thornton, 2003). Offenders with both related and unrelated victims were grouped as extrafamilial offenders. This type of grouping is commonly used in sex offender research (e.g., Barsetti, Earls, Lalumière, & Belanger, 1998; Hanson, 2002; Nunes et al., 2007).

Both groups had similar education levels. No non-sex offenders had below or equal to Grade 6 education, 71.4% ( $n = 15$ ) had between Grade 7 to Grade 13, and 23.8% ( $n = 5$ ) had some postsecondary education. For the child molester group, 5.7% ( $n = 2$ ) had below or equal to Grade 6 education, 60.0% ( $n = 21$ ) had between Grade 7 to Grade 13, and 25.7% ( $n = 9$ ) had some postsecondary education. Child molesters were significantly older than non-sex offenders,  $t(54) = 3.01, p = .004$ . In terms of sexual orientation, non-sex offenders uniformly reported a heterosexual orientation, whereas 60.0% ( $n = 21$ ) of child molesters reported a heterosexual orientation, 17.1% ( $n = 6$ ) reported a homosexual orientation, and 22.9% ( $n = 8$ ) reported a bisexual orientation.

The Static-99 scores and the Stable scores were transcribed either from the specialized sex offender assessment report or, if missing, from the treatment reports. In some cases, participants had two different Stable scores (e.g., before and after treatment

completion). In such cases, the most recent score was used in the analyses. The assessment dates ranged from January 2003 to June 2008 for the Static-99. The average Static-99 score of the child molesters was 4.55 ( $SD = 2.58$ ,  $n = 33$ ), which represents a moderate-high risk of sexual recidivism. Compared to other Canadian sexual offenders, this study's average participant (Static-99 score of 4) places in the 73.1 to 83.7<sup>th</sup> percentile. In other words, 73.1 to 83.7% of sex offenders in Canada scored lower or equal to the average participant in the current sample and 9.0% to 16.3% of sex offenders in Canada scored higher. With a follow-up of 5 years after release, sex offenders with the same Static-99 scores reoffended sexually at a rate of between 7.7% and 19.1% (Static-99 Clearinghouse, 2008). Considering that the average Static-99 score was 4.5 and that these estimates are based on a score of 4, we can expect the recidivism to be closer to the higher value of these approximations.

The average RRASOR score was 2.69 ( $SD = 1.69$ ,  $n = 32$ ). The Stable 2000 ( $n = 18$ ) and Stable 2007 ( $n = 12$ ) risk categories (i.e., low, moderate, high) were also combined to create more comprehensive picture of the offenders' dynamic risk. Child molesters were found to be either in moderate risk ( $n = 16$ , 55.2%) or high risk ( $n = 13$ , 44.8%) category on the Stable measures. The assessment dates for the Stable-2000 ranged from January 2003 to July 2007, whereas the assessment dates for the Stable-2007 ranged from January 2008 to June 2008.

The youngest victim (either from the index or prior offences) was, on average, 6.81 years old ( $SD = 2.69$ ,  $n = 34$ ). Nine (26.5%) child molesters offended against both boys and girls, whereas 11 (32.4%) only ever offended against boys, and 14 (41.2%) only ever offended against girls either in their index or, if any, prior sex offences. Lastly,

44.1% ( $n = 15$ ) had participated in a sex offender treatment program delivered by Correctional Service Canada after their index offence and 55.9% ( $n = 19$ ) had not completed a sex offender treatment program as of the testing date.

Table 6 presents the means, standard deviations, sample sizes, areas under the curve (AUCs) of the Receiving Operating Characteristic (ROC), and effect sizes for each main measure of sexual interest in children. AUCs are estimates of predictive accuracy and represent the chance that a randomly selected child molester would have a higher score on a measure of sexual interest (suggesting greater interest in children) than a randomly selected non-sex offender. AUCs of .56, .64, and .71 are considered, respectively, as small, medium, and large effect sizes (Rice & Harris, 2005). Figure 2 displays the ROC curve for each measure associated with the AUCs.

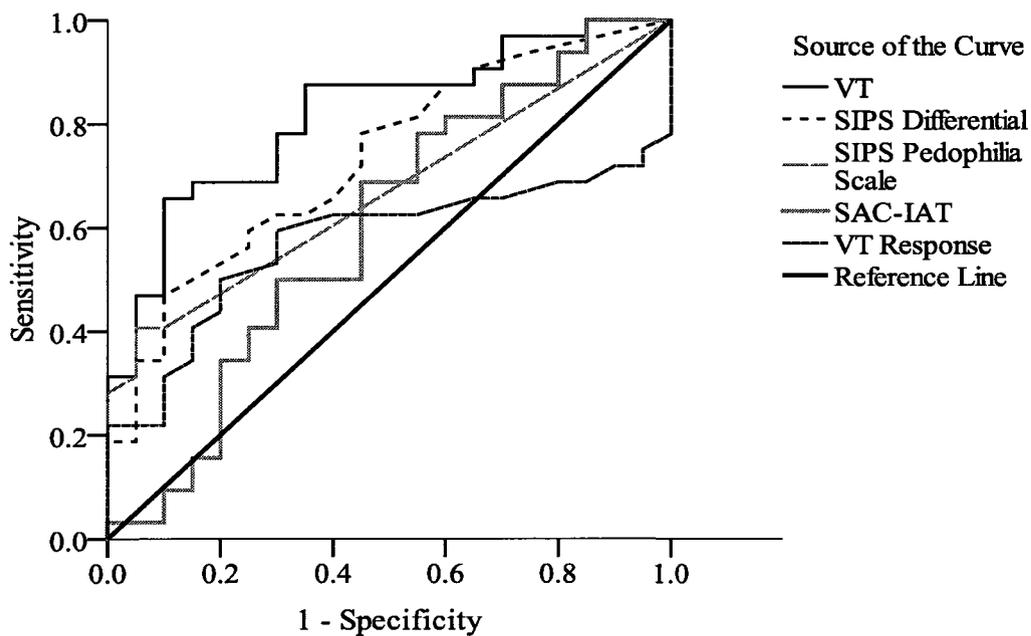


Figure 2. Receiver Operator Curves for measures of sexual interests.

*Table 6*  
Means, Standard Deviations, Cohen's *d*s, and AUCs for the Measures of Sexual Interest in Children

Measure	Child Molesters		Non-Sex Offenders		<i>d</i> (95% C.I.)	AUC (95% CI)
	Mean (SD)	<i>N</i>	Mean (SD)	<i>N</i>		
SAC-IAT	-0.20 (.37)	35	-0.38 (.47)	21	0.43 (-0.12 – 0.98)	.60 (.44 – .77)
VT <sup>a</sup>	0.11 (0.18)	31	-0.08 (0.14)	20	1.15 (0.55 – 1.75)*	.82 (.70 – .93)*
VT-AR	-0.92 (2.09)	31	-1.49 (1.11)	20	0.32 (-0.24 – 0.89)	.57 (.42 – .73)
SIPS Differential	-3.47 (2.17)	33	-5.03 (1.30)	22	0.83 (0.27 – 1.40)*	.73 (.59 – .87)*
SIPS Pedophilia	5.76 (4.47)	33	4.03 (0.13)	22	0.41 (-0.13 – 0.96)	.67 (.52 – .81)*

*Note.* SD: Standard deviation. C.I.: Confidence interval. SAC-IAT: Sexual Attraction to Children-Implicit Association Test. VT: Viewing Time. VT-AR: Viewing Time Attractiveness Ratings. SIPS: Sexual Interest Profiling System. Some group totals differ due to missing data. Child molesters were used as the reference group in the Cohen's *d* calculations.<sup>a</sup> Logarithmically transformed data presented for Viewing Time.\*  $p < .05$ .

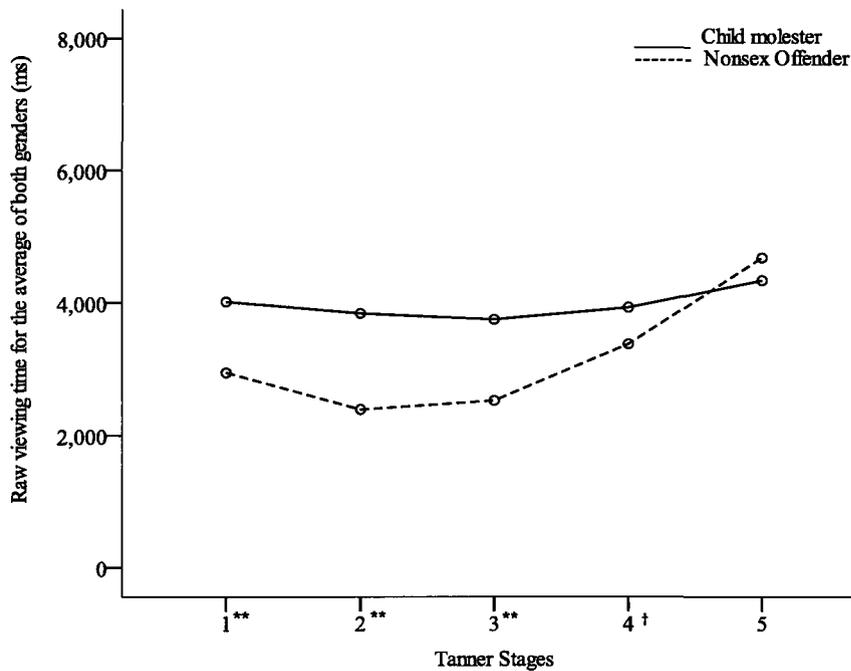
**Group differences on the SAC-IAT.** It was hypothesized that child molesters would demonstrate greater sexual interest in children as measured by the SAC-IAT than non-sex offenders. A two-way ANOVA was conducted to take into consideration the compatibility of the order of blocks presented. As aforementioned, the compatibility of the blocks (i.e., whether the first block you received, either child and sexy or adult and sexy, was compatible with your preference) has been found to influence *D* scores. Although group differences were in the expected direction (with child molesters showing greater sexual interest in children relative to non-sex offenders), the groups did not significantly differ in their SAC-IAT scores,  $F(1,54) = 2.43, p = .13$ . Correspondingly, the effect size was small, *Cohen's d* = 0.43 (95% C.I. = -0.12 – 0.98). The interaction between compatibility of the order of blocks and group membership (e.g., *adult* and *sexy* presented first versus *child* and *sexy* presented first) was not significant,  $F(1,54) = 0.48, p = .49$ . There was no relationship between the order of completions and the SAC-IAT scores,  $r(57) = -.19, p = .15$ .

**Cumulative meta-analysis.** As reported in the introduction, the meta-analysis of the previous studies that adapted the IAT measure to assess sexual interest in children (Banse et al., 2009; Brown et al., 2009; Gray et al., 2005; Mihailides et al. 2004; Nunes et al, 2007; Steffens et al., 2008) revealed a mean effect size of  $d = 0.73$  (95% C.I. = 0.54 – 0.93). As reported in Table 6, the effect size found in the current study was nonsignificant ( $d = 0.43$ , 95% C.I. = -0.55 – 0.98). The effect size was, however, not significantly different from the average effect size found in the previous studies ( $Q_{\Delta} = 1.03, df = 6, p > .05$ ). The new cumulative effect size, incorporating the current and past effect size, was  $d = 0.70$  (95% C.I. = 0.51 – 0.88). In addition, when the increase in variability ( $Q_{\Delta} = 1.03$ )

was added to the old variability ( $Q_{old} = 1.52$ ), the total variability was no more than would be expected by chance ( $Q_{new} = 2.55, p > .05$ ).

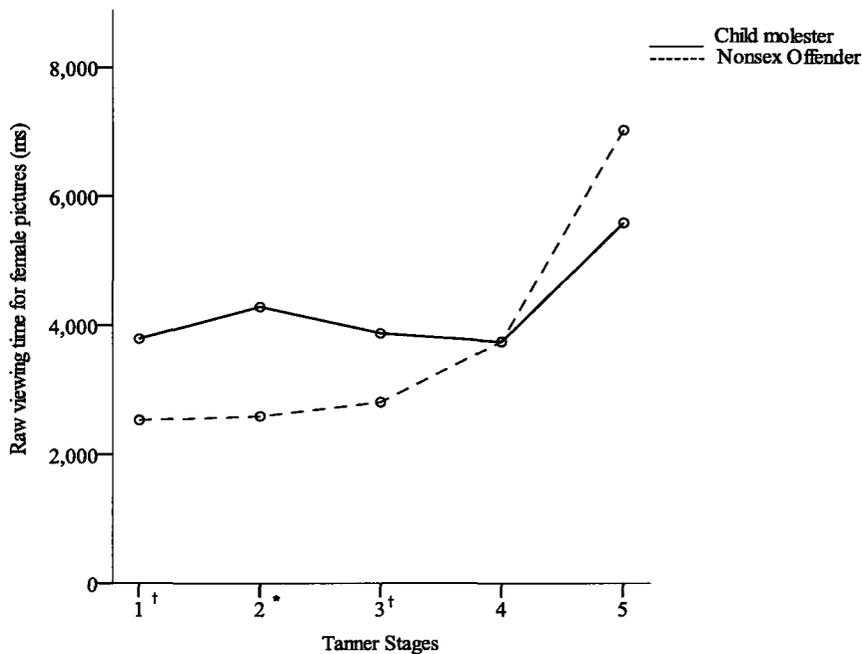
**Group differences on Viewing Time.** Viewing time was positively skewed and logarithmically transformed. Child molesters and non-sex offenders significantly differed on the Viewing Time differential scores,  $F(1,49) = 19.79, p < .001$ . This represented a large effect size, *Cohen's d* = 1.15 (95% C.I. = 0.55 - 1.75). Results presented in Table 6 were also based on the logarithmically transformed data.

Group differences in viewing time for each Tanner stage were examined using *t* tests. Figure 3 presents the raw viewing times as they are more easily interpretable compared to logarithmically transformed data, however, all *t*-tests were conducted on logarithmically transformed data because of severe skewness. Averaging gender, child molesters viewed pictures of Tanner 1, 2, and 3 significantly longer than non-sex offenders,  $t_{Tanner1(50)} = 2.38, p = .02, t_{Tanner2(50)} = 2.67, p = .01, t_{Tanner3(50)} = 2.86, p = .01$ . Tanner 4 approached significance,  $t_{Tanner4(50)} = 1.72, p = .09$  (see Figure 3).



*Figure 3.* Average viewing time for each Tanner scale. Log-transformed viewing time was used in t-tests. Analyses of raw data produced similar results, with the exception that the groups were not approaching a significant difference on Tanner 4 in the untransformed data. Tanner 1: ages 0-3, Tanner 2: ages 4-7, Tanner 3: ages 8-12, Tanner 4: ages 13-15, Tanner 5 ages 16 and older. †  $p < .10$ , \*\*  $p < .01$ .

As can be seen in Figure 4, child molesters viewed Tanner 2 females significantly longer than non-sex offenders,  $t(50) = 2.51, p = .02$ . Group differences between female pictures of Tanner 1 and Tanner 3 were approaching significance,  $t_{\text{Tanner1}}(50) = 1.91, p = .05$  and  $t_{\text{Tanner3}}(50) = 1.93, p = .08$ , respectively.



*Figure 4.* Viewing time for female pictures. Logged transformed reaction time was used in *t*-tests. Tanner 1: ages 0-3, Tanner 2: ages 4-7, Tanner 3: ages 8-12, Tanner 4: ages 13-15, Tanner 5 ages 16 and older. †  $p < .10$ , \*  $p < .05$ .

Lastly, child molesters consistently viewed pictures of males significantly longer than non-sex offenders, irrespective of Tanner stages,  $t_{\text{Tanner1}}(50) = 2.38, p = .02$ ,  $t_{\text{Tanner2}}(50) = 2.46, p = .004$ ,  $t_{\text{Tanner3}}(50) = 3.34, p = .03$ ,  $t_{\text{Tanner4}}(50) = 2.61, p = .03$ ,  $t_{\text{Tanner5}}(50) = 2.16, p = .03$ , respectively (see Figure 5).

