



All Theses and Dissertations

2015-06-01

Predictive Validity of the LOOK

Joy Wiechmann Cox
Brigham Young University - Provo

Follow this and additional works at: <http://scholarsarchive.byu.edu/etd>

 Part of the [Counseling Psychology Commons](#), and the [Special Education and Teaching Commons](#)

BYU ScholarsArchive Citation

Cox, Joy Wiechmann, "Predictive Validity of the LOOK" (2015). *All Theses and Dissertations*. Paper 5485.

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu.

Predictive Validity of the LOOK

Joy Wiechmann Cox

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Lane Fischer, Chair
G. E. Kawika Allen
Melissa Allen Heath
Aaron P. Jackson
Marleen Williams

Department of Counseling Psychology and Special Education

Brigham Young University

June 2015

Copyright © 2015 Joy Wiechmann Cox

All Rights Reserved

ABSTRACT

Predictive Validity of the LOOK

Joy Wiechmann Cox
Department of Counseling Psychology and Special Education, BYU
Doctor of Philosophy

The LOOK, an iOS app, is a viewing time measure used to assess sexual interest. The measure is based on the assumption that sexual interest can be assessed by the amount of time a participant spends looking at an image. The purpose of this study was to examine the ability of the LOOK, a newly developed viewing time instrument, to accurately screen and diagnose individuals with deviant sexual interest. The profiles of known sexual offenders were compared to norm-referenced profiles of an exclusively heterosexual, non-pedophilic, male, college student population. Researchers were not able to find a fair constant multiplier that would allow for a positive screen of our offender sample while not over identifying our non-offender sample. Instead a graph was generated which showed the trends of offenders were closely related to those of non-offenders using Fischer's Chi-square model. Additionally, when looking at the predictive validity of being able to identify victim demographics of known perpetrators based on Fischer's Chi Square residuals, only 15.9% were found to have offense histories that were consistent with their profiles on the LOOK. The LOOK, using Fischer's Chi-square model does not seem to be able to differentiate offenders from non-offenders. Future studies may include looking at the predictive nature of ipsative data.

Keywords: LOOK, sexual interest, viewing time, ipsative, chi-square, Lane Fischer

ACKNOWLEDGMENTS

I would like to thank Lane Fischer, my chair, for his unwavering support throughout my time in graduate school. He has been much more than a dissertation advisor; he is a dear friend, a cheerleader of all good things, and a listening, reassuring ear. I would also like to thank my committee members G. E. Kawika Allen, Melissa Allen Heath, Aaron P. Jackson, and Marleen Williams for their support in working on this study and my Affinity/LOOK team members who forged the path for this research to take place.

I am grateful to the Brigham Young University Graduate School for awarding me a Graduate Research Fellowship, which enabled me to fund my research. And for Mervyn Davies and his office manager, Tarita Propp, for sharing their data with us and helping us to meet the IRB requirement of the data being de-identified on site by staff members.

I will be forever grateful for the support my family has given to me throughout my life in all I have set out to do. And I am grateful for my sweet husband, Ryan, whom I met the same month I started my Ph.D. and whom has been supportive of my goals and aspirations. I love you so much.

TABLE OF CONTENTS

TITLE PAGE	i
ABSTRACT.....	ii
ACKNOWLEDGMENTS	iii
LIST OF TABLES.....	vi
LIST OF FIGURES	vii
DESCRIPTION OF DISSERTATION STRUCTURE	viii
Introduction.....	1
Theories on Sexual Abuse	2
Types of Assessment Used in the Field	3
Ipsative Measures.....	6
Predictive Validity	8
Current Study	9
Research Questions.....	10
Method	10
Procedures.....	10
Data Analysis.....	13
Results.....	14
Discussion.....	20
Questions and Limitations	20

Study Limitations.....	22
Future Directions	24
APPENDIX A: Literature Review.....	30
Recidivism	30
How Results from These Assessments Are Used.....	31
Theories of Sexual Abuse and Victim Choice.....	33
Types of Assessments Used.....	35
Self-Report/Clinical Interview Measures	36
Plethysmography.....	37
Viewing Time Measures.....	41
Neuroimaging	48
References.....	50

LIST OF TABLES

Table 1: <i>Participant Demographics and Offense Statistics</i>	16
---	----

LIST OF FIGURES

<i>Figure 1.</i> Positive identification of deviance of non-offending and offending samples based on Fischer's Chi-square model using a variety of fair constant multipliers.....	18
<i>Figure 2.</i> Graph of all participants Chi-square residuals from the offender sample (n=44).....	19
<i>Figure 3.</i> Graph of all participants Chi-square residuals from the non-offender sample (n=69).....	19

DESCRIPTION OF DISSERTATION STRUCTURE

This dissertation, *Predictive Validity of the LOOK*, is formatted differently than most dissertations as it is written in a hybrid format. The hybrid format allows for elements of both traditional dissertations and journal-ready formats to be combined. This dissertation is broken up in two parts, instead of having sections divided by chapters. The first part, the dissertation study, is written in a journal-ready format that is approximately the same length and style of studies published in most psychological journals. The second part includes an expanded literature review. Each part has a list of references included in that section. The preliminary pages are consistent with the traditional dissertation format.

Introduction

Sexual offending is a high frequency crime, which causes psychological damage to many innocent people who have been offended against. Based on The National Center for Victims of Crime report for 2013, 28% of the U.S. population of 14- to 17-year-olds reported that they had been sexually victimized at some point in their lifetime. In 2008 alone, 4.8% of males and 7.4% of females aged zero to 17 were sexually victimized (The National Center for Victims of Crime, 2013). As such, many researchers and practitioners have developed measures in an effort to better understand sexual abuse, the perpetrators of abuse, and to help make crucial decisions about their treatment and future risk to the community (Marshall & Fernandez, 2000).

Over the course of the last two decades, the use of structured risk assessments to predict recidivism has become standard procedure in evaluating sex offenders. Many hope to identify those perpetrators at risk for re-offense and to identify factors that may contribute to an individual's likelihood to reoffend. In response there has been new legislation passed in recent decades that includes longer prison sentences and mandatory risk assessments before allowing inmates to be released from prison. Twenty states' laws and many federal statutes include one particularly restrictive law, which allows sex offenders to be committed to a secured forensic hospital for an indefinite period of time. These individuals have been given the label Sexually Violent Persons (SVP) and have been identified as having a mental abnormality or other disorder, which contributes to the likelihood that they will reoffend.

While this type of assessment is becoming more widely accepted in the field, some of the research does not support recidivism risk assessment methods as being particularly accurate. Caldwell (2013) conducted one such study in reviewing the records of 198 juveniles who had committed sexually violent offenses that qualified them for possible commitment under the SVP

civil commitment law. As part of their evaluation, each youth was screened by at least two expert examiners. Of the 198 juveniles evaluated, 54 were found to meet criteria for an SVP petition. The remaining 144 were screened out. Researchers then collected follow-up data on subsequent criminal charges for an average of nearly five years on both those who qualified for SVP petitions and those who had been screened out of additional evaluation. The results of their study did not show significant differences between youth who had been screened out and those subject to petition in terms of prevalence rates for general sexual offending and felony sexual offending. Among petitioned youth, 11.76% were charged with a new sexual offense during the follow-up period including 9.80% who were charged with a felony sexual assault. Of the non-petitioned youth, 17.36% were charged with a sexual offense including 13.19% who were charged with a felony sexual assault.

Individual liberties are greatly impacted by these types of evaluations; more accurate methods to evaluate risk are needed. There is still much work to be done in the field to assess if these instruments are sensitive enough to identify individuals who pose a threat to the community.

Theories on Sexual Abuse

Most, if not all, of the measures used in the assessment process are based on the construct that sexual abuse is driven by sexual desire. While Singer's (1984) research would support this hypothesis, others in the field believe there are a variety of motivating factors behind sexual abuse besides the feelings of sexual attraction to the victim. Marshall (1996) stated that sexual abuse is not about sexual attraction but rather about issues around power and control, desire to humiliate, or an expression of aggression rather than stemming from sexual fantasy. Quinsey, Rice, Harris, and Reid (1993) have postulated that those with deviant sexual preferences will

engage in specific types of behavior because they are motivated to do so. However, they emphasize that sexual behaviors are not always an accurate reflection of sexual preference but may be more related to acting within social constraints and/or opportunities presented to them. Other studies have suggested that sex offenders may be more polymorphously perverse than focused on one stimulus category. Abel and Rouleau (1990) and Freund (1990) conducted studies in which they found that outpatient sex offenders had a high incident of multiple paraphilias compared to the general population.

Types of Assessment Used in the Field

Numerous approaches to evaluation are used in the assessment of sex offenders. These include the use of polygraphs, psychometrics, historical partner choice/clinical interview, pupillometry, self-report interviews, gaze tracking, stimulus rating, plethysmography, viewing time, and neuroimaging. Clinical interviews and self-report questionnaires aside, most methods for assessing pedophilic sexual preferences aim for a relatively objective measure of sexual interest or arousal. Despite attempting to develop objective measures, Glasgow, Osborne, and Croxen (2003) stated that he believed all methods of assessment have relative strengths and weaknesses.

Self-report measures and clinical interviews are often used in combination with other measures. The utilization of self-report measures is useful in gaining access to information that is not able to be assessed for in any other way. However, these measures are also prone to deception, impression management, and defensiveness as one must rely solely on the report of the offender (Grese, 2005); offenders are known to deny or minimize the abuse they inflicted. Obtaining access to victim statements, trial transcripts, and police reports may be necessary to

get an accurate picture of the abuse perpetrated (Marshall, 1996). As such, measures have been developed to try to gain a more objective understanding of individuals' sexual preferences.

Plethysmography, first used by Freund in 1957 to test sexual arousal in males, has been described as the most sensitive and reliable indicator of sexual attraction, though some have questioned whether plethysmography is a measure of sexual arousal or sexual attraction (Marshall & Fernandez, 2000). It has since gained popularity in the study of sexual offenders and is frequently used in psycho-sexual assessments conducted in the prison system. However, Quinsey and Chaplin (1988) reported that many studies have shown that men are able to enhance and inhibit their erectile responses using a number of faking strategies. When encouraged to do so, both men and women have demonstrated an ability to suppress their plethysmographically-assessed arousal (Israel & Strassberg, 2007). According to Abel et al. (2004) test subjects have disclosed falsifying their test results by not looking at the stimuli, only listening to a portion of the stimuli, or by using imagery to control their responses. Individuals have also been known to use a technique called "pumping" or the contracting of their perineal muscles to increase the appearance of arousal to socially appropriate stimuli. Other strategies reported in the literature include using cognitive processes to distract themselves from the erotic stimuli; one such strategy includes thinking about a difficult math problem.

