The Screening Scale for Pedophilic Interests (SSPI): Construct, Predictive, and Incremental Validity

Leslie Helmus, Caoilte Ó Ciardha[1], and Michael C. Seto

Accepted for publication in *Law and Human Behavior*

Author Note

Leslie Helmus, Forensic Assessment Group, Ottawa, ON, Canada. Caoilte Ó Ciardha, Centre of Research and Education in Forensic Psychology (CORE-FP), School of Psychology, University of Kent, UK. Michael C. Seto, Integrated Forensic Program, Royal Ottawa Health Care Group, Brockville, ON, Canada

We would like to thank R. Karl Hanson for permission to use these data and for his feedback on a draft of this paper.

Correspondence concerning this article should be addressed to Leslie Helmus, Forensic Assessment Group, 125-3290 Southgate Road, Ottawa, ON, Canada, K1V 8X2. Email: lesliehelmus@yahoo.ca

[1] Note to editors/copy-editors: The second author's surname is Ó Ciardha. The Ó should not be mistaken for an initial

Abstract

This study of 410 adult male sex offenders against children using data from the Dynamic Supervision Project (Hanson, Harris, Scott, & Helmus, 2007) examined the construct, predictive, and incremental validity of the Screening Scale for Pedophilic Interests (SSPI), a brief proxy measure of phallometrically-assessed sexual response to children that is based on sexual victim characteristics (Seto & Lalumière, 2001). As predicted, the SSPI was significantly related to the deviant sexual interests item on the STABLE-2007, a dynamic risk measure encompassing multiple domains, and with the deviant sexual interests item from its predecessor, the STABLE-2000. The SSPI was unrelated (or more weakly related) to items measuring general antisociality. In addition, the SSPI significantly predicted sexual recidivism, defined as new charges or convictions for sexual offenses, and a broader sexual recidivism outcome that included breaches of community supervision conditions that might involve sexually motivated behavior (e.g., being in the presence of children unsupervised). The SSPI did not add to the predictive accuracy of two actuarial risk measures, the Static-99R and Static-200R, but it did add to the predictive accuracy of the STABLE-2007. Additional analyses suggest the SSPI can serve as a substitute for the STABLE-2007 deviant sexual interests item, if necessary (e.g., in archival research), when assessing sexual offenders against children.

Keywords: deviant sexual interests, sexual offenders, prediction, recidivism, risk assessment

The Screening Scale for Pedophilic Interests (SSPI): Construct, Predictive, and Incremental Validity

Contrary to popular beliefs about sexual offending, not all offenders are sexually aroused by abusive, coercive, or violent sexual activity (Harris, Lalumière, Seto, Rice, & Chaplin, 2012; Seto, 2008). However, deviant sexual interests such as pedophilia or other paraphilias have been found to be among the single strongest predictors of sexual recidivism among sexual offenders (Hanson & Morton-Bourgon, 2004, 2005). As a result, when working with perpetrators of sexual abuse, it is important that clinicians and researchers have adequate measures of deviant sexual interest with which to inform risk management, treatment, and research. When assessing offenders against children, pedophilic sexual interest may be deduced using psychophysiological (e.g. phallometric assessment of penile responses), attentional (using viewing or reaction time tasks; for a review see Snowden, Craig, & Gray, 2011), or self-report approaches (using interviews or questionnaires), or by examining the offender's sexual history.

The Screening Scale for Pedophilic Interests

The Screening Scale for Pedophilic Interests (SSPI; Seto & Lalumière, 2001) was designed to offer a quick assessment of pedophilic interest based on offense history details among offenders with at least one child victim. The SSPI is not designed to replace measures such as phallometry. Rather, it is intended as a quick guide to aid in triaging cases or as a proxy measure of sexual interest where other measures are unavailable, such as when conducting archival research or when assessing offenders who refuse to participate in phallometry. A SSPI score is calculated from four items about the offender's sex offenses that have been linked to greater sexual interest in children (Seto & Lalumière, 2001). These items are: any male child victim, more than one child victim, any unrelated child victim, and any victims under 12. Child

victims refer to victims of age 13 and younger. Items are coded as 0 (absent) and or 1 (present), except for the male victim item which is given a weight of 2 when present because Seto and Lalumière (2001) found it to have a stronger relationship (approximately double that of the other items) with phallometrically-assessed sexual arousal to children. Consequently, total SSPI scores range between 0 and 5, with higher scores indicating a greater likelihood of sexual interest in children.