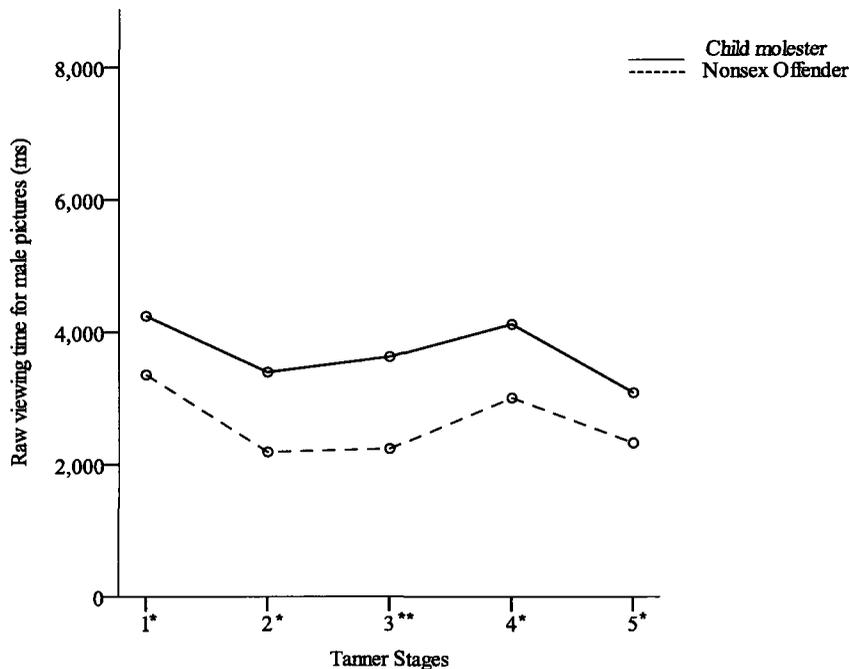


Figure 5. Viewing time for male pictures. Logged transformed reaction time was used in t-test. \*  $p < .05$ , \*\*  $p < .01$ .

#### Group differences on Viewing Time Attractiveness Rating. Sexual

attractiveness ratings of the pictures used in the viewing time procedure were summed for each Tanner stage and gender. Average ratings were also computed by averaging the sexual attractiveness rating of each Tanner stage (i.e., including both female and male pictures). With the exception of Tanner 1, child molesters consistently rated pictures (whether male or female) as more sexually attractive than non-sex offenders,  $t_{\text{Tanner2}(50)} = 3.02, p = .048$ ,  $t_{\text{Tanner3}(50)} = 2.81, p = .007$ ,  $t_{\text{Tanner4}(50)} = 3.53, p = .001$ ,  $t_{\text{Tanner5}(50)} = 2.43, p = .02$  (see Figure 6).

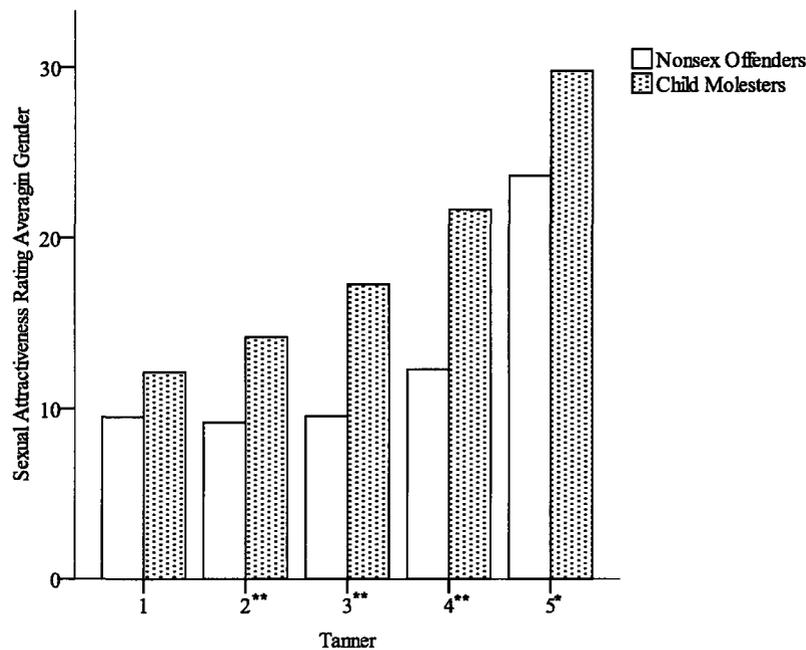


Figure 6. Average sexual attractiveness ratings. Untransformed data presented.  
 \*  $p < .05$ , \*\*  $p < .01$ .

Child molesters and non-sex offenders did not significantly differ in their attractiveness ratings for female pictures for all Tanner stages (see Figure 7). In contrast, groups differed in their responses for male pictures. With the exception of Tanner 1, child molesters consistently rated male pictures as more sexually attractive than non-sex offenders,  $t_{\text{Tanner}2}(50) = 2.07, p = .04$ ,  $t_{\text{Tanner}3}(50) = 2.85, p = .006$ ,  $t_{\text{Tanner}4}(50) = 3.69, p = .001$ ,  $t_{\text{Tanner}5}(50) = 3.33, p = .002$ , respectively (see Figure 8).

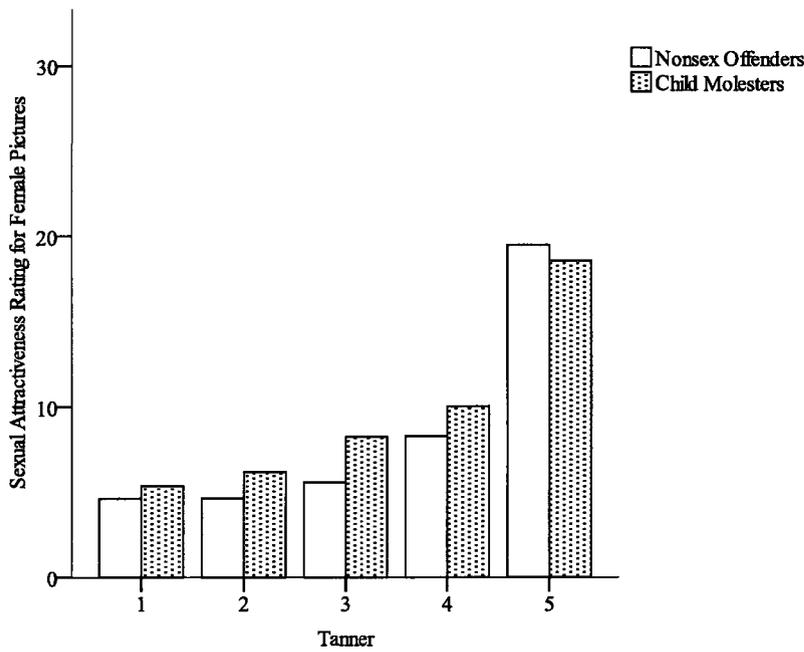


Figure 7. Sexual attractiveness ratings for female pictures. Untransformed data presented.

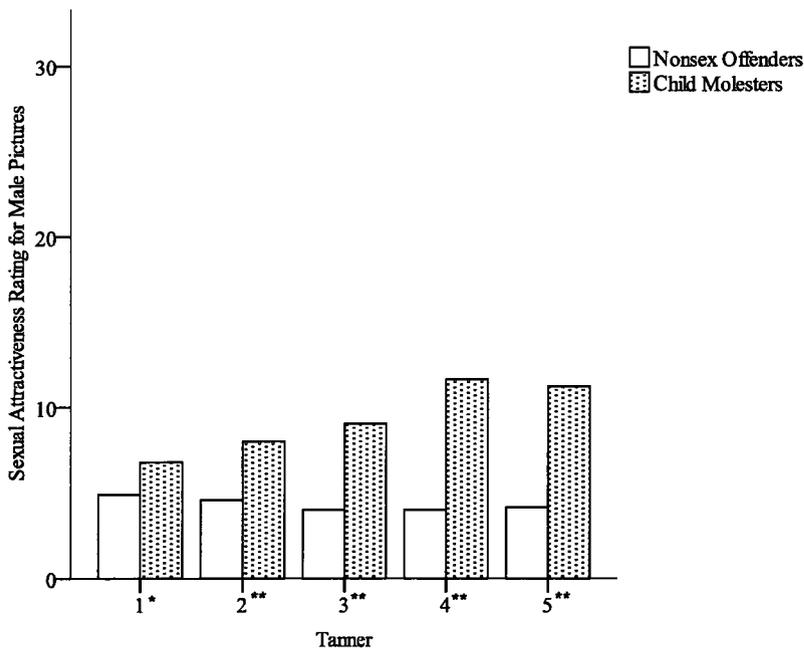


Figure 8. Sexual attractiveness ratings for male pictures. Untransformed data presented.  
 \*  $p < .05$ , \*\*  $p < .01$ .

**Viewing Time Attractiveness Ratings Differential.** Similar to the other measures of sexual interest, a differential score was computed for the attractiveness ratings of the pictures used in the Viewing Time procedure. The average attractiveness rating for each Tanner and each gender was first computed. Afterwards, a differential score was computed by subtracting the highest rated Tanner representing either adult males or female (Tanner 5, 16+) from the highest rated Tanner representing either boys or girls (Tanner 1 to 3). No outliers were found, however, the Welch correction was used because homogeneity of variance was violated. Groups did not significantly differ in their response on the Viewing Time response differential,  $F(1,47.55) = 1.62, p = .21$ .

**Group differences on the SIPS.** The SIPS differential score significantly differed between groups,  $F(1,51.92) = 10.24, p = .002$ . Table 7 presents group differences for the SIPS subscales and Composite Scores. The Pedophilia Composite scale had two outliers in the non-sex offenders group. The outliers were found and reduced to one unit above the next lowest score. Even after this change, outliers were present. As such, an inverse Logarithmic transformation was done for severe positive skewness. Despite these transformations, normality and homogeneity of variance was violated (although improved from untransformed data), and so the Welch correction was used. Using the transformed data (outlier reduced and with logarithmic transformation), the total pedophilia score was able to differentiate between child molesters and non-sex offenders,  $F(1,32.26) = 12.20, p < .001$ . The untransformed data are presented in Table 7.

Table 7

Means, Standard Deviations, and Cohen's *d*s for SIPS Subscales and Composite Scores

Subscales	Child Molester ( <i>N</i> = 34)		Non-Sex Offender ( <i>N</i> = 21)		<i>d</i> (95 % <i>CI</i> )
	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	
Adult Heterosexuality	46.26 (21.61)	60.19 (13.27)	-0.74 (-1.30 - -0.18)*		
Adult Homosexuality	24.53 (21.25)	10.67 (3.06)	0.82 (1.39 - 0.26)*		
Pedophilia Composite Score	14.88 (11.35)	10.08 (0.33)	0.54 (-1.09 - 0.02)†		
Female Pedophilia	15.38 (13.48)	10.05 (0.22)	0.50 (-1.05 - 0.05)†		
Female Incest	13.82 (11.34)	10.00 (0.00)	0.43 (-0.98 - 0.12)		
Male Pedophilia	16.18 (14.22)	10.29 (1.31)	0.52 (-1.08 - 0.03)†		
Male Incest	14.12 (11.15)	10.00 (0.00)	0.47 (-1.02 - 0.08)†		
Paraphilia Composite Score	14.04 (8.49)	10.63 (1.16)	0.51 (-1.06 - 0.05)†		
Voyeurism	15.21 (11.34)	10.05 (0.22)	0.58 (-1.13 - -0.02)†		
Exhibitionism	14.88 (10.97)	10.10 (0.30)	0.55 (-1.11 - 0.002)†		
Frottage	13.71 (10.52)	10.19 (0.87)	0.42 (-0.97 - 0.13)		
Transvestism	12.35 (5.09)	12.19 (4.70)	0.03 (-0.58 - 0.51)*		
Child Violence Composite Score	11.59 (2.57)	11.41 (1.78)	0.08 (-0.62 - 0.47)		
Rape	10.91 (3.16)	10.00 (0.00)	0.36 (-0.91 - 0.18)		
Sadism	10.21 (0.77)	10.00 (0.00)	0.35 (-0.89 - 0.20)		
Masochism	13.65 (6.29)	14.24 (5.33)	-0.10 (-0.45 - 0.64)		

Note. Child molester used as the reference group for the Cohen's *d*. Untransformed data are presented. The rape, sadism, and masochism subscales solely involve scenarios with children. Pedophilia Composite Score is the combined mean of the female and male pedophilia and incest subscales. Paraphilia Composite Score is the combined mean of the voyeurism, exhibitionism, frottage, and transvestism subscales. Child Violence Composite Score is the combined mean of the rape, sadism, and masochism subscales.

† *p* < .10, \* *p* < .05.

**Alternative explanations.** Table 8 presents the correlations between possible extraneous variables (e.g., age of participants, BIDR scores, vocabulary scores, and education level) and the measures of sexual interest. An ANCOVA conducted only if a significant correlation was found.

*Table 8*

Correlation between Possible Confounding Variables and the Measures of Sexual Interest in Children

Measures	SAC-IAT	SIPS Pedophilia	SIPS Differential	VT	VT-AR
Age <sup>1</sup>	.27* (57)	.09 (56)	.46** (56)	.14 (53)	.13 (53)
SIR	.20 (54)	.08 (53)	.38** (53)	.21 (51)	.21 (51)
BIDR Total	-.12 (57)	-.17 (56)	.03 (56)	-.31* (53)	.05 (53)
IM	-.01 (57)	-.06 (56)	.18 (56)	-.17 (53)	.11 (53)
SD	-.24 <sup>†</sup> (57)	-.26 <sup>†</sup> (56)	-.17 (56)	-.41** (53)	-.05 (53)
Education	-.15 (53)	-.05 (52)	-.36* (52)	-.004 (49)	-.21 (49)
Vocabulary	-.09 (57)	-.03 (56)	-.29* (56)	-.03 (53)	-.11 (53)

*Note.* SIR: Statistical Information on Recidivism. BIDR: Balanced Inventory of Desirable Responding. IM: Impression Management subscale. SD: Self-deception subscale. Education: Highest grade completed. Vocabulary: Clarke Vocabulary Scale. SAC-IAT: Sexual Attraction to Children-Implicit Association Test. VT: Viewing Time. VT-AR: Viewing Time Attractiveness Ratings. SIPS: Sexual Interest Profiling System. <sup>1</sup>Age at testing. Untransformed data presented. Concurrent with literature on social desirability and risk, there was a relationship between risk and social desirability. Offenders with low scores on Impression Management subscale (displaying less impression management) tended to have lower score on the SIR (representing a higher risk of recidivism),  $r(54) = .27, p = .046$ . <sup>†</sup> $p < .10$ , \*  $p < .05$ , \*\*  $p < .01$ .

Potential covariates specific to sex offenders (e.g., risk assessment instruments) were also examined through correlations. Table 9 presents these correlations.

Table 9

Relationship between Measures of Sexual Interest and Variables Specific to Child Molesters

Measures	RRASOR	Static-99	Male Victim	Age of Victim	Time Since Last Offence	Treatment Participation
SAC-IAT	-.02 (32)	-.02 (33)	-.11 (32)	-.27 (34)	.06 (34)	.39* (35)
SIPS Pedophilia	.18 (31)	.23 (32)	.23 (31)	-.14 (33)	.03 (33)	.15 (34)
SIPS Differential	-.02 (31)	.05 (32)	.18 (31)	.08 (33)	.12 (33)	.13 (34)
VT	.23 (29)	.33 <sup>†</sup> (30)	.05 (29)	-.08 (31)	.13 (31)	.12 (32)
VT-AR	.11 (29)	.21 (30)	.08 (29)	.11 (31)	-.28 (31)	.13 (32)

*Note.* SAC-IAT: Sexual Attraction to Children-Implicit Association Test. VT: Viewing Time. VT-AR: Viewing Time Attractiveness Ratings. SIPS: Sexual Interest Profiling System. The youngest age was when multiple victims existed. <sup>†</sup> $p < .10$ , \*  $p < .05$ .

#### ***Possible Explanations for Lack of Group Differences in the SAC-IAT.***

*Age.* Further analyses were conducted to examine whether age was a confounding covariate that could possibly explain the lack of group differences on the SAC-IAT. As can be seen from Table 8, age was correlated with SAC-IAT scores, with increases in age corresponding to greater sexual interest in children as measure by the SAC-IAT. An ANCOVA was therefore conducted with age as the covariate; the SAC-IAT was still unable to significantly differentiate between groups,  $F(2,56) = 2.33$ ,  $p = .11$ .