Marshall and Fernandez (2000) suggest that because there is no way to prevent all methods of deception, faking will always serve as a threat to the validity of phallometric assessments. Faking is not the only concern researchers have when measuring vaginal photoplethysmography (VPP) or penile plethysmography (PPG) responses, however. Results of these assessments have also been shown to be affected by lack of sleep, illness, or recent sexual activity. Additionally, ethical concerns are a problem in conducting plethysmographic research.

Plethysmography is a very invasive procedure and those that voluntarily take these assessments may not generate a normative sample of the population.

A less invasive method currently in use in the field is viewing time measures. The relationship between sexual attraction and time spent looking at images was initially described by Rosenzweig in 1942. Rosenzweig (1942) found that differences in viewing time of sexual images discriminated level of sexual interest. In 1956, Zamansky continued with this line of study. He presented men with pairs of images, with one picture in each pair being of a woman and one of a man. Based on how long the men spent viewing the pictures, Zamansky was able to distinguish between those with homosexual and heterosexual preferences (Zamansky, 1956).

The field of sexuality research has suggested that viewing time is a valid measure for understanding the sexual interest of those tested, however, few studies have established psychometric validity. Evidence for what Israel and Strassberg (2009) call “perceived validity” is the number of viewing time instruments that have been developed and made commercially available in the assessment of sex offenders. More research is needed, however, to know if these tests are accurate enough to make high stake decisions based on their results.

Faking is a concern in using viewing time measures. In a presentation at the Association for the Treatment of Sexual Abusers (ATSA) annual conference, Fischer, Baird, Hansen, Stephenson, and Veas-Wall (2012) presented a faking study in which they found that using a cognitive strategy such as pretending to be a different gender with different sexual preferences, or knowing that the test is actually measuring viewing time, can enable someone to fake the Affinity 2.5. These results suggest that sex offenders may be able to fake the assessment to look more socially acceptable in their sexual preferences.

One strategy that researchers are using to make these assessments harder to fake is the use of a dot probe. Requiring participants to attend to another stimulus seems to inhibit one's ability to use cognitive strategies to fake a test. Wright and Adams (1994) reported that a significant increase in latency occurred when individuals were shown pictures of subjects who matched their sexual preference. They hypothesized that when using a dot probe, the subjects were required to attend more to the stimuli, which interfered with their ability to use cognitive strategies, and as a result spent a longer time looking over the images that matched their sexual preference category.

Ipsative Measures

Interpreting the data correctly is of particular concern in the use of viewing time measurements. Most psychometric devices use norm-referenced data rather than ipsative data. As such many practitioners are less familiar with ipsative measures and incorrectly interpret the results from such measures. Ipsative measures only suggest one attribute is more salient than another attribute within an individual. Attempts have been made to use ipsative data to represent an ideal pattern of responding and to identify sexual deviance. However in order to use ipsative data in this manner, established norms are needed. Fischer and Meade (2010) stated that when using ipsative data, statements of normalcy or deviance are not supported in absence of a reference group.

The Affinity's viewing time measurements are reported in raw scores in either seconds or milliseconds then converted to mean rank scores (Glasgow et al., 2003). Images are then organized in terms of viewing time from longest to shortest and assigned a numerical value from one to 80. The rank scores are then averaged to provide a mean rank for each stimulus category. Viewing the data in this fashion allows a computer program to create a graphic representation of

each person's profile that is consistent from one participant to another in terms of format. While this is standard practice when using an ipsative measure, information about the individual may be lost in the representation. Data obtained from the Affinity's ipsative measures are considered intra-personal showing individual preference of one stimulus over another but nothing more.

The Abel Assessment for Sexual Interest (AASI), notably the most popular viewing time assessment, uses a different method for obtaining their results. Viewing time is not reported in ipsative form but in the form of z-scores that create a distribution of scores with a calculated mean and standard deviation. This procedure transforms the scores into a standard normal deviate of the distribution (Barbaree & Mewhort, 1994). The raw score data are then sent to Abel Screening and the results are returned to test administrators after statistical analysis has trimmed the data and outliers have been removed by an undisclosed formula. This procedure is problematic in that it does not allow for the computation of a coefficient alpha; this prohibits the evaluation of internal consistency. The removal of outliers is also problematic as they may be removing those who scored in the most deviant ranges. Fischer and Smith (1999) reported that because raw score means and standard deviation means are not provided, it could easily lead to the misinterpretation of data as being norm referenced and to clinicians concluding that the data is in interval form rather than ordinal form.

Fischer (Fischer, 2004; Fischer & Morgan, 2006) has postulated that the logic of the Chi-square goodness-of-fit test can be used to compare an individual's ipsative scores to an expected pattern of scores. Chi-square goodness-of-fit allows for the comparison of an observed pattern of responses to an expected pattern of responses and estimates the overall fit of the observed individual pattern to the expected pattern. In looking at Chi-square residuals, one can estimate which specific categories contribute to the overall Chi-square. Fischer (Fischer, 2004; Fischer,

Byrne & Glasgow, 2007; Fischer & Morgan, 2006) has developed and empirically tested a Chi-square application that uses ipsative patterns in a reference group procedure. Chi-square analyses have the following equation:

$$\chi^2_{(df)} = \sum_{i=1}^J \frac{(O_i - E_i)^2}{E_i}, \quad (1)$$

In this formula the differences of expected counts (E) and observed counts (O) both within each category are tested against chance. Fischer adapted this formula as follows:

$$\chi^2 = n * \sum_{i=1}^J \frac{(P_i - \pi_i)^2}{\pi_i}, \quad (2)$$

Fischer's Chi-square approach uses the observed proportion of time viewed for each category (P) and the expected proportion of time viewed for each category (π). Additionally, Fischer added n as a fair constant multiplier to help increase sensitivity and specificity of score variance.

Fischer's Chi-square approach seems to have overcome many of the difficulties associated with the use of ipsative scores in current clinical practice. The logic behind the Chi-square approach is based on a very different assumption than most instruments. Rather than looking for a typical deviant pattern, it assumes there are many ways in which a participant may have deviant sexual interest and instead looks to identify a typical expected pattern among non-pedophilic individuals. Using this expected pattern, one can identify those who are deviant from the typical pattern.

Predictive Validity

Stephenson (2014) conducted a predictive validity study using the Affinity 2.5. She collected archival data from an outpatient clinic in Colorado where 287 individuals undergoing psychosexual evaluations for sexual offenses were assessed using a variety of measures, including the Affinity 2.5. The results of these assessments were used in recommendations for

sentencing; this sample included the assessment results of 77 participants who were being re-evaluated for re-offense, five of whom were charged with a sexual re-offense.

Results of the study did not show that using Fischer's Chi-square statistical method on data from the Affinity was able to identify offenders as having deviant sexual interest at a rate better than chance (43.7%). Additionally, in using Fischer's Chi-square residuals as a diagnostic tool, this method only correctly identified 22.8% of study participants as having deviant interest in protected populations. Of the 22.8% of individuals having deviant residuals, only 12.0% of the 22.8% showed residual deviant interest that matched the demographic information of their victims. Stephenson concluded, "It seems that Affinity 2.5 and Fischer's Chi-square do not have sufficient discrimination between those who sexually offend against children and adolescents and those who do not."

The results of the Affinity 2.5 were reported in milliseconds. As such, a fair constant multiplier of 115 was used to make the numbers larger and easier to compare.

Current Study

The purpose of this study was to examine the ability of the LOOK, a newly developed viewing time instrument, to accurately screen and diagnose individuals with deviant sexual interest. In trying to find a fair constant multiplier that would differentiate the offenders from the non-offenders, it became apparent that finding a fair constant that identified offenders but did not over identify non-offenders was impossible. Thus, we had to change our research question to one of the screening efficacy of the LOOK and determining where the greatest differentiation occurred based on changing the fair constant multiplier. Additionally, Fischer's Chi-square scoring procedure was used in comparing the viewing-time profiles of individuals who have been convicted of a sex crime against a child to an expected profile. The expected profiles are

based on samples of exclusively heterosexual, non-pedophilic, college-age males. Individuals included in this sample were recruited using the SONA Research System, an experiment management system managed by the Department of Psychology at Brigham Young University (BYU) for a previous study.

Research Questions

1. What is the screening efficacy of using Fischer's Chi-square on a sample of offenders and non-offenders? Using a variety of fair constant multipliers where does the biggest differentiation happen between the offending and non-offending samples?
2. What percentage of convicted sexual offenders' offense histories is consistent with the diagnostic indicators on the LOOK using Fischer's Chi-square scoring procedure?

Method

Participants in this study included individuals undergoing psychosexual evaluation or re-evaluation from a private, outpatient clinic in Colorado after facing criminal charges of a sexual nature. Psychosexual evaluations were requested by the court or other interested parties to assist in making sentencing decisions. All participants were males 18 years of age or older. Additionally, a sample of heterosexual, non-pedophilic, college students served as our comparison group; the group was previously studied by the LOOK team to establish norms.

Procedures

The data collected for this study came from an existing data set collected by a practitioner in Colorado who has contracted with a state agency to do a large number of psychosexual evaluations. The LOOK was administered to individuals as part of their standard assessment battery. Demographic information from the psychosexual evaluations including gender, age at time of offense, age at time of testing, nature of offense, reason for current assessment, non-

sexual re-offense history, and any known victim demographics (age, gender, and number of victims) has been used by researchers in their data analysis. This data was de-identified and coded on site by a staff member of the private practice. The data from the LOOK was uploaded automatically to the LOOK server then distributed to researchers. This data did not contain any identifying information.

For the purpose of this analysis and for the prospect of establishing concurrent validity between the LOOK and the Affinity 2.5 in a future study, the definitions of offense categories used in this study are identical to the categories used by Stephenson (2014) in her study of the predictive validity of the Affinity 2.5. Offenses were categorized into five groups based on Stephenson's research: Incest, Child Pornography, Rape, Molestation, and Indecent Exposure. She defined these categories as follows.

Incest was defined as any form of parent-child or sibling-sibling relationship such as father/daughter, step-father/step-daughter, brother/sister, or mother/son. Offenses where the victim and offender were related but not in a parent-child or sibling relationship were not categorized as incest (e.g., cousins, uncle/niece). This distinction was made on the assumption that parent-child and sibling relationships are characterized by imputed trust and a power imbalance (Lesniak, Rudman, Rector, & David Elkin, 2006) whereas this assumption is harder to make in extended family situations.

Child Pornography included any offense where the offender had possession of nude or suggestive images of children and adolescents under the age of 18. This included possession of videos and images obtained surreptitiously, or text-based images from consenting minors, but did not include sending images of adults to minors.

Rape was defined as any offense that included consensual and illegal or nonconsensual penetration by sexual organs or objects. This included vaginal, anal, and oral penetration, but did not include digital penetration alone, which was categorized as molestation.