Construct validity. Using a large sample of 1,113 child molesters, Seto and Lalumière (2001) demonstrated that total SSPI scores were significantly and moderately correlated with phallometrically assessed indices of pedophilia (r = .34), suggesting construct validity. These findings were replicated in two smaller samples (Ns = 113, 145), albeit yielding slightly weaker correlations (rs = .27, .28; Seto, Harris, Rice, & Barbaree, 2004). Canales, Olver, and Wong (2009) reported that SSPI scores were significantly correlated with arousal to child stimuli and to pubescent male stimuli, but not to pubescent female stimuli. This shows support for the construct validity of the SSPI, especially given that having male victims is given extra weight in the measure. Additionally they found no relationship between SSPI scores and arousal to adult female stimuli, supporting the discriminant validity of the measure.

On the other hand, two other studies using the same sample of 206 contact sex offenders against victims aged 16 or younger found little (Moulden, Firestone, Kingston, & Bradford, 2009) or no (Kingston, Firestone, Moulden, & Bradford, 2007) relationship between SSPI and either phallometric results or diagnostic criteria for pedophilia (DSM-IV-TR; American Psychiatric Association, 2000). These studies, however, included offenders with victims up to and including age 16, which could affect their findings. On the one hand, adding offenders with older victims could increase the discriminative validity by increasing the variability in the

construct of interest (counteracting possible restriction of range issues in pedophilic interests). On the other hand, these studies applied the scale to offenders for whom the scale was not developed and intended for, which could weaken its validity, particularly if they modified the scoring of items to include child victims up to age 16 (which was not clear in these studies). Another possible explanation for the null results obtained by these authors is that they analyzed a dichotomous SSPI variable as opposed to the continuously distributed score, which would result in a loss of information.

Additional support for the construct validity of the SSPI has been found using less direct, reaction time approaches. For example in a meta-analysis of Implicit Association Tests (IATs) of sexual interest in children, Babchishin, Nunes, and Hermann (2013) report a moderate relationship between IATs and SSPI scores (fixed-effect $r = .28$, $k = 3$). Although the three studies examining this relationship had small sample sizes (combined N of 145), the effect size was very similar to the SSPI's relationship with phallometrically assessed deviance (Seto et al., 2004; Seto & Lalumière, 2001) and the relationship was consistent across the three IAT studies (i.e., the Q statistic was non-significant).

Several risk assessment measures incorporate estimates of deviant sexual interest in calculating risk of sexual re-offending. For example, the STABLE-2000 and STABLE-2007 (Hanson, Harris, Scott, & Helmus, 2007) both include an item for deviant sexual interest, which is defined broadly as sexual interest in activities that are illegal, inappropriate, or highly unusual (e.g., sexual interest in children, violence, fetishes). These items combine information from multiple sources including past behavior, self-report, and phallometric testing to estimate the presence or absence of deviant sexual interest.

Nunes and Babchishin (2012) found a non-significant correlation (r = .27, N = 20) between SSPI scores and the deviant sexual interest item from the STABLE-2000. However, they reported a very large and significant correlation (r = .81, N = 13) between SSPI scores and the STABLE-2007 deviant sexual interest item. Although the sample size of both analyses was small, this difference has intuitive appeal because the coding of the deviant sexual interest item was revised for STABLE-2007 to give greater weight to the offense history of the offender (Hanson et al., 2007), which should boost its relationship to the SSPI, which is measured solely by offense history.

Despite the overlap, the SSPI and STABLE-2007 item also have important differences in how they measure deviant sexual interest. For example the STABLE-2007 item does not consider the relationship between the offender and victim, unlike the SSPI. On the other hand, the SSPI does not include self-reported or suspected deviant preferences or the results of specialized testing (such as phallometry); the STABLE-2007 item is broader than pedophilia and so also includes interests or behaviors such as exhibitionism, voyeurism, humiliation of victims, or sex with animals. Taken in combination with evidence from phallometry and indirect measures, there is partial though equivocal support for the construct validity of the SSPI. Further research with larger samples is needed to strengthen this conclusion.