*Incest offenders.* As aforementioned, incest offenders have been found to demonstrate less sexual interest in children than extrafamilial offenders. If incest offenders have little to no sexual interest in children, then including them in the analysis may reduce the average SAC-IAT score of the child molester group. As such, incest

offenders were removed from the analysis and an ANOVA was conducted solely on extrafamilial child molesters. No significant differences on the SAC-IAT were found between non-sex offenders and extrafamilial child molesters,  $F(1,49) = 1.38, p = .25$ .

*Treatment participation.* The SAC-IAT was correlated to treatment participation (based on file review). This finding could be due to higher risk child molesters being given priority placement into treatment. Consistent with this explanation, after controlling for risk of recidivism (i.e., Static-99 scores), the observed relationship was no longer present,  $r_{\text{partial}}(30) = -.11, p = .55$ .

*Number of mistakes.* It is possible that number of mistakes on the SAC-IAT (i.e., not correctly sorting the stimuli into their respective categories) is a proxy for understanding and attentiveness. There were no outliers for non-sex offenders, that is, no non-sex offenders had significantly higher mistake counts on the IAT than other non-sex offenders. In contrast, two child molester had greater average mistake counts ( $M = 13.5$  and  $M = 17.5$ , respectively) than the child molester group ( $M = 3.33$ , with 32 of 34 participants averaging less than 8 mistakes). Perhaps due to the low number of offenders with extreme mistake counts, removing these offenders did not make a difference in the SAC-IAT ability to discriminate between group,  $F(1,53) = 2.52, p = .12$ .

*Gender preference and the IAT stimuli.* Despite the reduced ambiguity of pictures compared to words, results indicate no significant difference on this IAT measure between child molesters and non-sex offenders. One explanation of the lack of difference is the gender-neutral quality of our IAT measure. That is, in the SAC-IAT, pictures of both genders were used to represent the category *child* and *adult*. As such, the computation of the SAC-IAT score was based on the mean reaction time of both genders.

It is possible that the gender-neutrality clouded the SAC-IAT effect. In other words, since the majority of participants reported a heterosexual orientation, it is possible that pictures of men in the *adult* and *sexy* pairing would have actually increase errors and reaction time, hence potentially increasing their IAT scores and measurement error.

*Sexual orientation.* Homosexual and bisexual offenders were removed from the analyses so that the groups would be equally affected by the gender saliency of the measure. This analysis reduced the child molester group to 21 and, since all non-sex offenders reported a heterosexual orientation, the non-sex offender group continued to have 21 participants. The mean differences between heterosexual child molesters and heterosexual non-sex offenders were slightly larger,  $F(41) = 2.78, p = .10$ , however, not significantly different. Importantly, the average SAC-IAT score of heterosexual child molesters was not significantly different than the average SAC-IAT score of the original child molesters group (i.e., including heterosexual, homosexual, and bisexual participants),  $t(54) = 0.38, p = .71$ .

To further investigate the possible effect of gender saliency on SAC-IAT scores, two SAC-IAT scores were computed, one by selecting offenders' responses for only female pictures and one by selecting offenders' responses with only male pictures. These gender specific IAT measures did not distinguish between groups (SAC-IAT Female,  $F[1,55] = 2.29, p = .14$ ; SAC-IAT Male,  $F[1,55] = 2.67, p = .11$ ) and had comparable correlation coefficients with other measures of sexual interest compared to the original SAC-IAT (see Table 10). The SAC-IAT Male transformed data is presented, in which one outlier was reduced to the next highest score.

Table 10  
 Relationship between SAC-IAT Male and SAC-IAT Female and Other Measures of Sexual Interest

Measures	SAC-IAT	SAC-IAT		PPG	SIPS Differential	SIPS Pedophilia	VT	VT-AR	SSPI
		IAT Female	IAT Male						
SAC-IAT Male	-	-	-	.16 (21)	.34* (56)	.16 (56)	.38** (53)	.34* (53)	.11 (35)
SAC-IAT Female	-	-	.57** (57)	.37† (21)	.33* (56)	-.05 (56)	.31* (53)	.26† (53)	-.03 (35)
SAC-IAT	-	.83** (57)	.69** (57)	.19 (21)	.37* (56)	-.12 (56)	.41** (53)	.30* (53)	-.09 (35)

Note. SIPS: Sexual Interest Profiling System. SAC-IAT: Sexual Attraction to Children-Implicit Association Test. VT: Viewing Time. VT-AR: Viewing Time Attractiveness Ratings. SAC-IAT: Male correlations are based on untransformed data; transformed data produced similar, usually slightly larger, correlations. †  $p < .10$ , \*  $p < .05$ , \*\*  $p < .01$ .

**SIPS.** The SIPS differential had the most potential covariates of the measures of sexual interest. In other words, there were several possible alternative explanations as to why groups differed on the SIPS differential. As can be seen from Table 8, the SIPS differential score was correlated to age, education, vocabulary, and SIR scores. An ANCOVA was conducted using age as a covariate and the same pattern was found. The SIPS differential scores continued to differentiate between groups,  $F(2,52) = 9.06, p < .001$ . An ANCOVA was also conducted separately for vocabulary, education, and SIR scores. The SIPS differential continued to distinguish between groups,  $F(2,52) = 8.03, p < .001, F(2,48) = 4.83, p = .01, F(2,49) = 6.67, p = .003$ , respectively.

### **Secondary Hypotheses**

**Convergent validity.** The convergent validity of the measures is presented Table 11. Unlike the previous analyses, sex offenders with victims between 13 and 17, sex offenders with victim equal to or over 18, and non-sex offenders reporting perpetrating a sexual assault were also included in the subsequent analyses. This was done in order to (1) increase the sample size and (2) allow for greater range in responses, and because group differences were no longer of interest. Interestingly, although SAC-IAT could not differentiate between groups, it was correlated with other measures that could distinguish groups (i.e., Viewing Time and the SIPS).

Table 11  
 Convergent validity of the Measures Sexual Interest in Children

Measures	SAC-IAT	SIPS Pedophilia	SIPS Differential	VT	VT-AR	SSPI	PPG
SAC-IAT	-	-	-	-	-	-	-
SIPS Pedophilia	-.07 (73)	-	-	-	-	-	-
SIPS Differential	.35** (73)	-	-	-	-	-	-
VT	.33** (71)	.28* (71)	-	-	-	-	-
VT-AR	.22 <sup>†</sup> (71)	.35** (71)	.55** (71)	-	-	-	-
SSPI	-.05 (41)	.30 <sup>†</sup> (40)	.10 (40)	.30 <sup>†</sup> (39)	-	-	-
PPG	.19 (21)	.22 (20)	.02 (20)	.01 (20)	.06 (20)	.30 (21)	-

Note. SAC-IAT: Sexual Attraction to Children-Implicit Association Test. SIPS: Sexual Interest Profiling System. VT: Viewing Time. VT-AR: Viewing Time Attractiveness Ratings. SSPI: Screening Scale for Pedophilic Interests. PPG: Penile Plethysmography. Removing the low responders from PPG correlations reduced the sample size to 12 and increased some correlation coefficients (SAC-IAT,  $r[12] = .28$ ; VT,  $r[12] = .06$ ; VT-AR,  $r[12] = -.06$ ; SIPS Differential,  $r[12] = -.05$ ; SIPS Pedophilia,  $r[12] = .30$ ), and slightly reduced the SSPI ( $r[12] = .30$ ). Controlling for social desirability produced similar correlation between the SAC-IAT and the SIPS Differential ( $r_{partial}[67] = .32, p = .008$ ), and slightly larger correlations with VT-AR ( $r_{partial}[67] = .24, p = .046$ ) and the Pedophilia Composite Score of the SIPS ( $r_{partial}[67] = -.10, p = .4$ ). <sup>†</sup>  $p < .10$ , \* $p < .05$ , \*\* $p < .01$ .

Since the SAC-IAT was computed by averaging both genders, Viewing Time, Viewing Time Attractiveness Ratings, SIPS Differential, and PPG were also computed by averaging gender. This was done in order to examine whether greater correlation would be found between measures if the scores were computed similarly. The average Viewing Time for both women and men was subtracted from the mean viewing time for both boys and girls of Tanner equal to and below 3 (i.e., 0 to 12 years of age). The same computation was used for Viewing Time Attractiveness Ratings. Similarly, the PPG Z scores for both the women and men categories were subtracted from the mean viewing time for both boys and girls categories. Lastly, the SIPS Differential was computed by subtracting the average of the homosexual and heterosexual subscales from the average of the Female Incest, Male Incest, Female Pedophilia, and Male Pedophilia subscales. Table 12 presents the correlations between the SAC-IAT and the sexual interest measures that were computed averaging gender.

*Table 12*

Relationship between the SAC-IAT and Measures of Sexual Interest Computed by Averaging the Genders

Measures	SAC-IAT
SAC-IAT	–
VT	.33** (71)
VT-AR	.15 (71)
SIPS Differential	.28* (73)
PPG	-.10 (19)

*Note.* SAC-IAT: Sexual Attraction to Children-Implicit Association Test. VT: Viewing Time. VT-AR: Viewing Time Attractiveness Ratings. SIPS: Sexual Interest Profiling System. \* $p < .05$ , \*\* $p < .01$ .

**Measures of sexual interest in children and risk of recidivism.** Table 13

displays the relationship between measures of sexual interest and risk of recidivism. Only child molesters with victims under 12 were included in the correlations between the Stable deviance item and measure of sexual interest because deviant interest for rapists referred to interest in violent sex with an adult, rather than deviant interest in children. Stable 2000 and 2007 deviance items were combined to have a larger sample size (the coding rule for the deviance item for the Stable 2000 and 2007 are similar and mostly can be based on PPG results, victim type, and other relevant information; Harris, 2008).

As can be seen by Table 13, the SAC-IAT was not correlated to the Static-99. It is possible that including sexual offenders with victims over 16 years of age may dilute the correlations between SAC-IAT and Static-99 scores because these offenders may score high on the Static-99 due to antisociality rather than pedophilia. As such, some sex offenders with victims over the age of 16 may have high scores on the static-99 and low scores on the SAC-IAT. The correlation was therefore conducted only including sex offenders with victims under the age of 12. Including only child molesters with victims under 12, however, resulted in similar relationship,  $r(35) = -.10$ , as did only including those with victims under the age of 14,  $r(38) = -.08$ .

*Table 13*  
 Relationship between Measures of Sexual Interest in Children and Risk of Sexual Recidivism

Measures	SAC-IAT	SIPS Pedophilia	SIPS Differential	Viewing Time	VT-AR	PPG	SSPI
Static-99	-.13 (47)	.18 (46)	.02 (46)	.23 (45)	.10 (45)	.11 (24)	.58** (38)
RRASOR	-.13 (46)	.18 (45)	.02 (45)	.21 (44)	.08 (44)	.25 (24)	.63** (37)
Stable <sup>1</sup>	-.11 (42)	.19 (40)	.05 (40)	.23 (40)	.13 (40)	.06 (24)	.27 (32)
Stable-2000	-.12 (26)	.11 (24)	.16 (24)	.03 (25)	.21 (25)	.32 (17)	.46* (21)
Stable-2007	-.17 (20)	.20 (20)	.14 (20)	.38 (19)	.01 (19)	.34 (9)	.37 (13)
Deviance Item	.40* (29)	-.01 (28)	-.13 (28)	-.05 (27)	.03 (27)	.53* (21)	.13 (29)

*Note.* SAC-IAT: Sexual Attraction to Children-Implicit Association Test. SIPS: Sexual Interest Profiling System. VT-AR: Viewing Time Attractiveness Ratings. SSPI: Screening Scale for Pedophilic Interests. PPG: Penile Plethysmography. <sup>1</sup>Stable represents the risk category (low, moderate, or high) for either the Stable-2000 or Stable-2007. †  $p < .10$ , \*  $p < .05$ , \*\*  $p < .01$ .

## Discussion

The main purposes of the current study was to assess whether child molesters and non-sex offenders responded differently on the SAC-IAT, and to assess the relationship of the SAC-IAT to other more commonly used measures of sexual interest. In addition, the relationship of the SAC-IAT with risk of sexual recidivism was also examined. This study also examined the ability of the Viewing Time and the SIPS to distinguish between groups, as well as their relationship with other measures of sexual interest and risk of sexual recidivism.

Four primary hypotheses were proposed regarding the SAC-IAT. First, it was hypothesized that the SAC-IAT would differentiate child molesters from non-sex offenders. Contrary to the hypothesis, results demonstrate that child molesters did not hold greater association between children and sexual attractiveness than non-sex offenders as assessed by the SAC-IAT. The difference, however, was in the expected direction, with child molesters showing a nonsignificantly greater association between children and sexual attractiveness than non-sex offenders. It was also hypothesized that the SAC-IAT would be associated with other measures of sexual interest. The SAC-IAT was correlated with both the Viewing Time (indirect measure) and the SIPS differential (self-report measure). The relationship with SAC-IAT and PPG was not statistically significant; however, the SAC-IAT was positively and significantly correlated with the deviance item of the Stable measures (this analysis had a larger sample size).

The internal consistency of the SAC-IAT was  $\alpha = .64$  ( $\alpha = .70$ , when only child molesters with victims under 12 and non-sex offender without self-reported sex offences were included in the analysis), which is consistent with past research on other IAT

measures (.70 to .90; Hofmann et al., 2005; Nosek, Greenwald, & Banaji, 2006). Lastly, it was also expected that, among child molesters, greater scores on the SAC-IAT would be associated with greater risk for sexual recidivism, as measured by validated risk assessment instruments. Contrary to hypothesis and past research (e.g., Nunes et al., 2007), the SAC-IAT was not correlated with the Static-99. In addition, the SAC-IAT was not correlated with the RRASOR, which was presumed to assess sexual deviancy at a greater extent than the Static-99. Although Banse et al. (2009) found that the Girl Women IAT was significantly and positively correlated with the SSPI, the current study did not find any significant correlation.

Several other measures of sexual interest were also included in the current study to allow for convergent validity analyses. First, the ability of these measures to distinguish between groups was examined. It was hypothesized that the Viewing Time differential would be able to distinguish between groups. As expected, child molesters viewed pictures of children relative to adult significantly longer than non-sex offenders (i.e., Viewing Time differential). This difference was found whether the differential was computed by averaging the viewing time of both genders or by using the longest viewed gender. Furthermore, child molesters viewed pictures of males significantly longer than non-sex offenders for each Tanner stage. Consistent with findings based on PPG-assessed sexual interest, differences in viewing time were less pronounced in the female pictures (Seto, 2008). Child molesters viewed Tanner 2 (4-7 years of age) females significantly longer than non-sex offenders and group differences were only approaching significance for female pictures of Tanner 1 (0-3 years of age) and Tanner 3 (8-12 years of age), with child molesters viewing pictures longer than non-sex offenders. Group differences were

not found for Tanner 4 (13-15 years of age) and Tanner 5 females (16 years of age and older).

The Viewing Time differential was also found to have the highest accuracy in distinguishing between child molesters and non-sex offenders and was positively correlated with the SAC-IAT, the SIPS Differential, and the SIPS Pedophilia score. Although Letourneau (2002) found that viewing time was correlated with PPG, the current study did not find the Viewing Time Differential or the mean Viewing Time of each Tanner (computed using either both genders or the longest viewed gender) were correlated to PPG. This finding, however, could be due to a low number of child molesters with PPG data, and a large number of child molesters (nine of 21) that were classified as low responders (i.e., less than three millimeter changes). Viewing Time differential was also not correlated to the deviance item of the Stable 2000. Whereas Banse et al. (2009) found significant positive correlation between their viewing time measure and the SSPI; the current study found that the relationship between the Viewing Time differential and the SSPI was only approaching significance.

Child molesters and non-sex offenders did not differ in their attractiveness rating for female pictures, regardless of Tanners; however, child molesters significantly rated male pictures as more sexually attractive than non-sex offenders. The Viewing Time Responses Differential was unable to distinguish between child molesters and non-sex offenders; however, it was correlated to the Viewing Time, the SIPS Pedophilia Composite Score, and the SIPS Differential.