Molestation was defined as any unwanted or consensual but illegal touch that did not include penetration, except digital penetration as mentioned above. This included fondling genitalia, breasts, anus, or any other body part in a sexually suggestive way. Molestation also incorporated masturbation of the victim, the offender, or both.

Indecent Exposure included those offenses where the offender exposed him or herself to others in a sexual way without any physical touch. This included public masturbation and exhibitionism (Stephenson, 2014).

Currently there are two viewing time instruments on the market used to assess sexual interest. One is the Abel Assessment for Sexual Interest (AASI), and the other is the Affinity. A team of researchers are currently conducting or have in previous years conducted a number of research projects on the LOOK to establish norms, study the ability of individuals to fake their results, and establish the reliability of this new assessment.

The LOOK, developed by Cameron and Sierra Baird, uses touch screen technology in an iOS app available on the iPad Mini to gather test information from subjects. The LOOK begins with a priming task showing fourteen different outlines of both male and female body types of various ages. Participants are asked to select which body type they found to be most sexually attractive and then rate the rest of the figures from most sexually attractive to least sexually attractive. After the priming task is completed, participants are shown a series of 140 images of individuals that fall into fourteen categories and prompted to rate these individuals as sexually attractive or unattractive to them. The fourteen categories include elderly female (ELF), mature

adult female (MAF), adult female (ADF), juvenile female (JUF), pre-juvenile female (PJF), small child female (SCF), infant female (INF), elderly male (ELM), mature adult male (MAM), adult male (ADM), juvenile male (JUM), pre-juvenile male (PJM), small child male (SCM), and infant male (INM). The assessment is comprised of eleven images from each group; the first image from each category is thrown out as individuals get accustomed to the format of the test. Participants first must find a dot probe located on the screen in one of the four corners of the image. The dot randomly changes places from one picture to another. After locating the dot participants are asked to rate the images, which also appear in random order, on the seven point Likert scale from Very Attractive to Very Unattractive with a Neutral midpoint. After rating the image, the next image appears on the screen.

The images used in the LOOK were purchased from Shutterstock, an online stock photo company, and care was taken to screen the photos. Photos were chosen in which full body, clothed individuals were shown. Care was taken so that the image subjects were not in sexually provocative positions or making sexually expressive facial expressions. Participants' interaction with the LOOK program are recorded in microseconds from the time the image appears to when the dot is selected and from when the dot is selected to when the image is rated. The program also generates a total time score and the subjects' Likert score ratings are recorded. Data collected by the LOOK is retained in its raw state.

Data Analysis

Raw viewing time data from the LOOK was converted to difference scores by comparing the percentage of total time spent in each pre-defined category by each individual to the average percentage of time spent in each category by a non-offending, exclusively heterosexual, male comparison group. Patterns of differences were assessed for significance using Fischer's Chi-

square approach to scoring then a range of fair constant multipliers was tested to observe changes in positive identification of deviance. Statistically significant scores, those above 22.2, indicated significant deviance of the overall viewing-time pattern compared to the expected pattern. A graph showing the trend of positive identification for both the offender sample and the control sample were graphed. Additionally, we located the areas of greatest differentiation between the two samples based on the fair constant multiplier.

To answer the second research question, Fischer's Chi-square standardized residual was used. This standardizing procedure allows one to compare residuals across categories and examine whether these residuals will be consonant with the offense history of the sex offenders in the study.

Results

Researchers obtained the results of 82 potential participants who had previously been given the LOOK during a psychosexual evaluation. Of these participants, 38 were excluded for the following reasons: two participants did not have offense histories reported to researchers, 16 had only adult victims, seven were under the age of 18 when their crime(s) was committed, two were female, and an additional 12 had sexual offense histories that were 10 years old or more. The remaining sample of 44 participants had a mean age of 29.3 years old.

Of the 44 participants, 42 had convictions in only one offense category while two had convictions in two offense categories. Molestation was the most common offense with 47.7% (n=21) of participants having criminal charges in this area. Rape was the second most common offense with 34.1% (n=15) followed by incest at 11.4% (n=5), child pornography at 9.1% (n=4), and indecent exposure at 2.3% (n=1). The majority of participants in this sample had only one known victim (n=42) and had offended against pre-juvenile (n=22) or adolescent victims (n=22),

rather than small child victims (n=3). There were no offenses against infants among these participants. The majority of offenses were perpetrated against female victims (n=44).

Additional participant and victim demographics are found in Table 1.

Our comparison group was comprised of 69 males from an exclusively heterosexual, non-pedophilic sample. This reference group included males ranging in age from 18 to 30 with an average age of 22.01.

Table 1

Participant Demographics and Offense Statistics

	Number	Percentage	
Total Male Participants	44	100%	
Total Victims	47		
<hr/>			
Offense Category			
Incest	5	11.4%	
Rape	15	34.1%	
Molestation	21	47.7%	
Child Pornography	4	9.1%	
Indecent Exposure	1	2.3%	
Single known victim	34	77.3%	
Multiple known victims	9	20.5%	
Unknown number of victims	1	2.3%	
Has at least one PJ or SC victim	25	56.8%	
<hr/>			
Offended against identified victim*:			
Juvenile Female	13-17 years	20	43%
Juvenile Male	13-17 years	2	4%
Pre-juvenile Female	6-12 years	21	45%
Pre-juvenile Male	6-12 years	1	2%
Small Child Female	1-5 years	3	6%
Small Child Male	1-5 years	0	0%
Infant Female	0-12 months	0	0%
Infant Male	0-12 months	0	0%

Note. Excludes offenses of child pornography without an identified victim and includes multiple victims per participant.

Note. Age ranges have not been assigned to categories of the LOOK. As such each image was examined and ranges were assigned based on the perception of age by the principle investigator.

In trying to determine which fair constant multiplier would be used in our study, we tried to find a number, which made the offenders have deviant profiles while avoiding false positives among our sample of non-offenders. As we looked at different numbers it became evident that no number discriminated between these two populations without over identifying the non-offender sample as deviant. Instead we decided to graph what happens to the rate of positive identification as we adjusted the fair constant multiplier. We began with the fair constant set to 60 and continued to chart the increase of those with positive screens in 10 point increments up until we were able to capture the entire sample (identifying all participants as deviant). This number ended up being 1520.

We first started seeing individuals with a positive screen for deviance using a fair constant multiplier of 70. However, the first people identified as having deviant sexual interest were part of the control sample. The rate of identifying offenders as deviant seemed to follow a similar pattern to the identification of non-offenders as deviant. The two points where the most differentiation occurred were when the fair constant multiplier was set to 180 and 540. At 180, 15.9% of the non-offenders were identified as deviant and only 2.3% of the offenders were positively identified, thus obtaining a high rate of false positives compared to the number of positive identifications of convicted sex offenders. At 540, the offenders were identified as deviant at the rate of 84.1% and the non-offenders were identified at a rate of 71.0%. For additional data points see Figure 2 below.

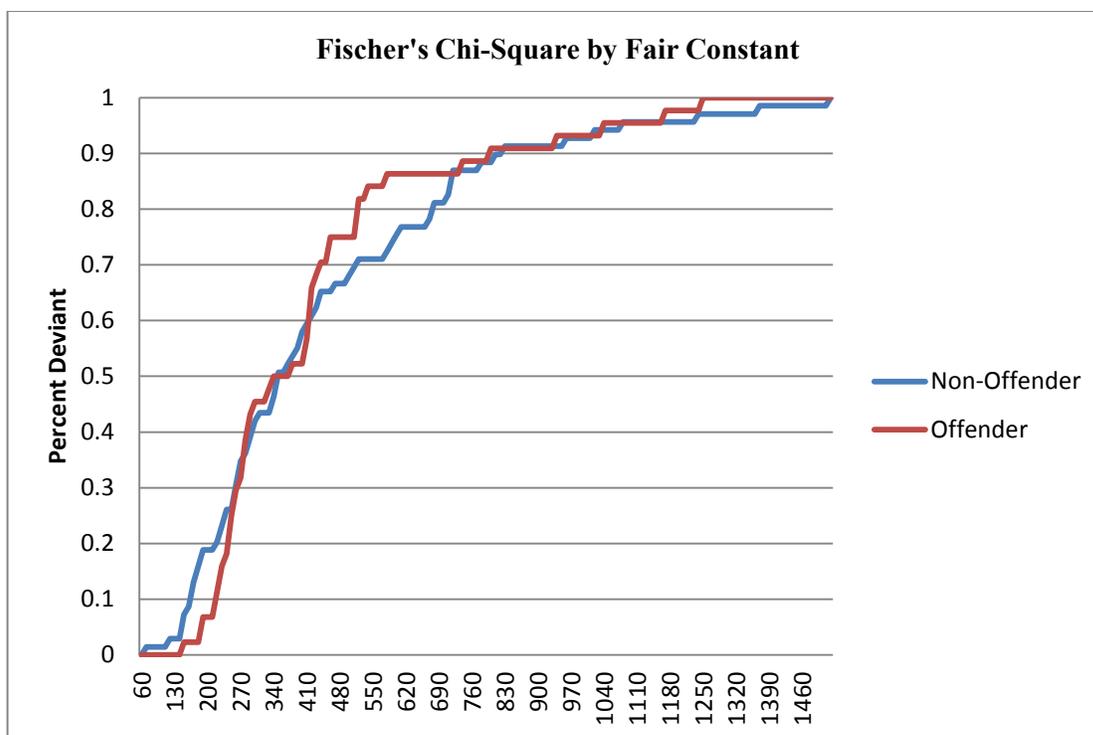


Figure 1. Positive identification of deviance of non-offending and offending samples based on Fischer's Chi-square model using a variety of fair constant multipliers.

When assessing the predictive validity of LOOK using Fischer's Chi-Square residuals to correctly identify their victim demographics, we found that only seven (15.9%) of the 44 participants had Chi-Square residuals that matched their offense histories. For example, one participant who offended against a 12-year-old female showed a Chi Square residual in the category of Pre-Juvenile Female. Those which were correctly identified had offended in three categories: molestation (42.9%), incest (28.6%), and rape (28.6%). This was determined by looking at graphs of the Chi-Square residuals. Two graphs showing the residuals of all of the offender and non-offender samples demonstrated that there did not seem to be any residual pattern established for either population. Figure 2 and 3 show the residuals when the fair constant multiplier was set at 115.