Predictive validity. Compared to the evidence for construct validity, there is much less research examining the predictive validity of the SSPI. In two samples (Ns = 113 and 145), Seto et al. (2004) found that SSPI scores were significantly and positively correlated with risk assessment measures (RRASOR, Static-99, VRAG, SORAG). While the scores were significantly related to violent recidivism (AUCs of .67 and .62), their relationship with sexual recidivism was non-significant in the first sample (AUC = .62; statistical power is generally

lower for sexual recidivism due to fewer recidivists). The second sample had a larger sample and a higher base rate of sexual recidivism than the first (19% versus 8%) and yielded a significant predictive relationship for the SSPI (AUC = .69).

In contrast, using a small sample of 79 offenders with one or more victims age 16 or younger, Canales et al. (2009) did not find the SSPI predictive of being charged (AUC = .47) or re-convicted for a sexual offense (AUC = .48). While Canales et al. (2009) reported a reconviction rate of 28% (n = 35) among their overall sample of 124 sexual offenders, including rapists, it is not clear what the recidivism rate was for offenders against children. Additionally, in a sample of 206 sex offenders against victims age 16 and younger, Moulden and colleagues (2009) found that neither continuous or dichotomous SSPI scores were related to sexual, violent, or any recidivism (correlations ranged between -.11 and -.04; odds ratios ranged between .84 and 1.03). As with the studies on construct validity, increased sample sizes will allow the development of a clearer picture of the ability of the SSPI to predict re-offending among sexual offenders.

Current Study

The SSPI is a promising measure of deviant sexual interest in children. It allows a quick and straightforward assessment based on criminal history information without requiring detailed file reviews, interviews with the offender, or specialized testing. However, further research is needed to better establish its psychometric properties, particularly construct and predictive validity. In the current study, a larger sample than those used in previous research facilitated the examination of the predictive validity of the SSPI in relation to sexual recidivism, both including and excluding sexually motivated technical violations (e.g., being with children alone, using pornography). Additionally, we wanted to use this larger sample to further examine the construct

validity of the SSPI. In addition to comparing the SSPI with the deviant sexual interest items from the STABLE-2000 and the STABLE-2007, we also compared it to the deviant sexual interest subscale of the Static-2002R (Phenix, Doren, Helmus, Hanson, & Thornton, 2008), which incorporates information (similarly to the SSPI) about victim gender, relatedness, and number, but also takes non-contact sexual offenses into account. This study also explored whether the SSPI added incremental predictive validity to previously established risk scales.

Method

Sample

The sample for this study came from the Dynamic Supervision Project (see Hanson et al., 2007). Data on the predictive accuracy of the risk scales and the deviant sexual interests item of the STABLE-2000 and STABLE-2007 have been previously reported; however, this study includes an updated follow-up period and is the first report on the SSPI and other individual static items in this sample. All offenders were adult male sex offenders starting a period of community supervision (probation or parole) in Canada. This study was restricted to male offenders who had at least one child victim less than 14 years old. Consequently, the overall sample available for analysis was 410 offenders, though sample size was further reduced for analyses of some variables, depending on the available information.

On average, offenders were 42.8 years old (SD = 14.4). Half the sample (51%) was classified as extrafamilial child molesters. Approximately one third (34%) had only incest victims, 11% had mixed adult and female victims, and 4% of the sample were offenders with exclusively non-contact offenses against children. Approximately 13% of the offenders self-identified as being of Aboriginal heritage, 12% had been hospitalized overnight for a psychiatric condition, and 6% had previously been diagnosed as developmentally delayed.

Measures

Screening Scale for Pedophilic Interests (SSPI). The SSPI (Seto & Lalumière, 2001) is a measure designed to assess sexual interest in children among individuals who have committed a sex offense against a child (defined as less than 14 years old). It consists of four items assessing previous sex offense child victim characteristics (male victim, unrelated victim, 2+ victims, and victim age 11 or younger), and total scores can range from 0 to 5. For 45 offenders, there was insufficient information to score one item (any victims 11 years old or younger). Consequently, a SSPI total score was calculated for offenders with complete information on all items. For the 45 offenders with one missing item, SSPI scores were also computed, assuming a score of 0 for the missing item; SSPI analyses using these cases are referred are hereafter referred to as approximated SSPI scores.