Furthermore, the present study found that the SIPS Differential was able to distinguish between groups, with child molesters showing a less pronounced preference

for adults over children compared to the non-sex offenders. The SIPS Differential had moderate accuracy in distinguishing child molesters from non-sex offenders and was also found to be correlated to Viewing Time and the SAC-IAT. The groups also differed on the SIPS Pedophilia Composite score. Although group differences were approaching statistical significance when using untransformed data, with child molesters rating sexual scenarios involving children as significantly more attractive than non-sex offenders. The Adult Homosexuality, Adult Heterosexuality, and the Transvestism subscales of the SIPS also significantly distinguished between groups. In other words, child molesters were more likely than non-sex offenders to rate scenarios involving an adult male and transvestism as more sexually attractive than non-sex offenders. This is consistent with higher self-reported homosexual and bisexual orientations in the child molesters than the non-sex offenders. In contrast, non-sex offenders were more likely to rate sexual scenarios involving an adult female as more sexual attractive than child molesters.

The SSPI had moderate, but non-significant, correlation coefficients to Viewing Time and the SIPS Pedophilia scale. Consistent with past research, the SSPI was also correlated with PPG (Seto et al., 2004; Seto and Lalumière, 2001). Despite a moderate correlation coefficient, this correlation was not statistically significant, perhaps due to a small sample size. Lastly, none of the measures of sexual interest, with the exception of the SSPI, were correlated to measures of sexual recidivism. The Viewing Time differential had a larger correlation with the Static-99, RRASOR, and Stable measures than did the SAC-IAT, Viewing Time Attractiveness Ratings, and SIPS; however, the correlation coefficients were small and non-significant.

### Alternate Explanations

One potential explanation for the lack of significant difference between groups on the SAC-IAT is the gender-neutral quality of the measure. The use of picture stimuli has been associated with faster reaction times compared to word stimuli (Mitchell, Nosek, & Banaji, 2003; Nosek et al., 2002). Such stimuli are also expected to produce similar if not better discrimination because they allow for a substantial reduction in classification ambiguity. Utilizing pictures to represent the *child* and *adult* categories, however, may have had a gender salient effect. Namely, in the SAC-IAT, pictures of both genders were used to represent the *adult* and *child* categories and gender was salient when assigning pictures into the respective categories. In contrast, the gender of the child and adult is not predetermined in IAT measures computed using solely words to represent the *adult* or *child* categories; the participant can think of their preferred gender when sorting words representing the *adult* and *child* categories. In the SAC-IAT, the computation of the *D* score was based on the mean reaction time of both genders. It is possible that this gender-neutrality clouded the SAC-IAT effect. In the current study, the majority of participants reported a heterosexual orientation. It is therefore possible that pictures of men in the *adult* and *sexy* pairing would have actually increased errors and reaction time, thereby increasing measurement error in the SAC-IAT. In fact, when computing the other measures of sexual interest using the average of the two genders, correlations between the measures were reduced and, although the viewing time continued to distinguish between groups, the SIPS Differential was no longer able to distinguish between groups. These findings suggest that better discrimination between groups may occur when genders are separated.

Removing homosexual and bisexual offenders (so that the groups would be equally affected by the gender saliency of the measure), however, did not make a difference; the SAC-IAT still did not distinguish between the groups. Although the mean differences between heterosexual child molesters and heterosexual non-sex offenders were slightly larger, the difference in the child molester group (whether including or excluding non-heterosexual offenders) was not significantly different. Computing the SAC-IAT using either only pictures of males or females showed some improvement; although the two IATs were unable to distinguish between groups. The SAC-IAT Female (computed using only female pictures) had a larger correlation coefficient with PPG and the mean difference of the SAC-IAT Female scores between child molesters and non-sex offenders was larger compared to the original SAC-IAT, although not statistically significant. Clearly, trying to control for gender of the pictures statistically by removing either female or male pictures is not a preferred method as it (1) cannot account for all the effect, (2) reduces the sample size or diminishes the number of IAT trials. IAT measures created using only pictures of females or IAT measures created using only pictures of males would allow for more confident determination of the effect of gender salience on IATs adapted to measure sexual interest in children. In cases where an IAT measure adapted to assess sexual interest in children included only one gender (i.e., Banse et al., 2009), only the IAT measures using female pictures produced significant differences between child molesters and non-molesters. The IAT adapted to include only male pictures could not distinguish between groups. Although preliminary, these findings suggest that if pictures are included in an IAT measuring sexual interest, only one gender, and perhaps only female pictures, should be included. That being said, Brown et al.

(2009) included pictures of both males and female children, and some pictures depicted groups of children. Despite this, Brown and colleagues (2009) found large significant group difference on their IAT measures. As such, it is difficult to make a conclusive statement on the effect of gender on IAT measures adapted to assess sexual interest in children. One can intuitively presume that gender and sexual interest are very much related. That is, some individuals are sexually attracted to female adults and not sexually interested in male adults. It is also possible that some individuals are attracted to female children and not male children. Lastly, it is also possible for someone to be attracted equally to both genders. As such, a measure of sexual interest in children relative to adults may be maximized if only one gender is included.

Although the SAC-IAT could not differentiate between groups, it had convergent validity with other measures that could successfully discriminate between groups. The significant correlation between the SIPS Differential and the SAC-IAT as well as the correlation between Viewing Time Attractiveness Rating and the SAC-IAT approaching significance is interesting considering the context of the literature on implicit measures. Namely, researchers have argued that implicit measures and self-report measures are distinct in what they measure (e.g., Greenwald & Farnham, 2000). Banse et al. (2009) also found that their Girls-Women IAT was correlated with a self-report measure created for the purpose of the study. These findings support a more recent meta-analysis showing moderate correlations between the IAT and explicit self-report measures (Hofmann et al., 2005). Namely, Hofmann et al. (2005) found small to moderate relationship (depending on the contrast of interest) between implicit and self-report measures. These authors therefore concluded that the methods were related; however, that the relationship may be

moderated by certain variables, such as social desirability. In the current study, controlling for social desirability produced similar correlation between the SAC-IAT and the SIPS Differential, and slightly larger correlations with VT-AR and the Pedophilia Composite Score of the SIPS. The moderating effect of social desirability between explicit and implicit measures, however, can be theorized to be moderated by situational factors. Participants in this study were assured confidentiality and that no consequences would result from how they answered the questionnaire. Perhaps the moderating effect of social desirability would be more pertinent in more adversarial settings (e.g., pre-treatment assessments, pre-court assessment). Moreover, this study found a significant relationship with risk for general recidivism as measured by the SIR and the Impression Management subscale of the BIDR, with lower score on the SIR (representing higher risk) being associated with lower impression management. As such, it is possible that the BIDR actually measures something other than social desirability and therefore controlling for high scores on the BIDR may not be particularly relevant (Mills et al., 2003).

### **Limitations**

Interpretations of the findings from the current study should take into consideration the small sample size and issues regarding the representativeness of the sample. The average sex offender in Correctional Services Canada has a Static-99 score of 2; this sample had a score of 4.5 (Static-99 Clearinghouse, 2008). There were several moderate correlation coefficients and group differences that may have been significant had the analyses been conducted on more participants (thus increasing power). Confidence on the validity of these findings was, however, increased by controlling for possible confounding variables. The lack of group differences on the SAC-IAT was not

found to be because of group differences on confounding variables, such as age, sexual orientation, socially desirable responding, and vocabulary. Treatment participation and time since last sex offences, among other variables specific to child molesters, were also not found to account for the lack of group differences on the SAC-IAT.

### **Conclusion**

The current study lends support to IAT measures designed to assess sexual interest in children. Despite the current study not finding significant group differences on the SAC-IAT, the cumulative meta-analysis suggests that IAT measures are able to distinguish between groups ( $d = 0.70$ , 95% C.I. = 0.51 – 0.88). This finding suggests that child molesters and non-molesters do indeed differ in their implicit associations between child and sexual attractiveness. Although the type of research design does not support a cause and effect conclusion (that is, having a greater association between child and sex causes child sexual abuse), it does suggest that those convicted of child sexual abuse are more likely to hold implicit association between children and sex than those not convicted of child sexual abuse. This finding is similar to those based on self-reported and PPG-assessed sexual interest in children; child molesters are more likely to report or display sexual interest in children.

The current study also lend strong support to viewing time measures and self-report measures of sexual interest in children. Both the Viewing Time and the SIPS differential could distinguish between groups, with a moderate to large effect size. These measures were also correlated to other measures of sexual interest in children, suggesting adequate convergent validity.

Moreover, the present study found the SAC-IAT was related to measures of sexual interest. These findings support the contention that the SAC-IAT is indeed measuring sexual interest in children. The SAC-IAT was, however, not related to PPG. As mentioned previously, only 21 sex offenders had PPG data, and a significant majority of them were low responders on PPG (i.e., did not display sufficient arousal for the test to be considered valid). As such, the current study does not offer any conclusive statement on whether there is a relationship between PPG and IAT measures adapted to assess sexual interest in children. Further research is necessary to further understand the relationship between the IAT measure and other measures of sexual interest in children.

This study suggests that IAT measures adapted using pictures may be affected by a gender salient effect. It would be interesting to examine whether an IAT measure created to only include female pictures (i.e., girls relative to women) can distinguish between groups better than IAT measures using only words or using pictures of both genders. The present study also suggests that the viewing time measure is a promising measure of sexual interest in children. Similarly to the SAC-IAT, the relationship between viewing time measures and PPG needs to be further investigated to gain greater confidence on the validity of viewing times measures.

The IAT measures offer several practical implications. Although still in its developmental stages, implicit measures assessing sexual interest in children can allow an easier, relatively low cost, and quick method to assess sexual interest in children. PPG is intrusive, expensive, and is time consuming. In addition, PPG stimuli have yet to be standardized and self-report measures are sensitive to impression management. These limitations have created interest in producing new measures that can assess in sexual

interest in children more efficiently and reliably. IAT measures of sexual interest in children, however, continue to be premature for applied evaluation settings.

Before being used in clinical setting, research with larger and more representative samples must be done to further assess the psychometric properties of IAT measures assessing sexual interest in children. Although these IAT measures can, in general, discriminate between groups, their relationship with other measures of sexual interest is still relatively unknown, as is their relationship with sexual offending. The current study as well as Banse et al. (2009) does suggest that IAT measures assessing sexual interest in children are related to both viewing time and self-report measures. As such, available information is promising, however, inconclusive without further replication. Lastly, studies linking IAT scores with recidivism would allow for conclusions on the predictive accuracy of the IAT measures assessing sexual interest in children.

The field of sex offender research is fittingly involved in an ongoing process of validation and revision of its assessment instruments. More precise measures are developed through research, therefore improving our understanding of the construct of interest and strengthening current offender management and rehabilitation practices. Preliminary studies suggest that the IAT measure is a promising tool with which to measure sexual interest in children and, possibly, can contribute to more accurate and comprehensive information in the assessment of child molesters. Further evaluation of the IAT measure's convergent and predictive validity, however, is required before its use in the treatment and management of child molesters.

## References

- Abel, G. G., Huffman, J., Warberg, B., & Holland, L. C. (1998). Visual reaction time and plethysmography as measures of sexual interest in child molesters. *Sexual Abuse: A Journal of Research and Treatment, 10*, 81-95
- 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 & Neglect, 25*, 703-718.
- Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83*, 380-393.
- Baldwin, K. & Roys, D. T. (1998). Factors associated with denial in a sample of alleged adult sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 10*, 211-226.
- Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, III & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder. Science conference series* (pp. 117-150). Washington, DC: American Psychological Association.
- Banse, R., Schmidt, A.F., & Clarbour, J. (2009). *Profiles of explicit and implicit sexual interest in child sex offenders and controls*. Manuscript submitted for publication.
- Barbaree, H. E., Baxter, D. J., Marshall, W. L. (1989). Brief research report: The reliability of the rape index in a sample of rapists and nonrapists. *Violence and Victims, 4*, 299-306.

- Barbaree, H. E. & Marshall, W. L. (1989). Erectile responses among heterosexual child molesters, father-daughter incest offenders, and matched non-offenders: Five distinct age preference profiles. *Canadian Journal of Behavioural Science, 21*, 70-82.
- Barbaree, H. E., Seto, M. C., Langton, C. M., & Peacock, E. J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior, 28*, 490-521.
- Barsetti, I., Earls, C.M., Lalumière, M. L., & Bélanger, N. (1998). The differentiation of intrafamilial and extrafamilial heterosexual child molesters. *Journal of Interpersonal Violence, 13*, 275-286.
- Beech, A; Friendship, C; Erikson, M and Hanson, RK (2002). The relationship between static and dynamic risk factors and reconviction in a sample of UK child abusers. *Sexual Abuse: A Journal of Research and Treatment 14*, 155-167.
- Beitchman, J. H., Zucker, K. J., Hood, J. E., DaCosta, G. A., Akman, D., & Cassavia, E. (1992). A review of the long-term effects of child sexual abuse. *Child Abuse and Neglect, 16*, 101-118.
- Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*, 828 - 841.
- Blanchard, R., Klassen, P., Dickey, R., Kuban, M. E., & Blak, T. (2001). Sensitivity and specificity of the phallometric test for pedophilia. *Psychological Assessment, 13*, 118-126.

- Blanchard, R., Kuban, M. E., Blak, T., Cantor, J. M., Klassen, P., Dickey, R. (2006). Phallometric comparison of pedophilic interest in nonadmitting sexual offenders against stepdaughters, biological daughters, other biologically related girls, and unrelated girls. *Sexual Abuse: Journal of Research and Treatment, 18*, 1-14.
- Bonta, J., Harman, W. G., Hann, R. G., & Cormier, R. B. (1996). The prediction of recidivism among federally sentenced offenders: A re-validation of the SIR scale. *Canadian Journal of Criminology, 38*, 61-79.
- Brown, A., Gray, N. S., & Snowden, R. J. (2009). *Implicit measurement of sexual preferences in child sex abusers. Sexual Abuse: A Journal of Research and Treatment, 21*, 166-180.
- Browne, A., & Finkelhor, D. (1986). Impact of child sexual abuse: A review of the research. *Psychological Bulletin, 99*, 66-77.
- Canadian Centre for Justice Statistic (2005). *Family violence in Canada: A statistical profile* (Catalogue number:85-224-XIE). Ottawa, ON: National Clearinghouse on Family Violence.
- Cumming, G. F., & McGrath, R. J. (2005). *Supervision of the sex offender: Community management, risk assessment & treatment* (2<sup>nd</sup> Ed.). Brandon, Vermont: The Safer Society Press.
- Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 121*, 163-170.
- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443-451.

- de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior, 28*, 235-251.
- Doren, D. (2004). Stability of the interpretative risk percentages for the RRASOR and Static-99. *Sexual Abuse: A Journal of Research and Treatment, 16*, 25-36.
- Engloff, B. Schwerdtfeger, A., & Schmuckle, S. T. (2005). Temporal stability of the Implicit Association Test-Anxiety. *Journal of Personality Assessment, 84*, 82-88.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115-141.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297-327.
- Fernandez, Y. (January, 2002). Phallometric testing with sexual offenders against female victims: Limits to its value. *CSC Forum, 14*.
- Finkelhor, D. (1984). *Child sexual abuse: New theory and research*. New York: Free Press.
- Finkelhor, D. (1994). Current information on the scope and nature of child sexual abuse. *The Future of Children, 4*, 31-53.
- Freud, K. & Watson, R. J. (1991). Assessment of the sensitivity and specificity of a phallometric test: An update of phallometric diagnosis of pedophilia.

*Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3, 254-260.

Freund, K., Watson, R. J., & Dickey, R. (1991). Sex offenses against female children perpetrated by men who are not pedophiles. *Journal of Sex Research*, 28, 409-423.

Gannon, T. A., Keown, K., & Polaschek, D. L. L. (2007). Increasing honest responding on cognitive distortions in child molesters: The bogus pipeline revisited. *Sexual Abuse: A Journal of Research and Treatment* 19, 5-22.

Gannon, T. A., Polaschek, D. L. L. (2006). Cognitive distortions in child molesters: A re-examination of key theories and research. *Clinical Psychology Review*, 26, 1000-1019.

Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology*, 49, 171-180.

Gawronski, B., Hofmann, W., Wilbur, C. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition: An International Journal*, 15, 485-499.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us?: Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, 2, 181-193.

Goldman, J. D. G. & Padayachi, U. K. (2000). Some methodological problems in estimating incidence and prevalence in child sexual abuse research. *Journal of Sex Research*, 37, 305-314.