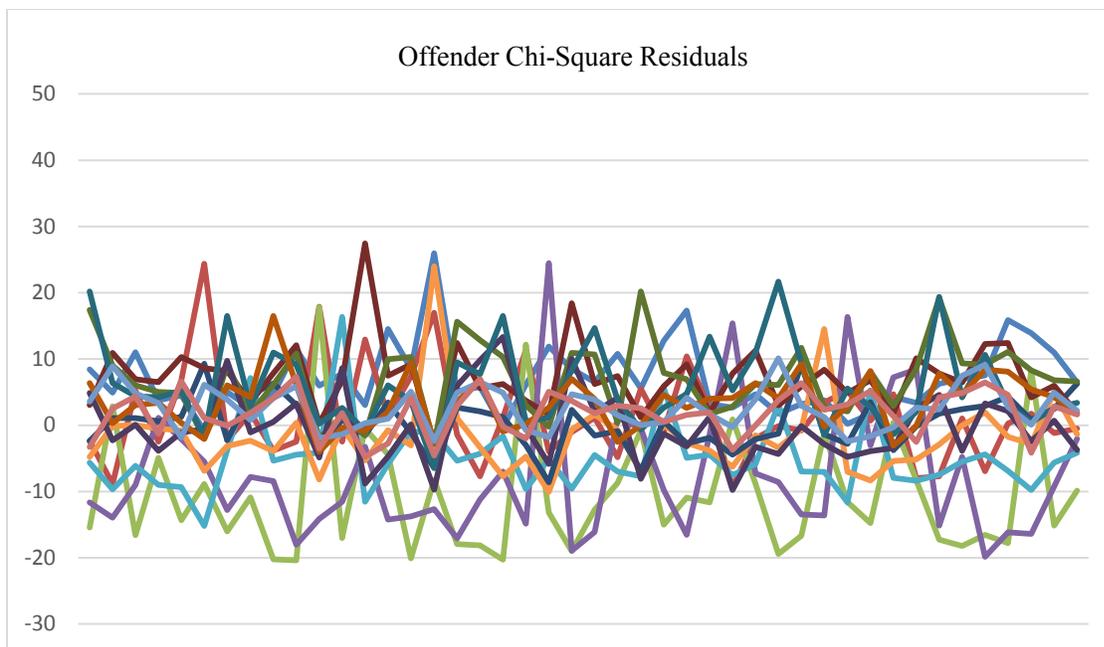


Figure 2. Graph of all participants Chi-square residuals from the offender sample (n=44).

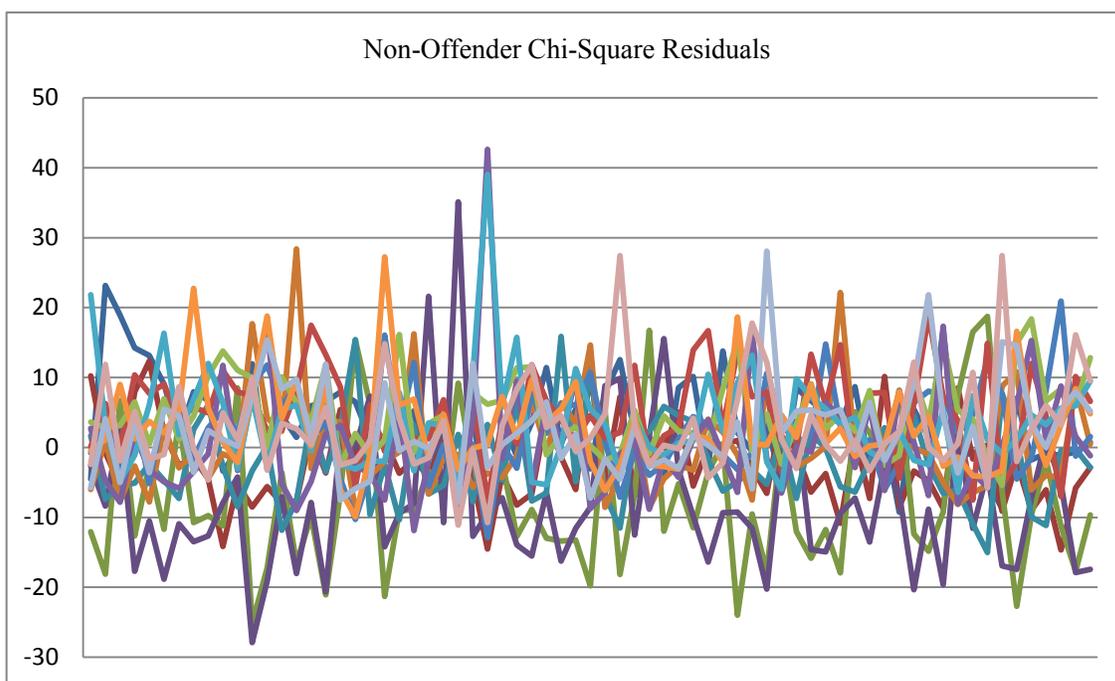


Figure 3. Graph of all participants Chi-square residuals from the non-offender sample (n=69).

Discussion

Using Fischer's Chi-square method, the LOOK does not appear to be a valid tool in screening for deviant sexual interest or for predicting whom perpetrators will offend against. This model identified non-offenders and offenders as having deviant sexual interest at similar rates. The areas with the most discrimination between the two samples produced high rates of false positive screens for the non-offender sample or produced higher rates of identification for non-offenders than offenders. For example, when 2.3% of offenders were identified as deviant, 15.9% of the non-offenders were identified as deviant. The other point where the most discrimination occurred between the two samples was when the offenders were identified as deviant at the rate of 84.1% and the non-offenders were identified at a rate of 71.0%. This model produces more false positives than would be acceptable, especially if used in job screening or to suggest sexual deviance in a criminal proceeding.

Additionally, when looking at the Chi-square residuals, only seven or 15.9% of the 44 participants had their victims' demographics predicted by the Chi-square residual. Thus, no predictive validity could be established either.

Questions and Limitations

A number of questions arise given the results of this study and others like it. One such question may be if viewing time can be considered an accurate measure of sexual attraction. The results of this study showed that a number of the participants spent longer than average looking at pictures of Elderly Adult Females (EAF). This trend would be considered unusual given that the average age of the offenders and non-offenders was between 20 and 30 years old. In looking at the significant residuals, 50% of the non-offender population spent more time than expected looking at EAFs whereas 20.3% spent significantly less time on the same images. Surprisingly,

88.6% of offenders had significant positive residuals when looking at this same sample and none of the offenders spent less time than would be expected when looking at pictures of EAFs.

One possible hypothesis is that individuals were surprised by these images and did not know how to rate these pictures in terms of their sexual attraction to them. Another may be that individuals were quick to rate those who fit their preferred category compared to those in categories they felt more unsure about. Similar hypotheses have been postulated about why women in previous LOOK and Affinity studies spent more time looking at pictures of other women than do their male counterparts at pictures of other men. Researchers in these studies hypothesized that these results may be due to women being more comfortable about rating other women as sexually attractive, that women have a more fluid view of sexuality, or that women spend more time looking at these images because they are comparing their own perceived attractiveness to the women in these images. Therefore, viewing time may be less about sexual attraction and more about social comparison, fluidity of sexuality, or the difficulty of rating the attractiveness of individuals outside of one's preferred category.

Another possible reason for the lack of significant findings in the study is that most sexual offenders are not deviant enough. The rate of true pedophiles, those who have an exclusive attraction to young children, is rare even among the population of convicted sex offenders. In an attempt to understand the causes and context of sexual abuse of minors by priests within the Catholic Church, The John Jay College Research Team conducted a study in which they found that only 5% of the priests would meet diagnostic criteria for pedophilia (Terry et al., 2011). Our study looked only at individuals who had offended against victims under the age of 18. In our sample, 29.5% of the sample offended against a child under the age of 11 and 18.2% offended against children under the age of eight.

Limitations of the instrument itself may also contribute to these findings. The 14 LOOK categories do not have age ranges assigned to them. As such, researchers had to estimate the ages that were depicted among the images of the LOOK. It also seems as that the LOOK did not provide enough of a continuation of age ranges to capture all variants of sexual interest. One such gap between the ages of those pictured was seen in the photos of pre-juvenile males who appear to be approximately between the age of six and 10 years old and the juvenile males who seem to be 14 years old and older. This gap also appears in the photos of pre-juvenile females and juvenile females. While this helps with differentiation between categories, it may not capture the full range of a perpetrator's sexual interest.

Study Limitations

There were a number of limitations of this particular study. The sample size was quite small; only 44 participants remained after screening was conducted. Individuals were excluded from the sample if they had only committed offenses against adults, had offense histories that were greater than 10 years old, or had committed offenses when they were minors. Additionally the two women in the sample were excluded for our analysis and two participants for whom we received LOOK data but were not provided with their offense histories.

The generalizability of the results of this study is also limited by the demographics of the participants. There were significant differences between the non-offender and offender samples. For example, the average age of the non-offender sample was 22.0 whereas the sample of offenders had an average age of 29.4. The non-offender sample self-identified their sexual preferences using a Kinsey Scale. This 7-point scale ranges from exclusively heterosexual to exclusively homosexual with gradient points in between. For the purpose of the study in which these data were collected, individuals who did not mark “Exclusively Heterosexual” were

excluded from the study. Self-identifying information about the offender's sexual preferences was not available. Additionally, while the offender sample was made up of individuals who had committed at least one sexual offense against a minor, it is unknown if anyone from the non-offender sample has committed a sexual offense or may have the propensity to do so in the future.

Another area where the samples differed was the level of control the researchers had over how the instructions were administered to each sample. Historical data were used with the offender sample and researchers had no direct control over how this instrument was administered or what directions were given. The LOOK gives all necessary instructions on one of the first screens seen but it is unknown whether or not the offenders actually read these instructions. Efforts were made to standardize the procedures of test administration among the non-offender sample. Researchers involved in collecting the research were trained, used standardized scripts to read to participants, and recorded any problems that participants encountered when using the instrument. In a study conducted by Baird (n.d.) for her dissertation, she was able to establish reliability by giving the non-offender sample the same test, two weeks apart. No second administration will be available with the offender population to establish the reliability of the instrument.

Offenders in this study were also undergoing psychosexual evaluations per the request of the court system; the pressure to fake good may have been higher for the offender than for the non-offending population. While this data was not used in decision-making or recommendations for their psychosexual evaluations, clients could have perceived this as a test that would influence their future sentencing recommendations.

Biases may exist within the normative sample of BYU students as well. Participants who elected to participate in this study most likely did so for extra credit in a psychology course and were exclusively university students; the educational background of the offender sample is unknown. There may also be vast differences among those willing to participate in an assessment of sexual interest and those who did not sign up.

Future Directions

This study adds to an existing body of research on whether ipsative data produced by viewing time measures such as the Affinity and the LOOK can be translated into norm-referenced scores and used to screen for or diagnose deviance. To try to solve this problem, much of the research done in recent years by Lane Fischer's research team at BYU has used a method of statistical analysis called Fischer's Chi-square. This method was the proposed solution to the problem of making statements of deviance based on ipsative data by establishing a norm from which to compare individual's scores. However, our study produced similar findings to a study conducted by Stephenson (2014) of the predictive validity of the Affinity 2.5 using Fischer's Chi-square approach to scoring. While this approach to scoring does not appear to work, different methods of statistical analysis may prove to be more effective.