Static-99R (Hanson & Thornton, 2000; Helmus, Thornton, Hanson, & Babchishin, 2012). Static-99R is an empirically derived actuarial risk assessment tool designed to predict sexual recidivism in adult male sex offenders (see also www.static99.org) and is the most commonly researched actuarial risk scale for sex offenders (Hanson & Morton-Bourgon, 2009). The scale has ten items assessing criminal history, victim characteristics, and relationship history. The total score (ranging from -3 to 12) is calculated by summing all item points. Items in this scale that may be indicative of deviant sexual interest include having a male victim or having an unrelated victim. Although these items have the same wording as the two SSPI items, they are scored slightly differently in that the Static-99R items consider all sex offense victims, whereas the SSPI items are restricted to child victims.

Static-2002R (Hanson & Thornton, 2003; Helmus et al., 2012). Static-2002R is also an empirical actuarial risk assessment tool for adult male sex offenders (see also

www.static99.org). Although it has less research support than the Static-99R, it was intended to have greater construct validity and has a subscale designed to assess deviant sexual interest (Hanson & Thornton, 2003). The scale has 14 items grouped into 5 main subscales: age at release (1 item), persistence of sex offending (3 items), sexual deviance (3 items), relationship to victims (2 items), and general criminality (5 items). The sexual deviance scale includes three items: having a non-contact sex offense, having a male victim, and whether the offender has at least two victims less than 12 years old, at least one of whom must be an unrelated victim. The total score for Static-2002R can range from -2 to 13. Similar to Static-99R and unlike the SSPI, victim items are scored for all victims unless the item specifies the age of the victims.

STABLE-2007 (Hanson et al., 2007). The STABLE-2007 is an actuarial risk scale designed to assess risk of sexual recidivism using putatively dynamic risk factors. It is based on the STABLE-2000 (see Hanson et al., 2007), but was modified to remove three items and alter three items. The Deviant Sexual Interest item in this scale was modified from the 2000 version to increase the structure of the scoring criteria to better incorporate offense history as an indicator of deviant sexual interests. In this version, the evaluator uses the coding manual guidelines to rate the offender as a 0, 1, or 2 in the following domains: total number of sex offense victims, number of deviant victims or activities involved in sex offenses (e.g., child victims, exhibitionism, sadism), evidence of deviant sexual interest (e.g., self-report, history, or other evidence), and results of specialized testing. The offender's score on the Deviant Sexual Interest item is the highest score in any of the domains. The total score on STABLE-2007 is a sum of all items and can range from 0 to 26.

STABLE-2000 Deviant Sexual Interest Item (Hanson et al., 2007). The original Deviant Sexual Interest item of the STABLE-2000 was also analyzed in order to further examine

the construct validity of the SSPI. Compared to the STABLE-2007 version, it can be considered more of an unstructured professional judgement about the level of deviant sexual interest, ideally incorporating diverse types of information (e.g., self-report, specialized testing, behavioral history).

Procedure

Supervising officers (e.g., probation or parole officers) completed and submitted assessments on the Static-99 and STABLE-2000 as part of their routine supervision practices. They also submitted general descriptive information of the offender's current and previous sex offense victims (e.g., age, gender, relationship to offender). This information (along with national criminal history records) was used to calculate additional items to obtain scores for the SSPI, Static-99R, Static-2002R, and STABLE-2007. Supervising officers attended a 2-day training session, although in rare cases, officers submitted data who had been trained by apprenticing with other local officers. Formal interrater reliability information is not available, but scores on training exercises and selective site visits suggested reasonable understanding and application of the scoring rules (for more information, see Hanson et al., 2007).