- Gorey, K. M., & Leslie, D. R. (1997). The prevalence of child sexual abuse: Integrative review adjustment for potential response and measurement biases. *Child Abuse & Neglect, 21*, 391-398.
- Govan, C. L. & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal of Experimental Social Psychology, 40*, 357-365.
- Gray, N. S., Brown, A. S., MacCulloch, M. J., Smith, J., & Snowden, R. J. (2005). An implicit association test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology, 114*, 304-308
- Gray, N. S., MacCulloch, M. J., Smith, J., Morris, M., & Snowden, R. J. (2003). Violence viewed by psychopathic murderers. *Nature, 423*, 497-498.
- Greenwald, A. G. (2004, January). *Revised Top 10 List of Things Wrong with the IAT*. Invited presentation at Attitudes Preconference of the 5th annual meeting of the Society of Personality and Social Psychology, Austin, TX. Retrieved September 10, 2007 from <http://faculty.washington.edu/agg/unpublished.html>
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79*, 1022-1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

- Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie*, 48, 85-93. Retrieved September 10, 2007 from [http://faculty.washington.edu/agg/iat\\_validity.html](http://faculty.washington.edu/agg/iat_validity.html)
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197-216.
- Gress, C. L. Z. (2005). Viewing time measures and sexual interest: another piece of the puzzle. *Journal of Sexual Aggression*, 11, 117-125.
- Gress, C. L. Z., & Laws, D. R. (2004). *NRP (Not Real People) Stimulus Sets A & B*. Victoria, BC: Pacific Psychological Assessment Corporation.
- Grossman, L. S., Cavanaugh, J. L., & Haywood, T. W. (1992). Deviant sexual responsiveness on penile plethysmography using visual stimuli: Alleged child molesters vs. normal control subjects. *Journal of Nervous and Mental Disease*, 180, 207-208
- Grubin, D. (1998). Sex offending against children: Understanding the risk. Police Research Series Paper 99. London: Home Office.
- Grubin, D. & Wingate, S. (1996). Sexual offence recidivism: Prediction versus understanding. *Criminal Behaviour and Mental Health*, 6, 349-359.
- Hall, G. C. N. & Hirschman, R. (1991) Toward a theory of sexual aggression: A quadripartite model. *Journal of Consulting and Clinical Psychology*, 59, 662-669.
- Hall, G. C. N., & Hirschman, R. (1992). Sexual aggression against children: A conceptual perspective of etiology. *Criminal Justice and Behavior*, 19, 8-23.

- Hanson, R. K. (1997). *The development of a brief actuarial risk scale for sexual offence recidivism* (User Report No. 1997-04). Ottawa: Department of the Solicitor General of Canada. Retrieved November 10, 2008 from <http://www.sgc.gc.ca/epub/corr/e199704/e199704.htm>
- Hanson, R. K. (2001). Note on the reliability of Static-99 as used by the California Department of Mental Health evaluators (Unpublished report). Sacramento, CA: California Department of Mental Health.
- Hanson, R. K. (2002). Recidivism and age: Follow-up data from 4,673 sexual offenders. *Journal of Interpersonal Violence, 17*, 1046-1062.
- Hanson, R. K. (2006). Does Static-99 predict recidivism among older sexual offenders? *Sexual Abuse: A Journal of Research and Treatment, 18*, 343-355.
- Hanson, R. K., & Broom, I. (2005). The utility of cumulative meta-analysis: Application to programs for reducing sexual violence. *Sexual Abuse: A Journal of Research and Treatment, 17*, 357-373.
- Hanson, R. K., Gizzarelli, R., & Scott, H. (1994). The attitudes of incest offenders: Sexual entitlement and acceptance of sex with children. *Criminal Justice and Behaviour, 21*, 187-202.
- Hanson, R. K., & Harris, A. J. R. (2000). *Sex Offender Need Assessment Rating (SONAR): A method for measuring change in risk levels* (User Report # 2000-1). Ottawa, ON: Public Safety Canada.
- 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*. Ottawa: Public Safety Canada.

- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis*. Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R. K., & Morton-Bourgon, K. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154-1163.
- Hanson, R. K., & Morton-Bourgon, K. (2007). *The Accuracy of Recidivism Risk Assessments for Sexual Offenders: A Meta-Analysis*. Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R. K., & Thornton, D. (1999). *Static -99: Improving actuarial risk assessments for sex offenders* (User Report 99-02). Ottawa, ON: Department of the Solicitor General of Canada.
- Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*, 119-136
- Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised*. Toronto, ON: Multi-Health System.
- Harris, A. J. R. (2008, October). *Stable-2007 and Acute-2007 training*. Pre-conference workshop delivered at the 27<sup>th</sup> Association for the Treatment of Sexual Abusers, Atlanta, Georgia.
- Harris, A. J. R., & Hanson, R. K. (2004). *Sex offender recidivism: A simple question*. Ottawa, ON: Public Safety and Emergency Preparedness.

- Harris, A. J. R., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static-99 coding rules: Revised 2003*. Ottawa, ON: Department of the Solicitor General of Canada.
- Harris, G. T., Rice, M. E., Cormier, C. A., Quinsey, V. L. (1998). Dissimulation in phallometric testing of rapists' sexual preferences. *Archives of Sexual Behavior*, 28, 223-232.
- Harris, G. T., Rice, M. E., Quinsey, V. L., & Chaplin, T. C. (1996). Viewing time as a measure of sexual interest among child molesters and normal heterosexual men. *Behaviour Research and Therapy*, 34, 389-394.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369-1385.
- Hogue, T. E. (1994). Sex Offence Information Questionnaire: Assessment of sexual offenders' perceptions of responsibility, empathy and control. *Issues in Criminological & Legal Psychology*, 21, 68-75.
- Horley, J. (2000). Cognitions supportive of child molestation. *Aggression and Violent Behavior*, 5, 551-564.
- Houben, K., & Wiers, R. W. (2006). A test of the salience asymmetry interpretation of the alcohol-IAT. *Experimental Psychology*, 54, 292-300.
- Hudson, S. M., Johnston, L., & Marshall, W. L. (1997). Cognitive distortions in sex offenders: An integrative review. *Clinical Psychology Review*, 17, 479-507.

- Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the Implicit Association Test to measure age differences in implicit social cognitions. *Psychology and Aging, 17*, 482-495.
- Jajodia, A. & Earleywine, M. (2003). Measuring alcohol expectancies with the implicit association test. *Psychology of Addictive Behaviors, 17*, 126-133.
- Johnson, K. A. & Tyler, K. A. Adolescent sexual onset: An intergenerational analysis. *Journal of Youth and Adolescence, 36*, 939-949.
- Kalmus, E., & Beech, A. R. (2005). Forensic assessment of sexual interest: A review. *Aggression and Violent Behavior, 10*, 193-217.
- Kazdin, A. E. (2003). *Research design in clinical psychology* (4<sup>th</sup> ed.). Boston, MA: Allyn and Bacon.
- Kessous, N. (2009). *Under the radar: Assessing attitudes towards child sexual abuse using the Implicit Association Test*. Unpublished master's thesis, Carleton University, Ottawa, Ontario, Canada.
- Kim, D. Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66*, 83-96.
- Koss, M. P. & Gidycz, C. A. (1985). Sexual Experiences Survey: Reliability and validity. *Journal of Consulting and Clinical Psychology, 53*, 422-423.
- Koss, M. P. & Oros, C. J. (1982). Sexual Experiences Survey: A research instrument investigating sexual aggression and victimization. *Journal of Consulting and Clinical Psychology, 50*, 455-457.

- Kroner, D. G. and Weekes, J. R. (1996). Balanced inventory of desirable responding: Factor structure, reliability and validity with an offender sample. *Personality and Individual Differences, 21*, 323–333.
- Lalumière, M. L., & Harris, G. T. (1998). Common questions regarding the use of phallometric testing with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 10*, 227-237.
- Langevin, R. & Watson, R. (1991). A comparison of incestuous biological and stepfathers. *Annals of Sex Research, 4*, 141-150.
- Launay, G. (1999). The phallometric assessment of sex offenders: an update. *Criminal behaviour and Mental Health, 9*, 254–274.
- Laws, D. R. (1986). *Sexual deviance card sort*. Unpublished manuscript, Florida Mental Health Institute, Tampa.
- 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.
- Laws, D. R., Hanson, R. K., Osborn, C. A., Greenbaum, P. E. (2000). Classification of child molesters by plethysmographic assessment of sexual arousal and a self-report measure of sexual preference. *Journal of Interpersonal Violence, 15*, 1297-1312.
- Letourneau, E. J. (2002). A Comparison of Objective Measures of Sexual Arousal and Interest: Visual Reaction Time and Penile Plethysmography. *Sexual Abuse: Journal of Research and Treatment, 14*, 207-223.

- Li, A. & Bagger, J. (2007). The Balanced Inventory of Socially Desirable Responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement, 67*, 525-544.
- Lisak, D. & Miller, P. M. (2002). Repeat rape and multiple offending among undetected rapists. *Violence and Victims, 17*, 73-84.
- Looman, J. (2006). Comparison of two risk assessment instruments for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 18*, 193-206.
- Looman, J., & Marshall, W. L. (2001). Phallometric assessments designed to detect arousal to children: The responses of rapists and child molesters. *Sexual Abuse: A Journal of Research and Treatment, 13*, 3-13.
- Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology, 14*, 405-415.
- Mann, R., Webster, S., Wakeling, H., & Marshall, W. L. (2007). The measurement and influence of child sexual abuse supportive beliefs. *Psychology, Crime & Law, 13*, 443-458.
- Marshall, W. L., Anderson, D., & Fernandez, Y. (1999). *Cognitive behavioural treatment of sexual offenders*. New York: Wiley.
- Marshall, W. L. & Eccles, A. (1991). Issues in clinical practice with sex offenders. *Journal of Interpersonal Violence, 6*, 68-93.
- Marshall, W. L., & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review, 20*, 807-822.

- McConaghy, N. (1999). Unresolved issues in scientific sexology. *Archives of Sexual Behavior, 28*, 285-318.
- McFarland, S. G. & Crouch, Z. (2002). A cognitive skill confound on the Implicit Association Test. *Social Cognition, 20*, 483-510.
- McGrath, R. J., Cumming, G. F., & Burchard, B. L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 nationwide survey*. Brandon, Vermont: The Safer Society Press.
- Mihailides, S., Devilly, G. J., & Ward, T. (2004). Implicit cognitive distortions and sexual offending. *Sexual Abuse: A Journal of Research and Treatment, 16*, 333-350.
- Mills, J. F., & Kroner, D. G. (2003). Anger as a predictor of institutional misconduct and recidivisms. *Journal of Interpersonal Violence, 18*(1), 282-294.
- Mills, J. F., & Kroner, D. G. (2006). Impression management and self-report among violent offenders. *Journal of Interpersonal Violence, 21*, 178-192.
- Mills, J. F., Loza, W., & Kroner, D. G., 2003. Predictive validity despite social desirability: Evidence for the robustness of self-report among offenders. *Criminal Behaviour and Mental Health, 13*, 144–154.
- Miner, M. H., West, M. A., & Day, D. M (1995). Sexual preference for child and aggressive stimuli: Comparison of rapists and child molesters using auditory and visual stimuli. *Behaviour Research and Therapy, 33*, 545-551.
- Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*, 455-469.

Nichols, H. R., & Molinder, I. (2000). *Multiphasic Sex Inventory: Second Edition*.

Tacoma, WA: Crime and Victim Psychology Specialists.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6, 101-115.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, 31, 166-180.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265-292). Psychology Press.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test. *Experimental Psychology*, 54, 14-29.

Nuffield, J. (1982). *Parole decision making in Canada: Research towards decisions guidelines*. Ottawa: Communication Division, Solicitor General Canada.

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 Behavior*, 34, 454-475.

Nuffield, J. (1982). *Parole decision making in Canada: Research towards decisions guidelines*. Ottawa: Communication Division, Solicitor General Canada.

- O'Donohue, W., & Letourneau, E. (1992). The psychometric properties of the penile tumescence assessment of child molesters. *Journal of Psychopathology and Behavioral Assessment, 14*, 123–174.
- Paitich, D. (1977). *The Clarke Vocabulary Scale: A Multiple-Choice Estimate of WAIS Vocabulary Level*. Clarke Institute of Psychiatry: Toronto, Ontario.
- Palfai, T. P. & Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: Assessing implicit response tendencies using the modified-IAT. *Behaviour Research and Therapy, 41*, 1149-1162.
- Paolucci, E. O., Genuis, M. L., & Violato, C. (2001). A meta-analysis of the published research on the effects of child sexual abuse. *The Journal of Psychology, 135*, 17-36.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598-609.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C. & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12*, 729-738.
- Pithers, W. D., & Laws, D. R. (1995). Phallometric assessment. In B. K. Schwartz & H. R. Cellini (Eds.), *The sex offender: Corrections, treatment and legal practice* (Vol. 1). Kingston, NJ: Civic Research Institute.
- Poehlman, T. A., Uhlmann, E. L., Greenwald, A. G., & Banaji, M. R. (2005). *Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity*. Manuscript submitted for publication.

- Priest, R. & Smith, A. (1992). Counselling adult sex offenders: Unique challenges and treatment paradigms. *Journal of Counseling & Development, 71*, 27-32.
- Quinsey, V. L. & Chaplin, T. C. (1984). Stimulus control of rapists' and non-sex offenders' sexual arousal. *Behavioral Assessment, 6*, 169-176.
- Quinsey, V. L., & Chaplin, T. C. (1988). Penile responses of child molesters and normals to descriptions of encounters with children involving sex and violence. *Journal of Interpersonal Violence, 3*, 259-274.
- Quinsey, V. L., Chaplin, T. C., & Carrigan, W. F. (1979). Sexual preferences among incestuous and nonincestuous child molesters. *Behavior Therapy, 10*, 562-565.
- Quinsey, V. L., Coleman, G., Jones, B., & Altrows, I. F. (1997). Proximal antecedents of eloping and reoffending among supervised mentally disordered offenders. *Journal of Interpersonal Violence, 12*, 794 – 813.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Quinsey, V. L., Ketsetzis, M., Earls, C., & Karamanoukian, A. (1996). Viewing time as a measure of sexual interest. *Ethology & Sociobiology, 17*, 341-354.
- Rice, M. E. & Harris, G. T. (2002). Men who molest their sexually immature daughters: is a special explanation required? *Journal of Abnormal Psychology, 111*, 329-339.
- Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's *d* and *r*. *Law and Human Behavior, 29*, 615-620.

- Roberts, C. F., Doren, D. M., & Thornton, D. (2002). Dimensions associated with assessments of sex offender recidivism risk. *Criminal Justice and Behavior, 29*, 569-589.
- Robinson J. P., Shaver P. R., & Wrightsman L. S. (1991). *Measures of personality and social psychological attitudes*. San Diego: Academic Pr.
- Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437-465.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime User's Guide. Psychology Software Tools, Inc., Pittsburgh.
- Seto, M. C. (2004). Pedophilia and sexual offenses against children. *Annual Review of Sex Research, 15*, 321-361.
- Seto, M. C. (2008). *Pedophilia and sexual offending against children: Theory, assessment, and intervention*. American Psychological Association: Washington, DC.
- Seto, M. C., Cantor, J. C. & Blanchard, R. (2006). Child pornography offenses are a valid diagnostic indicator of pedophilia. *Journal of Abnormal Psychology, 115*, 610-615.
- Seto, M. C., Harris, G. T, Rice, M. E., & Barbaree, H. E. (2004). The Screening Scale for Pedophilic Interests predicts recidivism among adult sex offenders with child victims. *Archives of Sexual Behavior, 33*, 455-466.

- Seto, M. C., & Lalumière, M. L. (2001). A brief screening scale to identify pedophilic interests among child molesters. *Sexual Abuse: A Journal of Research and Treatment, 13*, 15-25.
- Seto, M. C., Lalumière, M. L., & Kuban, M. (1999). The sexual preferences of incest offenders. *Journal of Abnormal Psychology, 108*, 267–272.
- Seto, M. C., Murphy, W.D., Page, J., & Ennis, L. (2003). Detecting anomalous sexual interests in juvenile sex offenders. *Annals of the New York Academy of Sciences, 989*, 118-130.
- Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and explicit attitudes toward cigarette smoking: The effects of context and motivation. *Journal of Social & Clinical Psychology, 22*, 13-39.
- Simon, W. T., & Schouten, P. G. (1991). Plethysmography in the assessment and treatment of sexual deviance: An overview. *Archives of Sexual Behavior, 20*, 75–91.
- Sjöstedt, G., & Langström, N. (2001). Actuarial assessment of sex offender recidivism risk: A cross-validation of the RRASOR and the Static-99 in Sweden. *Law and Human Behavior, 25*, 629-645.
- Snowden, R. J., Gray, N. S., Smith, J., Morris, M., & MacCulloch, M. J. (2004). Implicit affective associations to violence in psychopathic murderers. *The Journal of Forensic Psychiatry and Psychology, 15*, 620-641.
- Static-99 Clearinghouse (2008). *Norms*. Retrieved July 13, 2009 from [www.Static99.org](http://www.Static99.org)
- Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*, 165-179.