Future studies may include looking at the predictive nature of ipsative data. When looking at the ipsative data generated from the LOOK, there seemed to be a better agreement between offenders' ipsative data and their offense histories than was obtained after standardizing their scores and comparing them to established norms. More statistical analysis is required to determine if ipsative scores are more predictive of deviance and consistent with offense histories.

Establishing a wider range of normative patterns for LOOK may be another area of focus in future research. Norms for persons from different ethnicities, genders, age groups, and sexual orientations may help to establish a variety of reference groups for comparison.

References

- Abel, G. G., Jordan, A., Rouleau, J. L., Emerick, R., Barboza-Whitehead, S., & Osborn, C. (2004). Use of visual reaction time to assess male adolescents who molest children. *Sexual Abuse: Journal of Research and Treatment, 16*, 255-265.
doi:10.1177/107906320401600306
- Abel, G. G., & Rouleau, J. L. (1990) The nature and extent of sexual assault. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.). *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 9-21). New York: Plenum.
- Baird, S. (n.d.). *Expected profiles and temporal stability of the LOOK*. (Unpublished doctoral dissertation), Brigham Young University, Provo, Utah.
- Barbaree, H. E., & Mewhort, K. J. K. (1994). The effects of the z-score transformation on measures of relative erectile response strength: A re-appraisal. *Behaviour Research and Therapy, 32*, 547-558.
- Caldwell, M. F. (2013). Accuracy of sexually violent person assessments of juveniles adjudicated for sexual offenses. *Sex Abuse, 25*, 516-526.
- Fischer, L. (2004, October). Psychometric issues in viewing-time research, clinical, and forensic applications. In L. Fischer (Chair), *Accessible, standardized, viewing-time measurement*. Symposium conducted at the 23rd Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Albuquerque, NM.
- Fischer, L., Baird, S., Hansen, K., Stephenson, and Veas-Wall, R. (2012, October). Affinity 2.5. In L. Fischer (Chair), *Falsification of the Affinity 2.5*. Symposium conducted at the 31st Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abuse, Denver, CO.

- Fischer, L., Byrne, P. M., & Glasgow, D. V. (2007, October). *Real world decision-making using both ipsative and norm-referenced monarch PPG and affinity results*. Paper Presented at the 26th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, San Diego, CA.
- Fischer, L., & Meade, A. (2010). *Assessment of sexual interests: Reference-group solutions to problems of ipsative data*. Unpublished manuscript.
- Fischer, L., & Morgan, D. (2006, October). Norm-referenced clinical decision-making with affinity viewing-time results. In L. Fischer (Chair), *Current Research and Clinical Application of Affinity Viewing-Time Scores*. Symposium Conducted at the 25th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Chicago, IL.
- Fischer, L., & Smith, G. (1999). Statistical adequacy of the Abel assessment for interest in paraphilias. *Sexual Abuse: A Journal of Research and Treatment*, 11, 195-205.
- Freund, K. (1990) Courtship disorder. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.). *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 195-207). New York, NY: Plenum.
- Glasgow, D. V., Osborne, A., & Croxson, J. (2003). An assessment tool for investigating paedophile sexual interest using viewing time: An application of single case methodology. *British Journal of Learning Disabilities*, 31, 96-102. doi:10.1046/j.1468-3156.2003.00180.x
- Grese, C. L. Z. (2005). Viewing time measures and sexual interest: Another piece of the puzzle. *Journal of Sexual Aggression*, 11, 117-125. doi:10.1080/13552600500063666

- Israel, E., & Strassberg, D. S. (2009). Viewing time as an objective measure of sexual interest in heterosexual men and women. *Archives of Sexual Behavior, 38*, 551-558.
doi:10.1007/s10508-007-9246-4
- Lesniak, K. T., Rudman, W., Rector, M. B., & David Elkin, T. (2006). Psychological distress, stressful life events, and religiosity in younger african american adults. *Mental Health, Religion & Culture, 9*, 15-28.
- Marshall, W. L. (1996). Assessment, treatment, and theorizing about sex offenders: Developments during the past twenty years and future directions. *Criminal Justice and Behavior, 23*, 162-199. doi:10.1177/0093854896023001011
- Marshall, W. L. & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review, 20*, 807–822.
- Quinsey, V. L., & Chaplin, T. C. (1988). Preventing faking in phallometric assessments of sexual preference. *Annals of the New York Academy of Sciences, 528*, 49-58.
- Quinsey, V. L., Rice, M. E., Harris, G. T., & Reid, K. S. (1993). The phylogenetic and ontogenetic development of sexual age preferences in males: Conceptual and measurement issues. In S. M. Hudson (Ed.), *The Juvenile Sex Offender*. (pp. 143-163). New York, NY: Guilford Press.
- Rosenzweig, S. (1942). The photoscope as an objective device for evaluating sexual interest. *Psychosomatic Medicine, 4*, 150-158.
- Singer, B. (1984). Conceptualizing sexual arousal and attraction. *Journal of Sex Research, 20*, 230-240.
- Stephenson, H. (2014). *Screening and diagnostic validity of Affinity 2.5*. (Doctoral dissertation), Retrieved from <http://scholarsarchive.byu.edu/etd/4232/>

Terry, K. J., Smith, M. L., Schuth, K., Kelly, J. R., Vollman, B., & Massey, C. (2011). The causes and context of sexual abuse of minors by Catholic priests in the United States, 1950-2010. A Report Presented to the United States Conference of Catholic Bishops by the John Jay College Research Team.

The National Center for Victims of Crime. (2013). 2013 national crime victims' rights week resource guide. Retrieved from <http://ovc.ncjrs.gov/ncvrw2013/pdf/StatisticalOverviews.pdf>

Wright, L. W., & Adams, H. E. (1994). Assessment of sexual preference using a choice reaction time task. *Journal of Psychopathology and Behavioral Assessment*, 16, 221-231.
doi:10.1007/BF02229209

Zamansky, H. S. (1956). A technique for assessing homosexual tendencies. *Journal of Personality*, 24, 436-448.

APPENDIX A: Literature Review

Sexual offending is a high frequency crime that causes psychological damage to the many innocent people who have been offended against. Based on The National Center for Victims of Crime report for 2013 the U.S. population of 14- to 17-year-olds, 28% reported that they had been sexually victimized at some point in their lifetime. In 2008 alone, 4.8% of males and 7.4% of females aged zero to 17 were sexually victimized (The National Center for Victims of Crime, 2013). As such, many researchers and practitioners have developed a variety of assessments in an effort to understand sexual abuse, the perpetrators of abuse, and to help make crucial decisions about their treatment and risk to the community (Marshall & Fernandez, 2000).

Research in the field has primarily been focused on male populations of sex offenders. However, female and male sex offenders have been found to be different in several ways that suggest that research conducted on male sex offenders may not be accurately applied to female offenders. Women who are incarcerated for sexual offenses make up less than one percent of the female prison population whereas the prison population of male sex offenders is much higher. This makes research with female sex offender populations quite difficult. Some of the reported differences include that female sex offenders tend to be in caretaking roles of those they offend against whereas men are more likely to offend against strangers. Women are also more likely to be involved with a co-perpetrator when exploiting their victims (Beech, Parrett, Ward, & Fisher, 2009).

Recidivism

Over the course of the last two decades, the use of structured risk assessments to predict recidivism has become standard procedure in evaluating sex offenders (McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010). Many hope to help identify those perpetrators at risk of re-

offense and to identify factors that may contribute to an individual's likelihood to reoffend. A meta-analysis by Hanson and Morton-Bourgon (2005) concluded that sexual deviance (including pedophilia) was the strongest single predictor for sexual offense recidivism among sexual offenders. Risk factors prevalent in research include previous conviction for a sex offense, previous conviction for offenses of nonsexual violence, and repeated offending of any type. Thornton and Travers (1991) studied 313 sex offenders to see if the risk factors commonly found in the literature, stated above, were present among these individuals and what their recidivism rates would be. The men were placed into one of two groups: those who had one or more risk factors identified and those who did not have any. In a 10-year follow up study, 26% of the men who were identified as having one of the common risk factors had been reconvicted whereas only 5% of men who did not have any risk factors were reconvicted. They also found that men with those risk factors were more likely to commit violent offences than those who did not.

Marshall (1996) conducted a long-term follow up study of nearly 13,000 offenders of all types for years after their release from prison. Among these men 926 had committed a sexual re-offense. He found that nearly 7% of those with a history of sexual offending were responsible for 31% of the subsequent sex offense convictions. Of the men who were considered to be at high risk for re-offense, three out of four were not convicted of a sexual offense in the next 10 years; using risk factors to determine who will offend may produce a high number of false positives.

How Results from These Assessments Are Used

Measures of sexual interest are often used in judicial proceedings when an individual is accused or convicted of sexually assaulting another person. In these settings there are several methods of identifying deviant sexual interest for individuals who have been accused of a sexual crime. Two primary reasons for assessment of sex offenders are to determine their treatment

needs and to predict the possibility of recidivism (Grese, 2005). Sexual deviance can be defined either through *The Diagnostic and Statistical Manual of Mental Disorders* (DSM) diagnoses, or through criminal behavior patterns, though each definition may have different meanings. For example, there are many DSM paraphilic diagnoses that are not associated with illegal activity. However, sexual behavior that is criminal may also be diagnosable through the DSM. For the purposes of this paper, deviance will be restricted to those behaviors that are criminally deviant.

In recent decades new legislation has been passed in response to an increased awareness of sexual violence. Longer sentences are being imposed and many states have adopted laws to try to control the risk sex offenders pose to the community when they are released from prison. Twenty states' laws and many federal statutes include one particularly restrictive law, which allows sex offenders to be committed to a secured forensic hospital for an indefinite period of time. These persons have been given the label Sexually Violent Persons (SVP) and have been identified as having a mental abnormality or disorder than makes them likely to reoffend. Considering the ramifications of such decisions on individual's lives, methods of identifying those who are likely to reoffend need to be both highly reliable and valid to be used to make such permanent commitments.

Caldwell (2013) reviewed the records of 198 juveniles who had committed sexually violent offenses that qualified them for possible commitment under the SVP civil commitment law. As part of their evaluation, each youth was screened by at least two expert examiners. Of the 198 juveniles, 54 of the youth were found to meet criteria for an SVP petition. The remaining 144 were screened out. Researchers then collected follow-up data on subsequent criminal charges for an average of nearly five years on both those who qualified for SVP

petitions and those who were screened out. The results of their study did not show significant differences between youth who had been screened out and those subject to petition in terms of prevalence rates for general sexual offending and felony sexual offending. Among petitioned youth, 11.76% were charged with a new sexual offense during the follow-up period, including 9.80% who were charged with a felony sexual assault. Of the non-petitioned youth, 17.36% were charged with a sexual offense, including 13.19% who were charged with a felony sexual assault.