Recidivism

Analyses of predictive validity examined two recidivism outcomes. Sexual recidivism included any sexually motivated offense (this was determined based on the offense behavior and not the name of the charge or conviction). Sexual recidivism with breaches included all sex offenses as well as official sanctions for sexually motivated violations of the conditions of community supervision. Including sexual recidivism with breaches allows for consideration of a broader range of behavior that could be considered high risk for committing a sexual offense and may counteract some of the inherent biases in recidivism research due to low detection rates of

sex offenses (e.g., Craissati, Bierer, & South, 2011). In this sample, there were 12 offenders who were coded as having a sexually motivated breach who did not also have a sexual recidivism event. For all but three offenders, the breach was for having unsupervised contact with children or loitering around schoolyards and youth centers. The other three offenders engaged in the following behaviors: one offender was with a prostitute, one was caught following his former victim, and another offender engaged in a series of behaviors suggestive of grooming several young males at a children's center (including buying them candy, cigarettes, offering to buy drugs, and putting his arm around the children).

Analyses of discriminant validity examined two other types of recidivism. Non-sexually violent recidivism was defined as any offense that involved deliberate harm to an identifiable person (or threat of harm), which could not be classified as sexually motivated. Non-violent recidivism included any criminal behavior (excluding technical offenses) that was not classified as sexual or non-sexual violence.

Recidivism information was gathered from reviews of provincial and national criminal history records, as well as from supervising officers, local police jurisdictions, and searches of newspaper databases. Recidivism was considered to have occurred if the agency reporting the information believed that the offense occurred. For breaches, however, an official record of parole revocation or a new conviction for violation of conditional release was required. Given that criminal history records were the major source of recidivism information, the vast majority of recidivism events were linked to an officially recorded charge or conviction. In all but a few cases, information was obtained on the date and circumstances of new violent offenses (e.g., from the supervising officer or from police jurisdictions). This process also allowed for the

removal of cases of pseudo-recidivism (i.e., new charges/convictions for offenses that were committed before the offender's most recent sex offense)

The follow-up period was calculated from the date that the first assessment information was collected to the date of the last recidivism information received (or until death or deportation). For the few cases that did not appear on any official criminal record, the follow-up end date was set one month after the last assessment information was received. The offender start dates ranged from February 2001 to October 2005. Follow-up ranged from 0.2 to 10.0 years (M = 7.5, SD = 2.0, Mdn = 8.0).

Overview of Analyses

Correlations were measured as Kendall's tau as they are most appropriate for ordinal data (Field, 2009). Confidence intervals were calculated following the formulas presented in Bonett and Wright (2000). Predictive accuracy was measured using the Area Under the Receiver Operating Characteristic Curve (AUC; Swets, Dawes, & Monahan, 2000). The AUC can vary between 0 and 1, with .50 indicating the level of prediction that would be expected by chance, and AUCs above .50 demonstrating positive predictive accuracy (i.e., higher scores on a measure are associated with a higher likelihood of reoffending). AUCs can be interpreted as the probability that a randomly selected recidivist would have a higher score than a randomly selected non-recidivist. As a heuristic, an AUC of .56 corresponds to a small effect size, while .64 reflects a moderate effect, and .71 reflects a large effect size (Rice & Harris, 2005). An AUC value is statistically significant if the 95% confidence interval does not include .50.

To compare whether AUC values for predictors were different from each other, the DeLong test (DeLong, DeLong, and Clarke-Pearson, 1988) was used, which takes into account the correlation between the predictors. This test examines the absolute difference between two

AUC values and calculates a confidence interval for the difference score. The null hypothesis is that the difference should be 0, and a 95% confidence interval that does not include 0 indicates the two effect sizes are significantly different.

Analyses of incremental predictive accuracy used Cox regression (Allison, 1984), which estimates predictive accuracy (through hazard ratios) associated with one or more predictor variables from survival data with unequal follow-up times. When more than one predictor is entered in the model, the results are interpreted similar to a multiple linear regression. The hazard ratios reflect the accuracy of the predictor, after controlling for other predictors in the model. In other words, it reflects the unique predictive value of the scale, above and beyond the other scales included in the analysis. The hazard ratio indicates the relative risk of the negative outcome given a one-point change in the predictor variable, averaged across the follow-up period. For example, a hazard ratio of 1.20 indicates that the likelihood of recidivism increases by 20% for every unit increase in the predictor, averaged across the follow-up period.