- Steffens, M. C., Yundina, E., & Panning, M. (2008). Automatic associations with "erotic" in child sexual offenders: Identifying those in danger of reoffence. *Sexual Offender Treatment, 3*, 1-9.
- Stinson, J. D., & Becker, J. V. (2008). Assessing sexual deviance: A comparison of physiological, historical, and self-report measures. *Journal of Psychiatric Practice, 14*, 379-388.
- Studer, L. H., Aylwin, A. S., Clelland, S. R., Reddon, J. R., & Frenzel, R. R. (2002). Primary erotic preference in a group of child molesters. *International Journal of Law and Psychiatry, 25*, 173–280.
- Tan, L., & Grace, R. C. (2008). Social desirability and sexual offenders : A review. *Sexual Abuse: A Journal of Research and Treatment, 20*, 61-87.
- Teachman, B. A., Gregg, A. P., & Woody, S. R. (2001). Implicit associations for fear-relevant stimuli among individuals. *Journal of Abnormal Psychology, 110*, 226-235.
- Teachman, B. A. & Woody, S. R. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112*, 100-109.
- Terry, K. J. (2006) *Sexual offenses and offenders: theory, practice and policy*. Belmont: Thomson Wadsworth.
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14*, 137–151.
- Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., et al. (2003). Distinguishing and combining risks for sexual and violent recidivism. In R. A.

- Prentky, E. S. Janus, & M. C. Seto. (Eds.), *Sexually Coercive Behavior: Understanding and Management* (pp. 225-235). New York: New York Academy of Sciences.
- Uhlmann, E., & Swanson, J. (2004). Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence*, *27*, 41-52.
- Ward, T. (2000). Sexual offenders' cognitive distortions as implicit theories. *Aggression and Violent Behaviour*, *5*, 491–507.
- Ward, T. & Beech, A. (2006). An integrated theory of sexual offending. *Aggression and Violent Behavior*, *11*, 44–63.
- 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.
- Ward, T., & Keenan, T. (1999). Child molesters' implicit theories. *Journal of Interpersonal Violence*, *14*, 821–838.
- Ward, T., & Siegert, R. J. (2002). Toward a comprehensive theory of child sexual abuse: A theory knitting perspective. *Psychology, Crime & Law*, *8*, 319-351.
- Whitaker, D. J., Brenda, L., Hanson, R. K., Baker, C. K., McMahon, P. M., Ryan, G., Klein, A., & Rice, D. D. (2008). Risk Factors for the Perpetration of Child Sexual Abuse: A Review and Meta-Analysis. *Child Abuse & Neglect*, *32*, 529-548.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*, 815-827.
- Wormith, J. S. (1986). Assessing deviant sexual arousal: Physiological and cognitive aspects. *Advances in Behaviour Research & Therapy*, *8*, 101-137.

Yates, P. M., Goguen, B. C., Nicholaichuk, T. P., Williams, S.M., & Long, C.A. (2000).  
*National Sex Offender Programs (Volume II: Moderate Intensity)*. Ottawa, ON:  
Correctional Service of Canada.

## Appendix A

## Internal validity and reliability of IAT measures

Table 1A presents a summary of that affect the psychometric properties of the IAT measure and, when available, solutions are provided.

Table 1A. *Issues Affecting the Validity and Reliability of the IAT Measure*

Issues	Solution
Increased salience of the association under study due to self-report measures preceding the administration of the IAT.	Counterbalance the administration of the measures (Fazio & Olson, 2003; Nosek, Greenwald, & Banaji, 2005; Nosek et al., 2007).
Order of the combined task affects the IAT. That is, the first completed pairing tends to show a higher association compared to the second pairing, when the compatible version is presented first.	Increase the practice trial in block five from 20 to 40 rather and counterbalance the two combined tasks (Nosek, Greenwald, & Banaji, 2005; Nosek et al., 2007).
Experience with the IAT reduces the IAT effect.	The new scoring algorithm ( <i>D</i> ) reduces this effect. The use of a comparison group when many administrations of the IAT cannot be avoided (such as in longitudinal studies) is also suggested (Greenwald, Nosek, & Banaji, 2003; Teachman & Woody, 2003).
Exposure to positive or negative cues about the concept of interest before the administration of the IAT have been found to influence IAT effects in the expected direction (e.g., positive cues about a concept of interest has been found to increase positive associations as measured by the IAT). Word choice (e.g., valence as well as number of words to use per category have been found to influence the IAT). Uneven valence between categories has been found to alter IAT effects. Number of stimuli does not influence the IAT as long as there are two or more words per category.	Constructing an environment free of evaluation may help negate this effect.
Cognitive fluency (i.e., average response speed of participants) is related to age and has been found to positively influence the IAT. That is, individual lower in cognitive fluency (and usually older in age) are found to have larger IAT effects.	Attempt to select stimulus words to create equal valence between categories. The number of stimuli to be used per category is optimally four; however should not be lower than two (Greenwald, 1998; Nosek, Greenwald, & Banaji, 2005). New algorithm ( <i>D</i> ) reduces the effect of cognitive fluency and age (Greenwald, Nosek, & Banaji, 2003). Mean age of the groups should, however, be similar.

*Note.* Based on Revised Top 10 List of Things Wrong with the IAT (Greenwald, 2004) and Nosek et al. (2007).

### Counterbalancing the Administration of Measures

Several methodological issues have been raised regarding the order of administration of the IAT and self-report measures. Completing a self-report measure before the IAT has been posited to result in increased salience of the associations under investigation and, consequently, increased IAT effects (Fazio & Olson, 2003). The research, however, has found little difference between IAT measures administered prior or after the respective self-report measure. For instance, Hofmann et al. (2005) reported that when the self-report measure precedes the administration of the IAT measure, the correlation between measures was very similar ( $r = .24$ ; mean of 53 studies) to when the IAT measure preceded the administration of the self-report measure ( $r = .22$ , mean of 48 studies). An internet-based study conducted by Nosek, Greenwald, and Banaji (2005) included four different IAT measures and their respective self-report scales. Depending on the particular IAT measures, sample size ranged from 5,254 to 21,925 participants. This study also found little to no effect on magnitude of the correlation between implicit and explicit measures as a function of the order in which the measures were presented. Nevertheless, counterbalancing the administration of self-report measure(s) and the IAT measure is suggested in order to reduce the potential influence of increase in salience on the IAT effect by distributing the error associated with this effect evenly between participants (Fazio & Olson, 2003; Nosek et al., 2005; Nosek et al., 2007).

### Order of the Combined Task

The most commonly reported issue influencing the IAT effects is the order effect of combined tasks (Greenwald & Nosek, 2001). Specifically, Greenwald et al. (1998) noted that IAT effects are slightly biased toward indicating stronger associations in the first combined pairing compared to the second combined pairing, when the congruent task is completed first. For example, if Mr. X has a higher relative sexual interest in children, than this effect would be present if Mr. X first completed the *child* and *sexy* combination. The completion of the first combined task interferes with performance of the second combination task. Nosek et al. (2005), however, found that this particular order effect was significantly reduced by increasing the amount of trials in the fifth block (e.g., practice trial of the concept category, such as face of male/ face of female) from 20 to 40 trials. Consequently, the authors suggested that this new procedure be adopted. Nosek et al. (2007) also suggested the counterbalancing of the two combined tasks to allow for the statistical removal of this effect. Some research, however, suggests that counterbalancing compatible and incompatible blocks may result in added error and reduce correlations between self-report and implicit measures (e.g., Gawronski, 2002). Interestingly, Hofmann et al. (2005) found that the relationship between self-report measures and the IAT measure was positively influenced by counterbalancing the compatible and incompatible blocks in the IAT measure. That is, correlations were significant for studies in which compatible and incompatible blocks were counterbalanced (mean  $r = .25$ ) compared to studies that did not counterbalance compatible and incompatible blocks (mean  $r = .18$ ).

### Experience with the IAT

A participant's experience with the IAT measure, such as during longitudinal studies or during the administration of multiple IATs in a short interval, may reduce IAT

effects (Greenwald & Nosek, 2001; Greenwald et al., 2003). The updated scoring algorithm (*D*) significantly reduces this effect (Greenwald, Nosek, & Banaji, 2003). Nevertheless, experience with the IAT should be controlled for during studies administering multiple IATs (Nosek et al., 2007). For instance, the inclusion of a comparison group that would not be expected to change as a function of the intervention could be included to statistically account for the effect of experience on the IAT measure (Teachman & Woody, 2003). That being said, IAT measures have been successful in measuring change following a certain stimuli (e.g., before and after a violent video game; Uhlmann & Swanson, 2004) and treatment program (e.g., Teachman & Woody, 2003). Studies such as these suggest that experience with IAT measures may not be too strong as to reduce the clinical usefulness of an IAT measure by clouding (or increasing) treatment effect. That being said, studies continue to be needed to verify the test-retest reliability of IAT measures.

### **Cues Before Administration**

Past research suggests that exposure to positive or negative cues about the concept of interest before administering the IAT measure may influence the findings. For instance, Wittenbrink, Judd, and Park (2001) exposed participants to either a video of an African American family enjoying a joyful barbecue or a gang-related video before the administration of an IAT measure adapted to measure racial attitude toward African Americans. The video of a joyful family barbecue was found to increase the positive evaluation of African Americans. Similarly, Blair, Ma, and Lenton (2001) reported that participants who were instructed to imagine a counter-stereotypical female before the administration of their IAT measure demonstrated significantly lower gender stereotyping as opposed to participants envisioning a gender-neutral topic. It can be hypothesized that an environment that is either positive or negative to child molesters would similarly affect the IAT effect. As such, it would be beneficial to construct a neutral environment without any evaluation of the construct of interest in order to reduce this type of error.

### **Valence of Stimuli**

Valence of stimulus words (i.e., negative, neutral or positive meanings) may also influence IAT effects (Fazio & Olson, 2003). Govan and Williams (2004) found that the IAT effect can be influenced by varying the valence of stimulus words representing the concept categories *flowers* and *insects* to represent liked insects (e.g., butterfly) and disliked flowers (poison ivy). De Houwer (2001) balanced the amount of positive and negative stimulus words (e.g., Albert Einstein and Adolf Hitler were used as stimulus words to represent German category, and Princess Diana and Donald Shipman, a mass-murderer, were used to represent British category) in an IAT constructed to assess attitudes toward foreign individuals. In contrast to Govan and Williams' (2004) findings, the IAT effect was similar whether the valence of stimulus words was positive or negative (De Houwer, 2001). The construction of an IAT measure should take into consideration the possible effect of valence when selecting stimulus words and reduce discrepancy between categories as much as possible (Nosek et al., 2007). It is difficult, however, to use words that represent the category *sexy* that are not explicitly good or,

conversely, to use words that represent the *not sexy* category that are not explicitly negative.

### **Stimuli Modality**

Different modalities are used to distinguish between stimuli belonging to the two different category pairs, such as using pictures to represent the concept category pair (*adult* and *child*) and words to represent the attribute category pair (*sexy* and *not sexy*). More commonly, difference in fonts can also be used to differentiate between stimuli. Picture stimuli have been associated with faster reaction times and slightly smaller IAT effects compared to word stimuli (Mitchell, Nosek, & Banaji, 2003; Nosek et al., 2002). Nevertheless, Brown et al. (2008) and Banse et al. (2009) used picture in IAT measures designed to assess sexual interest in children and continued to be able to differentiate between child molesters and non-sex offenders. Therefore, when assessing child molesters or the construct of sexual interest, an IAT measure utilizing pictures is expected to produce similar, if not better, discrimination between groups than an IAT utilizing only stimulus words.

### **Number of Stimuli**

The number of stimuli to include in IAT measures is also of importance. Greenwald and colleagues (1998) found that using five as opposed to twenty-five stimuli per category did not result in any significant difference in the IAT effect. Correspondingly, Hofmann et al. (2005) reported that the correlation between implicit and self-report measures did not vary with respect to the number of target stimuli (mean  $r = .03$ ) or attribute stimuli (mean  $r = .06$ ). Nosek, Greenwald, and Banaji (2005) examined variance in IAT effects as a function of the number of stimulus words (ranging from one to six items per category). Their results indicated that at least four stimulus items per category is ideal, however two items per category is sufficient.

### **Cognitive Fluency and Age**

Cognitive fluency refers to the speed at which an individual process information, which can be measured by reaction time. Slow cognitive fluency (i.e., participants who, on average, have lower response speeds) has been associated with larger IAT effects as compared to those with faster response speeds (Greenwald et al., 2003; McFarland & Crouch, 2002). The age of participants has also been associated with IAT effect, perhaps due to the relationship between age and cognitive fluency, with older age corresponding to diminished cognitive fluency. More precisely, older participants tend to show larger IAT effects as compared to younger participants (Greenwald & Nosek, 2001; Hummert, Garstka, O'Brien, Greenwald, & Mellott, 2002). The issue of age and cognitive fluency is especially relevant to the current study because child molesters are usually older than non-sex offenders. For instance, of the six studies that adapted the IAT measures to assess sexual interest in children, the child molesters, on average, were at least ten years older than the comparison group. The updated scoring algorithm ( $D$ ) has been found to reduce the effect of age on the IAT measure (Greenwald et al., 2003). Nevertheless, the mean group age should be examined to verify consistency between groups, and if required, age should be used as a covariate in the analyses.

*Appendix B*  
SAC-IAT Stimuli

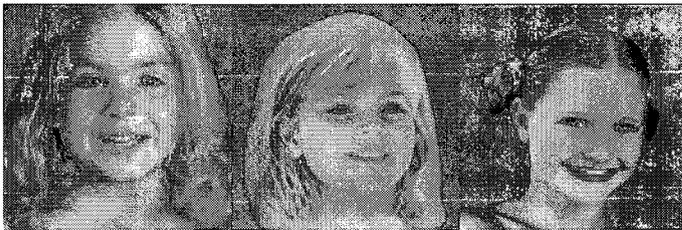
**Attribute Category**

*Sexy Words.* beautiful, love, attractive, kiss, smile, orgasm.

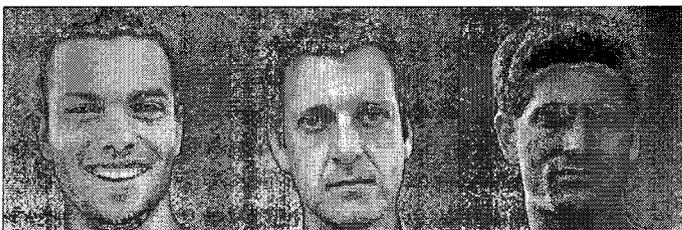
*Not Sexy Words.* yuck, disgusting, impotent, stink, gross, ugly.

**Concept Stimuli**

*Child Pictures.*



*Adult Pictures.*



*Appendix C*

## Balanced Inventory of Desirable Responding

Item	Not True					Very true	
1. My first impressions of people usually turn out to be right.	1	2	3	4	5	6	7
2. It would be hard for me to break any of my bad habits.	1	2	3	4	5	6	7
3. I don't care to know what other people really think of me.	1	2	3	4	5	6	7
4. I have not always been honest with myself.	1	2	3	4	5	6	7
5. I always know why I like things.	1	2	3	4	5	6	7
6. When my emotions are aroused, it biases my thinking.	1	2	3	4	5	6	7
7. Once I've made up my mind, other people can seldom change my opinion.	1	2	3	4	5	6	7
8. I am not a safe driver when I exceed the speed limit.	1	2	3	4	5	6	7
9. I am fully in control of my own fate.	1	2	3	4	5	6	7
10. It's hard for me to shut off a disturbing thought.	1	2	3	4	5	6	7
11. I never regret my decisions.	1	2	3	4	5	6	7
12. I sometimes lose out on things because I can't make up my mind soon enough.	1	2	3	4	5	6	7
13. The reason I vote is because my vote can make a difference.	1	2	3	4	5	6	7
14. My parents were not always fair when they punished me.	1	2	3	4	5	6	7
15. I am a completely rational person.	1	2	3	4	5	6	7
16. I rarely appreciate criticism.	1	2	3	4	5	6	7
17. I am very confident of my judgments.	1	2	3	4	5	6	7
18. I have sometimes doubted my ability as a lover.	1	2	3	4	5	6	7

- |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| 19. It's all right with me if some people happen to dislike me.       | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 20. I don't always know the reasons why I do the things I do.         | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 21. I sometimes tell lies if I have to.                               | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 22. I never cover up my mistakes.                                     | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 23. There have been occasions when I have taken advantage of someone. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 24. I never swear.  | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 25. I sometimes try to get even rather than forgive and forget.       | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 26. I always obey laws, even if I'm <b>unlikely</b> to get caught.    | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 27. I have said something bad about a friend behind his/her back.     | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 28. When I hear people talking privately, I avoid listening.          | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
-

## Appendix D

## Clarke Vocabulary Scale

INSTRUCTIONS: Please circle the response that has the most similar meaning to the underlined word.