Theories of Sexual Abuse and Victim Choice

Theories about sexual abuse, and many measures used in the assessment process, are based on the construct that sexual abuse is driven by sexual attraction. This is termed the sexual preference hypothesis. If this construct is accurate, trying to understand the sexual preferences of a sex offender is of great importance. If it is not accurate, the assessment of sexual offenders needs to change.

Many in the field believe there are other motivating factors behind sexual abuse besides sexual attraction. Marshall (1996) stated that sexual abuse is not about sexual attraction but rather about issues around power and control, desire to humiliate, and an expression of aggression. Quinsey, Rice, Harris, and Reid (1993) have postulated that those with deviant sexual preferences will engage in specific types of behavior because they are motivated to do so. They emphasize that sexual behaviors are not always an accurate reflection of sexual preference but may be more related to acting within social constraints and/or opportunities presented to them. Other theories and studies have suggested that sex offenders may be more polymorphously perverse than focused on one stimulus category. Abel and Rouleau (1990) and Freund (1990) conducted studies in which they found that outpatient sex offenders had a high

incidence of paraphilias. Able's study found that 61.4% of heterosexual pedophiles had three or more paraphilias and 55% of rapists demonstrated three or more.

Some effort has been made to categorize sex offenders by the characteristics of their victims given the sexual preference hypothesis. However, this idea is based on the assumption that sex offenders are stable in their choice of victims. Studies have found that sex offenders are not always consistent in choice of victims, often offending against a varied victim pool. Kleban Chesin, Jeglic, and Mercado (2012) conducted an archival records search of 789 incarcerated sex offenders. Among this sample, 279 had multiple victims listed in their records. Thirteen percent had victims of both genders and 14% had victims of various ages (e.g., child, adolescent, and adult). In looking at the records of individuals with past sexual convictions ($n = 208$), 20% were found to have offended against a prior victim of a different gender and 40% crossed over across age categories. The findings of this study suggest that sex offenders remain relatively stable with regard to victim gender, though offenders may be less stable with regards to age of their victims across crimes. This study also examined the relationship status (e.g., parent, acquaintance, stranger) between sex offenders and their victims; almost half of the offenders had victims who fell more than one relationship category.

Other studies have reported similar findings. One such study conducted by Heil, Ahlmeyer, and Simmons (2003) found that only 11% of the incarcerated offenders in their sample had abused only one type of victim and 70% admitted to having both adult and child victims. Heil et al. also found that 78% of sex offenders in this sample, who were thought previously to have only molested children, also admitted to the sexual abuses of adults. Weinrott and Saylor (1991) conducted a study with sex offenders identified as rapists or child molesters; their findings also demonstrated the variance in victims among perpetrators. Thirty-two percent

of rapists in their study also reported that they had engaged in sexual conduct with a child and, conversely, 12% of those classified as child molesters also disclosed having at least one adult victim. Levenson, Becker, and Morin (2008) examined a sample of child molesters and found that 22% of the offenders had offended against victims of both genders. They also found that 26% of offenders reported having had both child and adult victims.

Types of Assessments Used

The assessment of sexual offenders has focused primarily on the assessment of sexual interest. The underlying assumption in most of the assessments on the market today is that information about an offender's sexual interests and arousal patterns will help us understand the motivating factors behind the offender's behavior (Abel et al., 2004). Singer (1984) suggested that sexual arousal is a three-step process. First, an aesthetic response or emotional response to something seen as desirable occurs. Second, if increased attention is paid to desirable stimuli, the next naturally occurring step is a desire to be closer to the object of interest. Third, a genital response occurs if the stimuli is attended to and is in close proximity.

A number of approaches to assessment of sexual interest have been and are in use in the field of sex offender assessment. Some instruments use Singer's first step, an aesthetic response in trying to measure sexual attraction while other measures use genital response. Instruments used in the field include polygraphs, psychometrics, historical partner choice, pupilometry, self-report interviews, gaze tracking, stimulus rating/card sort, plethysmography, and viewing time. Clinical interviews and self-report questionnaires aside, most other methods for assessing pedophilic sexual interest can be categorized into either physiological or cognitive measures. All of these assessments aim for a relatively objective measure of sexual interest or arousal; each method has relative strengths and weaknesses (Glasgow, Osborne, & Croxen, 2003).

Marshall (1996) expressed his belief that—while measures of sexual interest are valuable tools in assessing the risk for re-offense and for providing treatment recommendations—these types of evaluations do not belong in criminal court proceedings. He stated, “In the present author’s opinion, evidence currently available does not justify, and perhaps never will, the use of psychological or psychiatric evaluations by either defense or prosecution counsel in such cases.” Marshall also asserted his opinion that the use of these measures to conduct risk-assessments is still in its infancy and needs more research to support this idea (Marshall, 1996).

For the purpose of this paper the strengths and weaknesses of the most popular assessment approaches found in the literature—self-report/clinical interview, plethysmography, and viewing time were examined. We have also briefly examined some of the findings of neuroimaging, as this technology is being incorporated into the field of assessing sex offenders.

Self-Report/Clinical Interview Measures. Self-report measures and clinical interviews may gain access to information that cannot be assessed in any other way. These measures, however, are also prone to deception, impression management, and defensiveness (Grese, 2005). Offenders tend to minimize the seriousness of their actions or deny the accusations against them. They also tend to shift responsibility from themselves to outside factors such as alcoholism or an unfulfilling sex life with their partner as reasons for their acting out. Therefore, having access to victim statements, trial transcripts, and police reports may be necessary to get an accurate picture of the abuse (Marshall, 1996). This may be of particular concern in clinical work with sex offenders who intentionally or unintentionally misrepresent their sexual interests. For these reasons, researchers and clinicians have attempted to create more objective measures of sexual interest or arousal rather than rely on the subjective information provided by the offender; newer

measures include plethysmography, viewing time measures, and neuroimaging (Israel & Strassberg, 2007).

Plethysmography. Freund was the first to use plethysmographic assessment procedures to test sexual arousal in males. He began this assessment process in 1957 using a device that measured volumetric (length and circumference) changes in the penis; this volumetric measure appeared to be more sensitive than later devices that only measured changes in circumference of the penis. Researchers have continued to use instruments which only circumference changes in their measurement as these devices are reported to be less complicated, more cost-efficient, and less prone to breaking than those which measure volume changes (Marshall, 1996; Marshall & Fernandez, 2000).

Vaginal photoplethysmography (VPP) and penile plethysmography (PPG) measure vascular changes in the genital area in response to sexual stimuli; VPP and PPG are the most commonly used objective measures of sexual interest arousal. Plethysmography has been described as the most sensitive and reliable indicator of sexual attraction. Many studies in the sexuality literature suggest a strong positive correlation between self-reports of sexual arousal and changes in penile circumference or volume as measured by a penile plethysmograph. However, based on the research conducted by Israel and Strassberg (2007), vaginal plethysmography does not appear to have as strong of a correlation between reported sexual arousal and physiological response.

Research by Suschinsky, Lalumie`re, and Chivers (2009) has examined the differences between male and female arousal to category-specific stimuli using plethysmography. They have found that genital arousal in males is more category-specific than is genital arousal in females. For example, when men were shown pictures of individuals who matched their self-

reported sexual preferences, they experienced their highest levels of sexual arousal. Women, on the other hand, experienced significant arousal to stimuli both when the stimuli matched their self-reported sexual preferences and when shown pictures of a sexual nature that did not fit their self-reported preferences. In a study by Suschinsky et al. (2009), women showed high genital responses to all three forms of human sexual interactions (i.e., male–male, female–female, and male–female) presented to them but their highest subjective responses were to the male–female stimuli. Researchers found that, even in circumstances where there is no apparent motivation to misrepresent their sexual arousal/interest, there was shown to be little correlation between vaginal photoplethymography (VPP) and self-reports of arousal or interest (Israel & Strassberg, 2009).

Another study conducted by Chivers, Rieger, Latty, and Bailey (2004) seems to show similar patterns regarding female sexual arousal. In this study heterosexual and homosexual men and women, and post-operative male-to-female transsexuals, were shown a variety of films depicting different sexual scenarios. Women in this study were found to experience their greatest genital arousal to sexually explicit material that did not match with their self-reported sexual preferences. However, when women reported on their feelings of sexual arousal, they reported that they felt the greatest arousal to films that corresponded with their self-reported sexual orientation and sex partner preferences. The men in this study and male to female transsexuals did not follow this same pattern.

Faking. Quinsey and Chaplin (1988) reported that many studies have shown that men are able to enhance and inhibit their erectile responses using a number of faking strategies. When encouraged to do so, both men and women have demonstrated an ability to suppress their plethysmographically-assessed arousal to stimuli that they report to be arousing or of sexual

interest (Israel & Strassberg, 2009). According to Abel et al. (2004) test subjects have disclosed falsifying their test results by not looking at the stimuli, only listening to a portion of the stimuli, or by using imagery to control their responses. Individuals have also been known to use “pumping”—the contracting of their perineal muscles—to increase the appearance of arousal to socially appropriate stimuli. Other strategies include using cognitive processes to distract themselves from the erotic stimuli. For example, some individuals successfully control their erectile responses by thinking of difficult math problems. It appears nearly impossible to prevent or detect faking of test results, especially when cognitive strategies are employed. Marshall and Fernandez (2000) suggest that because there is no way to prevent all methods of deception, faking will always serve as a threat to the validity of phallometric assessments.

Quinsey and Chaplin (1988) also conducted a PPG faking study in which they asked 15 non-pedophilic males to inhibit their response to desired stimuli. After the experiment researchers asked participants to describe the strategies they employed to control their penile tumescence. Participants admitted to substituting the audio stimuli in their minds with their own fantasies, causing physical pain to another part of their body, or using memorization to know when to push the button but not have to attend to the stimuli.

Adams, Motsinger, McAnulty, and Moore (1992) conducted a study into the voluntary control of penile tumescence under two conditions. In their study they found that participants produced significantly increased penile tumescence in the presence of their preferred sexual stimuli and produced no significant increase in the presence of non-preferred sexual stimuli when asked to focus on the stimuli presented. They also found that when given the instructions to try to falsify their sexual arousal patterns, the participants were able to suppress their penile

tumescence to a significant degree. On the other hand, the men in the study were mostly unsuccessful when trying to increase their arousal to a non-preferred stimulus.

In addition to faking concerns, PPG and VPP responses have been shown to be affected by lack of sleep, illness, or recent sexual activity.