Results

For the overall sample (N = 410), the sexual recidivism rate was 8.8% (n = 36). Including sexually motivated technical violations, the recidivism rate was 11.7% (n = 48).

Construct Validity

Total scores on the SSPI were moderately related to the STABLE-2000 deviant sexual interests item (r = .27, 95% CI of .20 to .34, n = 303) and strongly related to the STABLE-2007 deviant sexual interests item (r = .55, 95% CI of .49 to .60, n = 303). Including the approximated SSPI total scores (could include some missing information) yielded the same pattern (with STABLE-2000 item, r = .25, 95% CI of .18 to .32, n = 336; with STABLE-2007 item, r = .51, 95% CI of .45 to .56, n = 336).

Although it had not yet been examined in previous research, we also explored the relationship between the SSPI and the Static-2002R deviant sexual interests subscale. This subscale was very strongly related to both complete SSPI scores (r = .70, 95% CI of .66 to .74, n = 349) and the approximated SSPI score (r = .68, 95% CI of .64 to .71, n = 387). Examining in isolation the Static-2002R item that was intended to be an indicator of pedophilia (this item assigns one point for having two or more victims less than 12 years old, at least one of whom is unrelated; Phenix et al., 2009), this item was also strongly related to both the complete SSPI score (r = .57, 95% CI of .52 to .61, n = 349) and the approximated SSPI score (r = .55, 95% CI of .50 to .60, n = 387), though correlations were notably lower than with the full Static-2002R sexual deviance subscale.

Discriminant Validity

Table 1 presents the correlations between the SSPI and items of risk scales that are more likely to be measuring antisociality or general criminality. Both the complete and approximated SSPI scores were unrelated to non-sexual violence (either as part of their current sex offense or a prior conviction), prior sentencing dates (on Static-99R), cooperation with supervision, impulsive acts, and any conviction for Break and Enter (rs ranged between -.04 and .07). Both the complete and approximated SSPI scores were, however, significantly related to prior sentencing occasions for any type of offense (on Static-2002R), poor cognitive problem-solving, and negative emotionality/hostility (rs ranged between .08 and .18). The complete SSPI score was also significantly related to having a breach of conditional release (r = .08). Although some of these correlations were significant, the highest correlation (r = .18) was still lower than the relationship between SSPI and the STABLE-2007 deviant sexual interests item (rs between .25

and .27) and substantially lower than the other potential measures of deviant sexual interest (rs of .51 and higher).

The SSPI total score was also not significantly predictive of non-sexually violent (AUC = .531, 95% CI of .439 to .623) and non-violent recidivism (AUC = .505, 95% CI of .419 to .590). Similar results were found for the approximated SSPI score (for non-sexual violence, AUC = .544, 95% CI of .455 to .633; for non-violent recidivism, AUC = .520, 95% CI of .440 to .601). These findings suggest that the SSPI is more strongly related to items measuring deviant sexual interests than general criminality/antisociality.

Predictive Validity

Table 2 presents the predictive accuracies of the SSPI total scores and items. For comparative purposes, predictive accuracies are also presented for other scale items that are indicative of deviant sexual interest, and total scores on the risk scales. Most AUCs (13 out of 15) are higher for predicting sexual recidivism that included sexually motivated breaches compared to sexual recidivism without these technical violations.

The complete SSPI total score was significantly related to both sexual recidivism (AUC = .62) and sexual recidivism with breaches (AUC = .64); these values approached or reached a moderate effect size. Incorporating the additional 45 cases with missing information on one item resulted in significant and slightly higher predictive accuracy (AUC = .64 for sexual recidivism and AUC = .66 for sexual recidivism including breaches).

On their own, the individual SSPI items demonstrated a small relationship with sexual recidivism (including or excluding sexual breaches; AUCs at or below .64). The strongest item was having two or more child victims, which was significantly related to both outcomes. Having an unrelated victim was significant only for sexual recidivism including breaches (AUC = .62);

the remaining items did not reach significance, but were all predicting in the expected direction. The lowest effect sizes were found for having a child victim age 11 or under, which was trivially related to both outcomes (AUC = .52).