Item	Possible responses				
1. a book	A. is to read	B. is a ball	C. is blue	D. is to run	
2. a car	A. is a dance	B. is to drive	C. is black	D. is to eat	
3. a horse	A. has four	B. is small	C. lives in a	D. flies	
4. a clock	A. keeps you dry	B. is deep	C. tells time	D. is for cats	
5. a hammer	A. grows on a tree	B. is a tool	C. is for you leg	D. has a shell	
6. a pillow	A. is an animal	B. is hard	C. is to open	D. is to sleep on	
7. a glove	A. goes on your hand	B. goes on your head	C. is to work with	D. is a tree	
8. a saw	A. is to cut with	B. is to sit on	C. is to point with	D. is a page	
9. fur	A. is blue	B. is animal	C. is sweet	D. is on birds	
10. jewel	A. round	B. sky	C. flower	D. diamond	
11. connect	A. spill	B. part	C. join	D. feel	
12. shovel	A. farm	B. spade	C. cry	D. ace	
13. weapon	A. sword	B. correct	C. lip	D. pecan	
14. nuisance	A. baby	B. help	C. worse	D. bother	
15. ridiculous	A. silly	B. ruined	C. poor	D. defenceless	
16. nimble	A. fat	B. active	C. brave	D. sick	
17. wager	A. court	B. dice	C. gamble	D. strike	
18. magnify	A. make	B. distant	C. make	D. handle	
19. fable	A. fur	B. watch	C. rhyme	D. story	
20. facetious	A. hungry	B. witty	C. weighty	D. ugly	
21. blemish	A. flaw	B. abuse	C. vaunt	D. clean	
22. strife	A. try	B. verse	C. battle	D. wine	
23. seclude	A. ensure	B. dream	C. isolate	D. obtain	
24. recede	A. withdraw	B. deny	C. repeat	D. dine	
25. profusion	A. glitter	B. plenty	C. noise	D. glory	

---

26. adversary	A. opponent	B. explosive	C. light	D. grammar
27. effrontery	A. bombing	B. boldness	C. triviality	D. belief
28. impending	A. valuable	B. prominent	C. imminent	D. obese
29. abyss	A. confusion	B. suicide	C. monastery	D. chasm
30. diverge	A. asset	B. vacation	C. distress	D. wander
31. devoid	A. planet	B. vacuum	C. missing	D. insect
32. arduous	A. laborious	B. shapeless	C. sterile	D. proud
33. vigilant	A. watchful	B. careful	C. literal	D. winning
34. relinquish	A. food	B. plague	C. distress	D. surrender
35. catacomb	A. deep valley	B. Japanese garden	C. ruined city	D. underground cemetery
36. indemnity	A. penalty	B. subsidy	C. amity	D. ordinance
37. reprobation	A. approval	B. censure	C. slander	D. verification
38. dilatory	A. slow	B. expanded	C. selfish	D. uncertain
39. requisite	A. admonish	B. love	C. repay	D. commend
40. imprecation	A. inaccuracy	B. scruple	C. curse	D. rarity

---

*Appendix E*  
Background Questionnaire

Question	Response
1. Do you speak English fluently?	Yes No
2. Do you understand written English?	Yes No
3. Do you understand spoken English?	Yes No
4. Is English your first language?	Yes No
5. Did you attend an English grade school (grades 1 to 8)?	Yes No
6. Did you attend an English high school?	Yes No
7. Who are you most sexually attracted to?	Women, Men, or Both men and women
8. Have you ever used physical force or threats to make an adult have sexual contact with you when he or she didn't want to?	Yes No
9. Have you ever had sexual contact with a child under the age of 16 when you were at least 4 years older than the child?	Yes No
10. Have you ever had sexual contact with a child under the age of 14 when you were at least 4 years older than the child?	Yes No
11. Have you ever had sexual contact with a child under the age of 12 when you were at least 4 years older than the child?	Yes No
12. Have you ever had sexual contact with an animal?	Yes No
13. Have you ever been charged for a sex offence against someone under the age of 12?	Yes No
14. Have you ever been charged for a sex offence against someone between the ages of 12 and 14?	Yes No
15. Have you ever been charged for a sex offence against someone aged 16 or older?	Yes No
16. Have you ever received any treatment for sex offending?	Yes No

*Appendix F*

## Consent Form for the Preliminary Interview

An informed consent form outlines the procedure, identifies any potential negative consequences, and provides contact information should you have any questions or concerns about the research that cannot be addressed by the experimenter. The main purpose of a consent form is to INFORM you of what you will be required to do so that you have adequate information to decide whether or not you wish to participate.

This research is being conducted by Kelly Babchishin and Nicolas Kessous under the supervision of Dr. Kevin Nunes (Assistant Professor, Department of Psychology, Carleton University, 613-520-2600, ext 1545; [kevin\\_nunes@carleton.ca](mailto:kevin_nunes@carleton.ca)). For questions or concerns about this study please contact Kevin Nunes. Should you have any ethical or any other concerns about this study, please contact Dr. Avi Parush, Chair of the Carleton University Ethics Committee for Psychological Research (613-520-2600, ext. 6026; [avi\\_parush@carleton.ca](mailto:avi_parush@carleton.ca)) or Dr. Anne Bowker, Chair of the Psychology Department (613-520-2600, ext. 2648; [psychchair@carleton.ca](mailto:psychchair@carleton.ca)).

The purpose of this research is to make a list of words related to things like children, adults, asexual attractiveness, and sexual offending. The words provided by you will be put together with words from other participants to create measures for use in later research. Other more general questions will be asked about your opinion of child sexual abuse and other topics. Your participation will involve speaking with the researcher for about one hour.

Information obtained will NOT be shared with Correctional Service Canada (CSC) staff and will NOT be put on any institutional file. You will not be identified as a participant in the publication or presentation of the results. Your name and other identifying details will not be written on information you provide.

Because some of the activities deal with sensitive material, such as reading descriptions of sexual acts and viewing pictures of naked people, they may be embarrassing or offensive. You are free to withdraw from the project at any time, refuse to participate, and refuse to answer questions. Refusal to participate will **NOT** affect your treatment by CSC in any way. Participation in this study will **NOT** affect any administrative decisions concerning you, such as your institutional placement or parole.

The data collected will be kept in a secure manner at Carleton University for a period of 10 years in a locked filing cabinet and password-protected computer, with any identifying information removed, and will be accessible only to the researchers working on this research.

*I have read and understood the information above. My signature indicates that I agree to participate in this study. There are two copies of the consent form, one of which I may keep.*

Participant Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Witness Signature: \_\_\_\_\_ Date: \_\_\_\_\_

*Appendix G*

## Preliminary Structure Interview Questions

**I) Words**

1. What does the word “sexy” make you think of?
2. What do the words “not sexy” make you think of?
3. What are some words you would use that mean the same thing as “sexually attractive”?
4. What are some words you would use that mean the same thing as “not sexually attractive”?
5. What are some words you would use that mean the same thing as “sexually arousing”?
6. What are some words you would use that mean the same thing as “not sexually arousing”?
7. What words come to mind when you think of an adult having sexual contact with a child?
8. What words come to mind when you think of an adult having sexual contact with another adult?
9. What kind of activities does the phrase “sex with a child” make you think of?
10. What kind of activities does the phrase “sex with an adult” make you think of?

*Appendix H*

## Consent Form for Rating Questionnaire

An informed consent form outlines the procedure, identifies any potential negative consequences, and provides contact information should you have any questions or concerns about the research that cannot be addressed by the experimenter. The main purpose of a consent form is to **INFORM** you of what you will be required to do so that you have enough information to decide whether or not you wish to participate.

This research is being conducted by Kelly Babchishin (613-520-2600, ext. 2649; [kbabchis@connect.carleton.ca](mailto:kbabchis@connect.carleton.ca)) and Nicolas Kessous (613-520-2600, ext. 2649; [nkessous@connect.carleton.ca](mailto:nkessous@connect.carleton.ca)) under the supervision of Dr. Kevin Nunes, Assistant Professor, Department of Psychology, Carleton University, (613-520-2600, ext. 1545; [kevin\\_nunes@carleton.ca](mailto:kevin_nunes@carleton.ca)). For questions or concerns about this study, please contact Kevin Nunes. Should you have any ethical or any other concerns about this study, please contact Dr. Avi Parush, Chair of the Carleton University Ethics Committee for Psychological Research (613- 520-2600, ext. 6026; [avi\\_parush@carleton.ca](mailto:avi_parush@carleton.ca)) or Dr. Anne Bowker, Chair of the Psychology Department (613-520-2600, ext. 2648; [psychchair@carleton.ca](mailto:psychchair@carleton.ca)).

The purpose of the research is to see how incarcerated men view words or pictures related to adults, children, sexual attractiveness, and sexual offending. You will be asked to fill out a questionnaire that will take about 15 minutes to complete. You will also be asked to answer a few questions about your conviction(s). Your answers combined with many other people's answers will be used to select words for use in future research.

Your answers are anonymous. That is, nobody will know how you answered because your name will not be written on your questionnaire. Your answers will **NOT** be shared with Correctional Service Canada (CSC) staff or put on any institutional file. You will **NOT** be identified as a participant in the publication or presentation of the results. You are free to withdraw from the project at any time, refuse to participate, and refuse to answer questions. Refusal to participate will **NOT** affect your treatment by CSC in any way. Participation in this study will **NOT** affect any administrative decisions concerning you, such as your institutional placement or parole.

The data collected will be kept in a secure manner at Carleton University for a period of 10 years in a locked filing cabinet and password-protected computer and will be accessible only to the researchers working on this research.

*I have read and understood the information above. My signature indicates that I agree to participate in this study. There are two copies of the consent form, one of which I may keep.*

Participant Signature: \_\_\_\_\_ Date: \_\_\_\_\_

## Appendix I

## Rating Questionnaire with Responses Included

## Instructions:

- On the following pages, you will be asked to categorize some words and pictures. There are no right or wrong answers.
- You will also be asked a few questions about your conviction(s).
- Your answers are anonymous. That is, nobody will know how you answered because your name will not be recorded.
- Please indicate if each word is “sexy” or “not sexy” by circling one of the options. For example, if you think “ass” is sexy, you would circle “sexy”. If you think “ass” is not sexy, you would circle “not sexy”.

Word	Rating	
1. ASS	Sexy 57.8% (n = 26)	Not Sexy 42.2% (n = 19)
2. KISS	Sexy 90.1% (n = 40)	Not Sexy 9.1% (n = 4)
3. STINK	Sexy 0% (n = 0)	Not Sexy 100% (n = 45)
4. AROUSING	Sexy 77.3% (n = 34)	Not Sexy 22.7% (n = 10)
5. UGLY	Sexy 2.2% (n = 1)	Not Sexy 97.8% (n = 44)
6. VOMIT	Sexy 2.2% (n = 1)	Not Sexy 97.8% (n = 44)
7. HORNY	Sexy 72.7% (n = 32)	Not Sexy 27.3% (n = 12)
8. ORGASM	Sexy 84.1% (n = 37)	Not Sexy 15.9% (n = 7)
9. IMPOTENT	Sexy 0% (n = 0)	Not Sexy 100% (n = 45)
10. YUCK	Sexy 0% (n = 0)	Not Sexy 100% (n = 45)
11. COLD	Sexy 2.2% (n = 1)	Not Sexy 97.8% (n = 44)
12. FUCK	Sexy 44.4% (n = 20)	Not Sexy 55.6% (n = 25)
13. SEX	Sexy 82.2% (n = 37)	Not Sexy 17.8% (n = 8)
14. ROTTEN	Sexy 2.3 % (n = 1)	Not Sexy 97.7% (n = 43)

15. BEAUTIFUL	Sexy 91.1% ( <i>n</i> = 41)	Not Sexy 8.9% ( <i>n</i> = 4)
16. LUST	Sexy 72.1% ( <i>n</i> = 31)	Not Sexy 27.9% ( <i>n</i> = 12)
17. ATTRACTIVE	Sexy 88.9% ( <i>n</i> = 40)	Not Sexy 11.1% ( <i>n</i> = 5)
18. BLOW JOB	Sexy 57.8% ( <i>n</i> = 26)	Not Sexy 42.2% ( <i>n</i> = 19)
19. GROSS	Sexy 4.4% ( <i>n</i> = 2)	Not Sexy 95.6% ( <i>n</i> = 43)
20. NAKED	Sexy 82.2% ( <i>n</i> = 37)	Not Sexy 17.8% ( <i>n</i> = 8)
21. LOVE	Sexy 91.1% ( <i>n</i> = 41)	Not Sexy 8.9% ( <i>n</i> = 4)
22. DISGUSTING	Sexy 0% ( <i>n</i> = 0)	Not Sexy 100% ( <i>n</i> = 45)
23. UNATTRACTIVE	Sexy 2.2% ( <i>n</i> = 1)	Not Sexy 97.8% ( <i>n</i> = 44)
24. SMILE	Sexy 88.9% ( <i>n</i> = 40)	Not Sexy 11.1% ( <i>n</i> = 5)
25. FRIGID	Sexy 9.1% ( <i>n</i> = 4)	Not Sexy 90.9% ( <i>n</i> = 40)

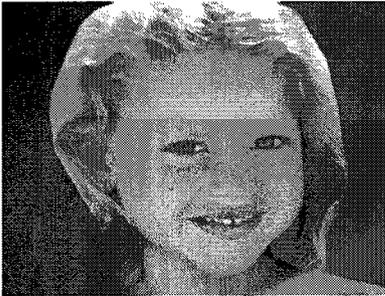
INSTRUCTIONS: Please indicate if each word is “good” or “bad” by circling one of the options. For example, if you think “ass” is good, you would circle “good”. If you think “ass” is bad, you would circle “bad”.

<i>Word</i>	<i>Rating</i>	
1. ASS	Good 57.1% ( <i>n</i> = 24)	Bad 42.9% ( <i>n</i> = 18)
2. KISS	Good 97.6% ( <i>n</i> = 40)	Bad 2.4% ( <i>n</i> = 1)
3. STINK	Good 2.4% ( <i>n</i> = 1)	Bad 97.6% ( <i>n</i> = 40)
4. AROUSING	Good 78.6% ( <i>n</i> = 33)	Bad 21.4% ( <i>n</i> = 9)
5. UGLY	Good 4.8% ( <i>n</i> = 2)	Bad 95.2% ( <i>n</i> = 40)
6. HORNY	Good 71.4% ( <i>n</i> = 30)	Bad 28.6% ( <i>n</i> = 12)
7. VOMIT	Good	Bad

	4.8% ( <i>n</i> = 2)	95.2% ( <i>n</i> = 40)
8. ORGASM	Good	Bad
	90.5% ( <i>n</i> = 38)	9.5% ( <i>n</i> = 4)
9. IMPOTENT	Good	Bad
	2.4% ( <i>n</i> = 1)	97.6% ( <i>n</i> = 40)
10. YUCK	Good	Bad
	0% ( <i>n</i> = 0)	100% ( <i>n</i> = 42)
11. COLD	Good	Bad
	7.1% ( <i>n</i> = 3)	92.9% ( <i>n</i> = 39)
12. FUCK	Good	Bad
	42.9% ( <i>n</i> = 18)	57.1% ( <i>n</i> = 24)
13. SEX	Good	Bad
	97.6% ( <i>n</i> = 40)	2.4% ( <i>n</i> = 1)
14. ROTTEN	Good	Bad
	2.4% ( <i>n</i> = 1)	97.6% ( <i>n</i> = 41)
15. BEAUTIFUL	Good	Bad
	97.6% ( <i>n</i> = 41)	2.4% ( <i>n</i> = 1)
16. LUST	Good	Bad
	61.9% ( <i>n</i> = 26)	38.1% ( <i>n</i> = 16)
17. ATTRACTIVE	Good	Bad
	97.6% ( <i>n</i> = 41)	2.4% ( <i>n</i> = 1)
18. BLOW JOB	Good	Bad
	52.4% ( <i>n</i> = 22)	47.6% ( <i>n</i> = 20)
19. GROSS	Good	Bad
	2.4% ( <i>n</i> = 1)	97.6% ( <i>n</i> = 40)
20. NAKED	Good	Bad
	85.7% ( <i>n</i> = 36)	14.3% ( <i>n</i> = 6)
21. LOVE	Good	Bad
	100% ( <i>n</i> = 42)	0% ( <i>n</i> = 0)
22. DISGUSTING	Good	Bad
	4.8% ( <i>n</i> = 2)	95.2% ( <i>n</i> = 40)
23. UNATTRACTIVE	Good	Bad
	9.5% ( <i>n</i> = 4)	90.5% ( <i>n</i> = 38)
24. SMILE	Good	Bad
	100% ( <i>n</i> = 42)	0% ( <i>n</i> = 0)
25. FRIGID	Good	Bad
	12.2% ( <i>n</i> = 4)	87.8% ( <i>n</i> = 36)

---

Please indicate if each picture fits better with “child” or “adult” by circling one of the options. For example, if you think the first picture looks like a child, you would circle “child”. If you think the picture looks like an adult, you would circle “adult”.