Other Research Limitations. Marshall and Fernandez (2000) suggest that these tests need to be standardized in order to get reliable estimates of their reliability and validity. Currently, no consistent presentation of stimuli is used in studies measuring VPP or PPG. Marshall (1996) reports that studies using plethysmography have had such varied methodologies that drawing conclusions about validity between studies is not justified and may be misleading. While many studies have used the same stimulus categories, the nature of the stimuli presented is often varied; stimuli are presented in a number of modalities including film, videotapes, slides, audio recordings or print images. Within these modalities the stimuli also vary in terms of black-and-white versus color stimuli, the background used in the presentation, and whether more than one subject was presented (Marshall & Fernandez, 2000). In a study conducted by Abel and Blanchard (1976), they found that using videos as the mode of presentation generated the highest levels of arousal. While using videos produces strong sexual arousal among participants, researchers warn that using such strong stimuli may affect differential responding.

Other factors that make it difficult to compare one study to another include how sexual arousal is measured and how much historical information is gathered about participants. As a form of measurement, some PPG studies measure volume changes where others measure circumference changes. Additionally, some studies use average tumescence as their reported scores while other studies only record the peak rating tumescence rating of the respondent's sexual arousal. Studies must also clearly define both the parameters of the test as well as the

histories of the test subjects. Demographic information about participants including offense histories, nature of the offense, victim demographics, and details about under what condition tests were administered would be important in comparing one study of sex offenders to another. All of these variables contribute to the difficulty in comparing one study to another. For example, Freund's early volumetric studies and Quinsey's circumferential studies appeared to be able to identify deviant sexual attraction to children of non-familial child molesters whereas these deviant attraction patterns did not appear among those who had committed incest.

Ethical concerns also are a problem in conducting plethysmographic research. Plethysmography is a very invasive procedure; obtaining an adequate comparison sample that is representative of the general population is a challenge. Another issue raised in the research is the nature of the stimuli presented; some researchers worry that exposing individuals to deviant stimuli may only encourage deviant tendencies.

Viewing Time Measures. Viewing time measures of sexual attraction work under the assumption that if the participant finds the person in the image to fit their preferred category of sexual interest, they will spend more time looking at the image. Conversely, if the person in the image does not fit their preferred category, the participant will not spend much time looking at the image (Israel & Strassberg, 2009; Quinsey, Ketsetzis, Earls, & Karamanoukian; Quinsey et al., 1993, 1996). Freund (1990) hypothesized that viewing time may be a measure of sexual interest because it mimics the early stages of courtship, namely the evaluation of a potential partner.

The field of sexuality research has suggested that viewing time is a valid measure for understanding the sexual interest of those tested, however, few studies have established psychometric validity. Evidence for what Israel and Strassberg (2009) call "perceived validity"

is the number of viewing time instruments that have been developed and made commercially available in the assessment of sex offenders. More research is needed, however, to know if these tests are accurate enough to make high stake decisions based on their results. Israel and Strassberg (2007) stated, “The methodological limits of both self-report and genital measures have led some researchers and clinicians to look for alternative ways to assess sexual interest and arousal. One approach that holds promise is viewing time.” Israel and Strassberg postulated that one of the possible advantages that viewing time might have over plethysmography is that it may be less consciously manipulated since viewing time is an indirect measure of sexual interest.

The relationship between sexual attraction and time spent looking at images was initially described by Rosenzweig in 1942. He found that differences in the viewing time of sexual images discriminated level of sexual interest. In 1956, Zamansky continued with this line of study. He presented a group of men with pairs of images depicting both men and women. Based on how long the men spent viewing the pictures, Zamansky was able to distinguish between those with homosexual and heterosexual preferences. Wright and Adams (1994) found similar results in their study of both heterosexual and homosexual men and women; participants spent more time looking at images that matched their self-identified sexual orientation. In their 1999 follow-up study, they again found viewing time to be significantly correlated to sexual orientation for heterosexual men, gay men, and lesbians. However their findings did not support such a correlation for heterosexual women (Wright & Adams, 1999).

Viewing time measures also seems to use people’s natural tendency towards categorizing stimuli to help simplify and streamline the perception process. Macrae and Bodenhausen (2000) postulated that, once these categorical representations are triggered, cognitive content associated with knowledge of the category is generated. These theories are found within the construct of

viewing time measures. When a subject knows that the person does not fit their preferred sex partner category, they may easily dismiss the image. However, when subjects find the stimuli to be sexually arousing, they may spend more time looking at the image trying to discriminate within their preferred category of interest.

Gender differences were also assessed by Quinsey et al., (1996). They hypothesized that because men tend to be more visually stimulated than women, heterosexual men's viewing time would be longer when looking at pictures of women than would heterosexual women's viewing time while looking at pictures of men. A study conducted by Wright and Adams (1999) seemed to confirm this hypothesis. They found that men spent more time viewing images of their self-reported preferred sexual partners than did females of their self-reported partner choice. Quinsey et al. (1996) found that heterosexual men and women viewed pictures of the opposite sex longer than they did pictures of their own sex. Researchers also predicted that individuals would look longer at pictures of individuals who are in their prime for childbearing.

Quinsey et al. (1993) argued that based on the hypothesis that viewing time reflects sexual interest, viewing times should correlate with other ratings of sexual attractiveness including phallometrically-measured age and gender preferences. They found positive correlations between viewing time and ratings of sexual attractiveness among both male and female subjects; this trend appeared to be higher among male test subjects. Abel and Rouleau (1990) also found a positive correlation between subjects' self-reported level of sexual arousal and their viewing time. More recently, Abel et al. (2004) reported that viewing time was comparable to plethysmography in the assessment of sexual interest of adult child molesters as they found the results of viewing time measures and plethysmography to be significantly correlated.

Israel and Strassberg (2009) conducted a study using 106 self-identified heterosexual men and women in which they asked participants to rate the sexual appeal of a number of sexually provocative pictures while covertly measuring the length of time each picture was viewed. Several of their hypotheses proved to hold true for this sample, including that men and women looked at pictures of the opposite sex significantly longer than those of their own sex. They also found that male participants viewed opposite sex pictures significantly longer than did female participants and that females viewed pictures of their own sex significantly longer than the males in the study. One surprising finding in this study was that the ratings of sexual attraction did not correlate with the measured viewing times of either male or female participants.

Faking and Viewing Time. Fischer et al. (2012) conducted a faking study using the Affinity. They recruited 240 subjects, 120 men and 120 women, and assigned participants to one of four conditions. In the first condition, participants were asked to take the assessment as quickly as possible while still attending to the images. The second condition had participants take the assessment as quickly as possible without attending to the images presented to them. The results for these two conditions suggested that even when told to not attend to the images or to rate them as quickly as they could, they still spent more time looking at images that were consistent with their self-reported sexual preference.

For the third condition, participants were asked to pretend to be the opposite gender and rate the images as they thought the opposite gender would. Finally, in the fourth condition, participants were told briefly about the covert measurement of viewing time and told to look longer at images that men or women would find appealing while pretending to be the opposite gender. Results from these conditions indicated that using a cognitive strategy, pretending to be

a different gender with different sexual preferences, or knowing what the test is actually measuring can help someone fake the test.

One strategy that researchers are using to make these assessments harder to fake is through the use of a dot probe. A dot probe is a dot placed on each of the images, which the participants are required to find before rating the image. Requiring participants to attend to another stimulus seems to inhibit their ability to use cognitive strategies to fake a test.

The environment in which the test is given also appears to be important to obtaining uninhibited results. Brown, Amoroso, Ware, Pruesse, and Pilkey (1973) conducted a study where subjects were placed in one of two conditions and given a viewing time measure. Some subjects were given the test individually and others were asked to complete the test in the presence of others. Results indicated that participants who were alone spent longer and more variable intervals in looking at the slides than when in the presence of others. Brown, et al. (1973) suggested that these results might be due to impression management.

Limitations of the Research. There appears to be great variability between the images used in viewing time measures. Some measures, such as the Affinity and the Able Assessment for Sexual Interests (AASI), use fully-clothed models. However, other studies use nude or pornographic images in the measuring of viewing time.

In response to studies that suggest individual variability and gender differences in responding to viewing sexual stimuli, Rupp and Wallen (2009b) hypothesized that this variability may be, in part, due to the use of varied types of stimuli. They conducted a study using exclusively heterosexual participants who were asked to view 216 pictures that depicted various sexual activities. Some of the variation among the images includes differences in the gaze of the female actor, and the proportion on the image that was dedicated to the genital

region. Overall, they found that men and women did not differ in either their subjective ratings or viewing times of the images. However, there appeared to be some preferences for specific types of images within the sample that were more gender specific. Both sexes reported finding pictures of the opposite sex receiving oral sex as least sexually attractive and both sexes looked longer at pictures showing the female actor's body. Participants did not spend as long looking at close-ups of genitals and participants who reported that they were taking oral contraceptives rated genital images as less sexually attractive. Differences were shown in female ratings of pictures where the female actor was looking directly versus looking off camera while viewing time in men did not seem to be affected by the female actor's gaze.

This study adds to the literature demonstrating sex-specific preferences for the specific types of stimuli (Rupp & Wallen, 2009b). Brown et al. (1973) also found that viewing time tends to increase as images become increasingly pornographic. He also suggested that participants looked at images they found to be unpleasant and distasteful for less time, possibly out of disgust. A number of factors have been shown to be influential in how long test subjects looked at the pictures presented to them. Rupp et al. (2009a) cited research that demonstrated among heterosexuals, hormonal states, relationship goals, and participant's social situation all influence how they respond to images of the opposite sex. Recent sexual activity also seemed to have an effect on the subjective reports of attraction. Rupp's research provided some evidence to support these claims. Individuals who did not report any recent sexual activity found the stimuli to be more sexually attractive than did participants who have reported having engaged in sexual activity within the preceding month (Rupp & Wallen 2009b). He also found that subjective ratings were not influenced by partner status for either men or women. They did find, however, that women who did not report involvement with a current sexual partner spent more time

viewing the photos than did women who reported having a current sexual partner. Partner status for men did not appear to influence the time they spent looking at the images. Rupp hypothesized that, for women, being in a relationship suppresses their interest in looking at alternative partners (Rupp et al., 2009a).

Viewing time may also be dependent on characteristics of the perpetrator. Abel et al., (2004) conducted a study using the AASI in which they found that male adolescent sex offenders who had molested children looked significantly longer at slides depicting images of children that did non-offending youth. They also found that the viewing time of images of children by those who had molested children were correlated with the numbers of victims they had and the number of acts of molestation committed. Researchers concluded that these results support the AASI as a valid measure of assessing sexual interest in adolescent male child molesters. Harris, Rice, Quinsey, and Chaplin (1996) conducted a study in which they found that sexual attractiveness ratings and viewing times were highly correlated for men in a non-pedophilic control group, but that the correlations were much lower among those who had been convicted of child molestation.