Similarly, the other indicators of deviant sexual interests showed small relationships with the outcomes. Both outcomes were significantly predicted by the unrelated victim item of Static-99R, the deviant sexual interests subscale of the Static-2002R, and the deviant sexual interests item of the STABLE-2007 (but not the STABLE-2000). The deviant sexual interest subscale of Static-2002R had pretty comparable predictive accuracy to the SSPI, whereas the deviant sexual interests item of STABLE-2007 had slightly lower accuracy than the SSPI and Static-2002R subscale.

The DeLong test was used to compare the SSPI scale (both the total score and approximated score) to the four other global indicators of deviant sexual interests: the STABLE-2000 and STABLE-2007 items, as well as the Static-2002R item for having multiple child victims under 12, at least one of whom is unrelated, and lastly, the Static-2002R deviant sexual interests subscale score. Analyses examined both recidivism outcomes (analyses not reported but are available upon request). None of the differences were significant, indicating that all these items had comparable predictive accuracy.

AUCs were also presented for the total risk scale scores, for comparison purposes. All AUCs were significant and moderate to large (AUCs varied between .68 and .78). All scale AUCs were meaningfully higher than the predictive accuracies for individual items/measures of deviant sexual interests, demonstrating that although deviant sexual interests is an important risk-related construct, it is not sufficient information for a comprehensive assessment of risk.

Table 3 presents the results of the Delong tests comparing the total SSPI score with the risk scales. For the prediction of sexual recidivism both including and excluding technical violations, the SSPI had significantly lower predictive accuracy than Static-99R and Static-2002R, but did not significantly differ from the STABLE-2007. The same pattern of results was found comparing the approximated SSPI score with these risk scales (analyses available upon request).

Incremental Accuracy with Risk Scales

Cox regression analyses were used to test whether the SSPI scale added incremental predictive accuracy to any of the risk scales: Static-99R, Static-2002R, and STABLE-2007 (Table 4). In other words, does the SSPI add unique predictive information above and beyond what is already being measured by these risk scales? The SSPI did not add incremental predictive accuracy to either Static-99R or Static-2002R, suggesting that the construct measured by the SSPI is already sufficiently measured in these scales.

The SSPI did add significant incremental validity to the STABLE-2007 for predicting sexual recidivism with breaches (but not for predicting sexual recidivism without the sexually motivated breaches). In other words, for predicting sexual recidivism including breaches, accuracy could be improved by considering SSPI in addition to the STABLE-2007.

Substituting SSPI Scores for STABLE-2007 Deviant Sexual Interests Item

Given that predictive accuracy was comparable for the different indicators of deviant sexual interests and that the SSPI score added incrementally to the STABLE-2007 for predicting sexual recidivism with breaches, the next set of analyses explored whether the SSPI could be used to substitute for the deviant sexual interests item in the STABLE-2007. Though this is unlikely to occur in real-world assessments, it may be common in research contexts to score the

STABLE-2007 from archival data with limited information about deviant sexual interest (save for offense history information).

This required first converting the SSPI total score into a score on a 0-2 scale, to match the metric of the other STABLE-2007 items. Table 5 presents the sample sizes and recidivism rate for each SSPI total score (n = 365). Table 5 suggests that the most meaningful cut-points on the SSPI to create a trichotomous item were as follows: 0 = 0 (offenders with this score had a sexual recidivism rate of 2.6%), 1 to 3 = 1 (offenders with these scores had a sexual recidivism rate of 8.5%), 4 to 5 = 2 (offenders with these scores had a sexual recidivism rate of 19.3%).

Table 6 presents AUC values for the trichotomous SSPI item, as well as the original STABLE-2007 score and the STABLE-2007 score with the SSPI trichotomous item substituted for the deviant sexual interests item. Trichotomizing the SSPI total score reduced its AUCs only to a small extent compared to the full SPPI scores (for sexual recidivism, AUC went from .62 to .61; for sexual recidivism with breaches, the AUC went from .64 to .62; both AUCs remained significant). Substituting the SSPI trichotomous item for the deviant sexual interests item resulted in trivial increases in the accuracy of the scale (the absolute difference in the AUCs was < .01). Differences in accuracy between the original STABLE-2007 scale and the scale with the SSPI substitution were non-significant (analyses available upon request).