	Picture	Rating	
1.		Child 100% ( $n = 45$ )	Adult 0% ( $n = 0$ )
2.		Child 100% ( $n = 45$ )	Adult 0% ( $n = 0$ )
3.		Child 28.9% ( $n = 13$ )	Adult 71.1% ( $n = 32$ )
4.		Child 100% ( $n = 45$ )	Adult 0% ( $n = 0$ )
5.		Child 0% ( $n = 0$ )	Adult 100% ( $n = 45$ )

6.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

7.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

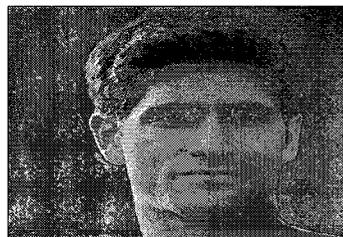
8.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

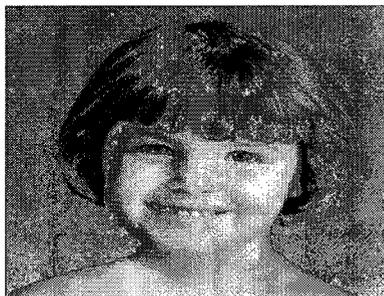
9.



Child  
0% ( $n = 0$ )

Adult  
100% ( $n = 45$ )

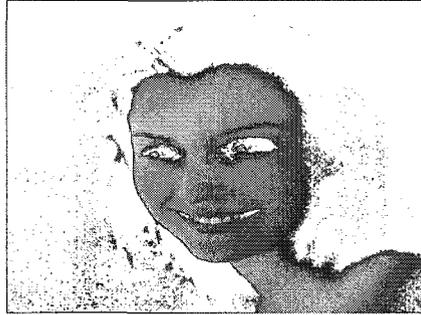
10.



Child  
100% ( $n = 44$ )

Adult  
0% ( $n = 0$ )

11.



Child  
4.5% ( $n = 2$ )

Adult  
95.5% ( $n = 42$ )

12.



Child  
100% ( $n = 44$ )

Adult  
0% ( $n = 0$ )

13.



Child  
100% ( $n = 44$ )

Adult  
0% ( $n = 0$ )

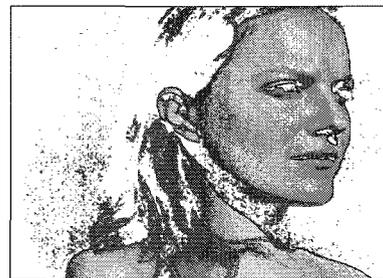
14.



Child  
0% ( $n = 0$ )

Adult  
100% ( $n = 45$ )

15.



Child  
0% ( $n = 0$ )

Adult  
100% ( $n = 44$ )

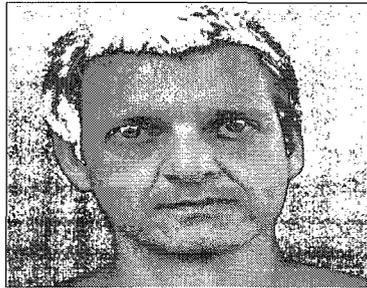
16.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

17.



Child  
0% ( $n = 0$ )

Adult  
100% ( $n = 45$ )

18.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

19.



Child  
4.4% ( $n = 2$ )

Adult  
95.6% ( $n = 43$ )

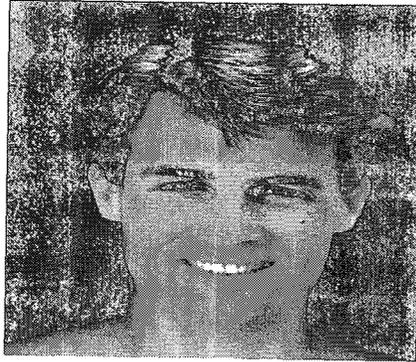
20.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

21.



Child  
2.2 % ( $n = 1$ )

Adult  
97.8% ( $n = 44$ )

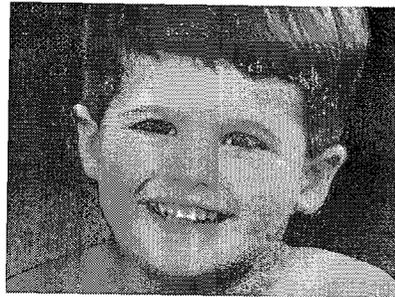
22.



Child  
0% ( $n = 0$ )

Adult  
100% ( $n = 45$ )

23.



Child  
0% ( $n = 45$ )

Adult  
100% ( $n = 0$ )

24.



Child  
8.9% ( $n = 4$ )

Adult  
91.1% ( $n = 41$ )

25.



Child  
0% ( $n = 44$ )

Adult  
100% ( $n = 0$ )

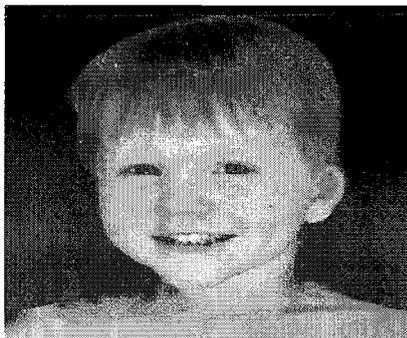
26.



Child  
100% ( $n = 44$ )

Adult  
0% ( $n = 0$ )

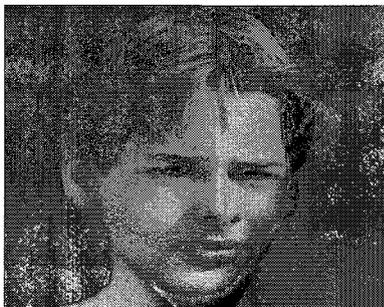
27.



Child  
100% ( $n = 44$ )

Adult  
0% ( $n = 0$ )

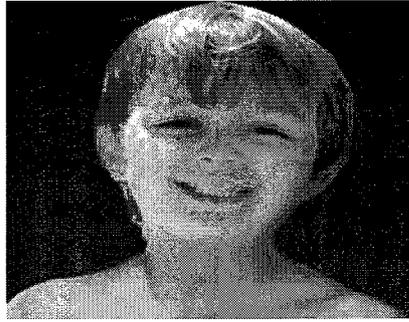
28.



Child  
77.3% ( $n = 34$ )

Adult  
22.7% ( $n = 10$ )

29.



Child  
100% ( $n = 45$ )

Adult  
0% ( $n = 0$ )

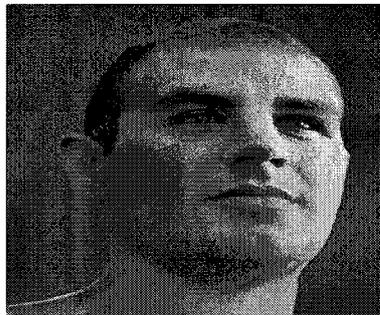
30.



Child  
4.4% ( $n = 2$ )

Adult  
95.6% ( $n = 43$ )

31.



Child  
4.4% ( $n = 2$ )

Adult  
95.6% ( $n = 43$ )

32.



Child  
97.8% ( $n = 44$ )

Adult  
2.2% ( $n = 1$ )

33.



Child  
8.9% (*n* = 4)

Adult  
91.1% (*n* = 41)

Please circle your answer.

Have you ever been convicted for a sexual offence against:

A person less than 12 years of age?	YES 47.6% ( <i>n</i> = 20)	NO 52.4% ( <i>n</i> = 22)
A person between the ages of 12 and 15 years of age?	YES 40.5% ( <i>n</i> = 17)	NO 59.5% ( <i>n</i> = 25)
A person 16 years of age or older?	YES 56.1% ( <i>n</i> = 23)	NO 43.9% ( <i>n</i> = 18)

You have completed the questionnaire.

Thank you very much!

*Appendix J*

## Consent Form for Pilot Test and Study Proper

An informed consent form outlines the procedure, identifies any potential negative consequences, and provides contact information should you have any questions or concerns about the research that cannot be addressed by the experimenter. The main purpose of a consent form is to **INFORM** you of what you will be required to do so that you have adequate information to decide whether or not you wish to participate.

-----

This research is being conducted by Kelly Babchishin and Nicolas Kessous under the supervision of Dr. Kevin Nunes (Assistant Professor, Department of Psychology, Carleton University, 613-520-2600, ext 1545; [kevin\\_nunes@carleton.ca](mailto:kevin_nunes@carleton.ca)). For questions or concerns about this study please contact Kevin Nunes. Should you have any ethical or any other concerns about this study, please contact Dr. Avi Parush, Chair of the Carleton University Ethics Committee for Psychological Research (613-520-2600, ext. 6026; [avi\\_parush@carleton.ca](mailto:avi_parush@carleton.ca)) or Dr. Anne Bowker, Chair of the Psychology Department (613-520-2600, ext. 2648; [psychchair@carleton.ca](mailto:psychchair@carleton.ca)).

The purpose of the research is to see if incarcerated men who have not committed sex offences are different from men who have committed sex offences in the way they think about children and sex offending. The researchers need information from men who have not committed sex offences as well as those who have been arrested for sex offences.

Your participation will consist of one session, lasting approximately 1 hour and 30 minutes. You will be asked to complete some categorization tasks and questionnaires on a computer. For one of these questionnaires you will be asked to rate how attractive you find a wide range of sexual acts, some of them involving children and violence. You will also be asked to view pictures of naked women, men, girls, and boys, and rate the attractiveness of these pictures.

You are asked to consent to the disclosure of specific information from your institutional files, such as your criminal record, intake assessment, and psychological assessment information. This information will be used by the researchers to see if there is a relationship between specific file information and the tasks completed in the present study. Your file information will be accessed for the current study and for follow-up studies up to a maximum of 10 years from now.

The contents of your files will be used only for research purposes and your confidentiality will be respected and protected. The information you provide will **NOT** be shared with Correctional Service Canada (CSC) staff or put on any institutional file. You will not be identified as a participant in the publication or presentation of the results.

Your name, FPS number, and other personal information will **NOT** be written on the information you provide to protect your confidentiality.

Because some of the activities deal with sensitive material, such as reading descriptions of sexual acts and viewing pictures of naked people, they may be embarrassing or offensive. You are free to withdraw from the project at any time, refuse to participate, and refuse to answer questions. Refusal to participate will **NOT** affect your treatment by CSC in any way. Participation in this study will **NOT** affect any administrative decisions concerning you, such as your institutional placement or parole.

The data collected will be kept in a secure manner at Carleton University for a period of 10 years in a locked filing cabinet and password-protected computer, with any identifying information removed, and will be accessible only to the researchers working on this research.

*I have read and understood the information above. My signature indicates that I agree to participate in this study. There are two copies of the consent form, one of which I may keep.*

Participant Signature: \_\_\_\_\_

Date: \_\_\_\_\_

Witness Signature: \_\_\_\_\_

Date: \_\_\_\_\_

## Appendix K

## Order of Presentation of Measures

Sequences	Blocks	Measures	Within-Block Order	Within-Sequence Order	
1	1	SAC-IAT	Counterbalanced	Counterbalanced	
		CSA-IAT <sup>a</sup>			
	2	Evaluation of Child Sexual Abuse Scale <sup>a</sup>	Counterbalanced		
		Sex With Children Scale <sup>a</sup>			
	3	Balanced Inventory of Desirable Responding	Counterbalanced		Sequential
		Background Questionnaire			
2	4	Clarke Vocabulary Scale	Sequential	Sequential	
	5	Sexual Interest Profiling System	Counterbalanced		
		Viewing Time			
	6	Debriefing	Sequential		

*Note.* SAC-IAT: Sexual Attraction to Children IAT. CSA-IAT: Attitudes toward Child Sexual Abuse IAT. <sup>a</sup> Measure is not used in the current study. Two sequences represented the two computer files that will be administered to participants during the testing session.

*Appendix L*  
Debriefing Script

The following script was showed to participants via the laptop computer at the end of the testing procedures. Each paragraph was shown separately, with the presentation of the next paragraph following a ten second delay. The debriefing therefore lasted approximately 2 minutes.

Thank you very much for taking part in the study. This research would not be possible without your participation. I hope the following information addresses any questions and concerns you may have.

What Am I Trying to Learn in this Research?

I'm doing this research for my Master's thesis. The purpose of my research is to develop measures of sexual interest and attitudes related to sexual offending against children. One of those measures is called the Implicit Association Tests (IAT) - this was the one where you categorized words and pictures.

We are trying to find out if men who have been convicted of sex offences against children are (a) more attracted to children and (b) view offending against children more positively than men who have not been convicted of such offences.

Why Is This Research Important?

Finding out more about these sexual interests and attitudes can help us to better understand the motivation behind these offences and, ultimately, how to reduce them through assessment and treatment.

What if I Have Questions or Concerns?

Please direct any questions or concerns about this research to the people listed on your copy of the consent form.

If you experience any distress as a result of this study, please seek help from staff in Psychology.

Thank you again for taking the time to participate in my study.

*Appendix M*

## A Brief Description of the D Algorithm

The D algorithm uses only the data from block 3 (practice combination of child and sexy), block 4 (combination of child and sexy), block 6 (practice combination of adult and sexy), and block 7 (combination of adult and sexy).

Table L1

*D algorithm outlined in Greenwald et al. (2003)*

Step	Procedure
1.	Remove extreme trials, which are defined as trials in which the participant took more than 10,000 milliseconds to sort the stimulus into the appropriate category.
2.	Participants with more than 10% of trials with reaction times below 300 milliseconds are removed from the analysis as these participants are likely not attempting to sort the stimuli accurately.
3.	The correct mean latency is then calculated for each block. <ul style="list-style-type: none"> <li>• If participants did not make any mistakes then the correct mean latency is simply the mean reaction of the particular block (e.g., when child and sexy were paired).</li> <li>• If error(s) were present, then this step involves adding an error penalty of 600ms.</li> </ul>
4.	Two difference scores are calculated. <ul style="list-style-type: none"> <li>• The first by subtracting the mean latency of block 3 from block 6.</li> <li>• The second by subtracting the mean latency of block 4 from block 7.</li> </ul>
5.	The difference score for blocks 3 and 6 is then divided by the pooled standard deviation for all trials in block 3 and block 6. The difference score for blocks 4 and 7 is also divided by the pooled standard deviation from block 4 and block 7.
6.	The average of two scores derived from step 5 represents the IAT effect.

Appendix N

Internal Consistency Analyses

a) Internal consistency method described in Nosek, Greenwald, and Banaji (2005):

- Calculate a *D* coefficients for the combination of child and sexy (block 3 and block 4) and for the combination of adult and sexy (block 6 and block 7).
  - The calculation of a *D* score involves several steps that were explained in detail in Appendix L.
  
- Correlated the two *D* coefficients for an estimate of internal consistency.