Another research limitation is the recruiting of a sample population for research. Giving personal information about one's own sexual experience, viewing erotic stimuli, and revealing sexual arousal states are aspects of sexuality research that participants often find uncomfortable. As such, many would-be participants do not agree to engage in such experiments. Strassberg and Lowe (1995) studied a sample of over 1,000 introductory psychology students. As part of a two-phase study, students were asked to participate in a variety of sexuality studies. Researchers then looked at the difference between those willing and unwilling to participate in the studies. Those who were willing to participate, compared to those who were not, had a more positive attitude towards sexuality, less sexual guilt, and reported more sexual experiences. These results

held true for both male and female test subjects. Findings also suggested that participants were more likely to be willing to participate in interviews than to participate in viewing erotic video and even less willing to participate in plethysmograph studies.

Lastly, viewing time measures that use sexually explicit images run up against child pornography laws. As such, many measures in the field use fully clothed models in sexually neutral positions. Others use computer-generated images such as the Not Real People Set to construct images of nude models to circumvent child pornography laws.

Neuroimaging. Neuroimaging is a growing field of study, which is expanding into many areas of study including sex offender research. Functional magnetic resonance imaging (fMRI) technology is used to measure brain activity and works by measuring changes in the blood oxygenation and blood flow that occur in relation to a neural activity. An fMRI is often used to create maps showing what parts of the brain are used in a specific mental process. Schiffer et al. (2008) conducted a study using fMRIs to analyze the processing of visually sexually arousing material. Samples of homosexual pedophilic patients as well as a homosexual control group were studied. Researchers found that when the subjects were presented with sexually arousing pictures of homosexual and pedophilic content, both groups of subjects showed activation in the occipito-temporal and prefrontal cortices, areas known to be involved in the processing of visual stimuli with emotional content. The sample of pedophiles also showed brain changes in the globus pallidus, substantia nigra and striatum, key areas identified as being involved in mediating sexual arousal and behavior. The control sample did not have these same changes seen in their brains (Schiffer et al., 2008).

Suchy, Wilson-Whittaker, Strassberg, and Eastvold (2009) conducted a study in hopes of gaining more clarity into the inconsistent findings of neurological tests within the population of

sex offenders. They conducted a study where they used 60 male participants and divided them into three categories: non-offenders, pedophilic child molesters, and non-pedophilic child molesters. The distinction between pedophilic and non-pedophilic child molesters was based on a self-report to a therapist, a score of four or above on the Screening Scale for Pedophilic Interests, or response to PPG data that showed a sexual preference for children under the age of 13. They found that pedophilic child molesters exhibited slower processing speed, nonpedophilic child molesters exhibited poor semantic knowledge, and both molester groups showed executive weakness in comparison to the control group.

Neuropsychological testing of sex offenders, however, is still in its infancy. Joyal, Beaulieu-Plante, and de Chantérac (2013) suggest that neuropsychological data on sex offenders is still too scarce to be able to draw conclusion, confirm trends, or test more precise hypotheses.

References

- Abel, G. G., & Blanchard, E. B. (1976). The measurement and generation of sexual arousal in male sexual deviates. In M. Hersen, R. Eisler, & P.M. Miller (Eds.), *Progress in Behavior Modification*, 11, 99–136. New York: Academic Press.
- Abel, G. G., Jordan, A., Rouleau, J. L., Emerick, R., Barboza-Whitehead, S., & Osborn, C. (2004). Use of visual reaction time to assess male adolescents who molest children. *Sexual Abuse: Journal of Research and Treatment*, 16, 255-265.
doi:10.1177/107906320401600306
- Abel, G. G., & Rouleau, J. L. (1990). The nature and extent of sexual assault. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.). *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 9-21). New York: Plenum.
- Abel, G.G., Rouleau, J., Lawry, S.S., Barrett, D.H., and Camp, N.L. (1990) *A non-invasive physiologic measure of adult sexual preference*. Sixteenth Annual Meeting of the International Academy of Sex Research, Sigtuna, Sweden.
- Adams, H. E., Motsinger, P., McAnulty, R. D., & Moore, A. L. (1992). Voluntary control of penile tumescence among homosexual and heterosexual subjects. *Archives of Sexual Behavior*, 21, 17-31.
- Beech, A. R., Parrett, N., Ward, T., & Fisher, D. (2009). Assessing female sexual offender's motivations and cognitions: an exploratory study. *Psychology, Crime & Law*. 15, 201-216.
- Brown, M., Amoroso, D., Ware, E., Pruesse, M., & Pilkey, D. (1973). Factors affecting viewing time of pornography. *Journal of Social Psychology*, 90, 125-135.

- Caldwell, M. F. (2013). Accuracy of sexually violent person assessments of juveniles adjudicated for sexual offenses. *Sex Abuse, 25*, 516-526.
- Chivers, M. L., Rieger, G., Latty, E., & Bailey, J. M. (2004). A sex difference in the specificity of sexual arousal. *Psychological Science, 15*, 736–744.
- Freund, K. (1990) Courtship disorder. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.). *Handbook of sexual assault: Issues, theories and treatment of the offender* (pp. 195-207). New York: Plenum.
- Glasgow, D. V., Osborne, A., & Croxen, J. (2003). An assessment tool for investigating paedophile sexual interest using viewing time: An application of single case methodology. *British Journal of Learning Disabilities, 31*, 96-102. doi:10.1046/j.1468-3156.2003.00180.x
- Grese, C. L. Z. (2005). Viewing time measures and sexual interest: Another piece of the puzzle. *Journal of Sexual Aggression, 11*, 117-125. doi:10.1080/13552600500063666
- Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristic of persistent sexual offenders: a meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73*, 1154-1163.
- 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.
- Heil, P., Ahlmeyer, S., & Simons, D. (2003). Crossover sexual offenses. *Sexual Abuse: A Journal of Research and Treatment, 15*, 221-236.

- Israel, E., & Strassberg, D. S. (2009). Viewing time as an objective measure of sexual interest in heterosexual men and women. *Archives of Sexual Behavior, 38*, 551-558.
doi:10.1007/s10508-007-9246-4
- Joyal, C. C., Beaulieu-Plante, J., & de Chantérac, A. (2014) The Neuropsychology of Sex Offenders: A Meta-Analysis. *Sexual Abuse, 26*, 149 -177
- Kleban, H., Chesin, M. S., Jeglic, E. L., & Mercado, CC. (2013) An exploration of crossover sexual offending. *Sexual Abuse-A Journal Of Research And Treatment, 25*, 427-443. doi: 10.1177/1079063212464397
- Levenson, J. S., Becker, J., & Morin, J. W. (2008). The relationship between victim age and gender crossover among sex offenders. *Sex Abuse, 20*, 43-60.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*(1), 93.
- Marshall, P. (1994) Reconviction of imprisoned sexual offenders. Research Bulletin. 36. London Home Office Research and Statistics Department.
- Marshall, W. L. (1996). Assessment, treatment, and theorizing about sex offenders: Developments during the past twenty years and future directions. *Criminal Justice and Behavior, 23*, 162-199. doi:10.1177/0093854896023001011
- Marshall, W.L., & Fernandez, Y. M. (2000). Phallometric testing with sexual offenders: limits to its value. *Clinical Psychology Review, 20*, 807-822.
- McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). Current practices and emerging trends in sexual abuser management. The Safer Society Press: Brandon, Vermont.

- Quinsey, V. L., & Chaplin, T. C. (1988). Preventing faking in phallometric assessments of sexual preference. *Annals of the New York Academy of Sciences*, 528, 49-58.
- Quinsey, V. L., Ketsetzis, M., Earls, C., & Karamanoukian, A. (1996). Viewing time as a measure of sexual interest. *Ethology & Sociobiology*, 17, 341-354. doi:10.1016/S0162-3095(96)00060-X
- Quinsey, V. L., Rice, M. E., Harris, G. T., & Reid, K. S. (1993). The phylogenetic and ontogenetic development of sexual age preferences in males: Conceptual and measurement issues. In S. M. Hudson (Ed.), *The Juvenile Sex Offender*. (pp. 143-163). New York, NY US: Guilford Press.
- Rosenzweig, S. (1942). The photoscope as an objective device for evaluating sexual interest. *Psychosomatic Medicine*, 4, 150-158.
- Rupp, H. A., Librarch, G. R., Feipel, N. C., Ketterson, E. D., Sengelaub, D. R., & Heiman, J. R. (2009a) Partner status influences women's interest in the opposite sex. *Human Nature*, 20, 93-104. doi: 10.1007/s12110-009-9056-6
- Rupp, H. A., & Wallen, K. (2009b) Sex-specific content preferences for visual sexual stimuli. *Archives of Sexual Behavior*, 38, 417-426. doi: 10.1007/s10508-008-9402-5
- Schiffer, B., Krueger, T., Paul, T. MD; de Greiff, A., Forsting, M., Leygraf, N., Schedlowski, M., & Gizewski, E. (2008) Brain response to visual sexual stimuli in homosexual pedophiles. *Journal of Psychiatry & Neuroscience*. 33, 23-33.
- Singer, B. (1984). Conceptualizing sexual arousal and attraction. *Journal of Sex Research*, 20, 230-240.
- Strassberg, D.S., & Lowe, K. (1995). Volunteer bias in sexuality research. *Archives of Sexual Behavior*, 24, 369-382.

- Suchy, Y., Wilson Whittaker, J., Strassberg, D. S., & Eastvold, A. (2009). Neurocognitive differences between pedophilic and nonpedophilic child molesters. *Journal of International Neuropsychological Society*, *15*, 248-257.
- Suschinsky, K. D., Lalumie`re, M. L., & Chivers, M. L. (2009). Sex differences in patterns of genital sexual arousal: measurement artifacts or true phenomena? *Archives of Sexual Behavior* *38*, 559–573.
- Thornton, D. & Travers, R. (1991) A longitudinal study of the criminal behaviour of convicted sexual offenders. Proceedings of the Prison Psychologists' Conference, HM Prison Service.
- The National Center for Victims of Crime. (2013). 2013 national crime victims' rights week resource guide. Retrieved from <http://ovc.ncjrs.gov/ncvrw2013/pdf/StatisticalOverviews.pdf>
- Weinrott, M. R., & Saylor, M. (1991). Self-report of crimes committed by sex offenders. *Journal of Interpersonal Violence*, *6*, 286-300.
- Wright, L. W., & Adams, H. E. (1994). Assessment of sexual preference using a choice reaction time task. *Journal of Psychopathology and Behavioral Assessment*, *16*, 221-231.
doi:10.1007/BF02229209
- Wright, L. W., & Adams, H. E. (1999). The effects of stimuli that vary in erotic content on cognitive processes. *Journal of Sex Research*, *36*, 145-151.
- Zamansky, H. S. (1956). A technique for assessing homosexual tendencies. *Journal of Personality*, *24*, 436-448.