Discussion

In summary, the relationship between the SSPI and scale items/subscales assessing deviant sexual interest from the STABLE-2000, STABLE-2007, and Static-2002R lend support to the construct validity of the SSPI. While common variance due to common variables undoubtedly contributed to the relationships between the SSPI and these items, its strong relationship with the STABLE-2007 item in particular, which includes additional sources of

information (as well as its much smaller relationships with measures of antisociality), is a promising indicator of construct validity.

In terms of its ability to predict sexual recidivism (including or excluding sexually motivated breaches), the SSPI yields significant AUCs and demonstrated comparable accuracy to other indicators of deviant sexual interest. This is encouraging and supports the predictive validity of the scale given the limited and mixed evidence about this question. The effect sizes in the current study were consistent with a previous study (Seto et al., 2004), although the current study had a larger sample size and greater statistical power.

The SSPI did not, however, add incrementally to the predictive ability of the Static-99R or to the Static-2002R risk scales. This suggests that the Static scales are adequately incorporating deviant sexual interest in children. However the SSPI did add incrementally to the STABLE-2007 for the prediction of sexual recidivism including breaches, suggesting that this scale may benefit by giving greater weight to deviant sexual interests, at least for sex offenders against children (this study cannot address whether the scale adequately incorporates this construct for other types of offenders, given the items all pertain to child victims).

Coding missing items in the SSPI as 0 allowed us to increase our sample size, and thus our statistical power, without affecting the overall patterns of findings. We also found that for researchers who might only have limited file data, but who nonetheless wish to calculate STABLE-2007 scores, mapping the SPPI scores onto a 0-2 scale and substituting for the deviant sexual interest item maintained the predictive accuracy of the scale. However, we do not recommend this substitution for applied risk assessments (given that the item is intended to measure the construct using more diverse sources of information, including those that could be more dynamic in nature). However, for research purposes, if the complete information is not

available, our analyses suggest the SSPI is a reasonable substitute in research with sex offenders with child victims.

The SSPI demonstrated significantly less predictive accuracy than Static-99R and Static-2002R (but not the STABLE-2007). This is unsurprising as recidivism risk is predicted by a lot more than deviant sexual interest. Antisociality indicators play an important role as well, as reflected in variables such as younger offender age and prior criminal history (Hanson & Morton-Bourgon, 2005). While future refinement of the SSPI may improve its predictive ability somewhat, this will always be bounded by the absolute contribution of deviant interests to future re-offending.

It should be noted that sexual recidivism including breaches is a broader outcome than other studies typically examine because it includes high-risk behavior that is not an actual sex offense. However, the low sexual recidivism rate substantially reduced statistical power. Including the sexually motivated breaches increased statistical power while still examining a similar type of behavior (even though it does not meet the legal threshold of a sex offense), and a behavior that was illegal for the offenders in the sample. Future research using considerably larger samples may consider the ability of the SSPI to predict different types of sex offenses, as the specific nature of the scale as an index of pedophilic interest suggests it should be a better predictor of sexual offenses against children, compared with rape, for example. Given the low base rate of sexual reoffending overall, this was impossible to explore in the current data.

Another interesting question we could not address in this study is whether this predictive relationship was due to the greater statistical power to detect an association with a higher base rate outcome that included breaches, or whether the SSPI may be able to differentially predict breaches that suggest sexually motivated behavior such as being with children without a

responsible adult. Follow-up research focusing on the ability of the SSPI and other risk-related measures to specifically predict breaches is needed to answer this question.

An important limitation of the current study is that it used archival data, which restricted the variables that were available for explorations of construct validity. In particular, the indicators of deviant sexual interest were limited to items from risk assessment scales. The inclusion of additional measures of sexual interest (e.g., phallometric data, viewing time, implicit association tasks) would have provided stronger support for the construct validity of the SSPI.

Implications

Taken together, our results build on previous literature demonstrating the construct and predictive validity of the SSPI. We conclude that the SSPI is useful as an brief measure of sexual interest in children. It performs equally as well as indicators of deviant sexual interests from other risk scales. Additionally, for research purposes, the SSPI could be a reasonable substitute for the STABLE-2007 deviant sexual interests item if insufficient information was available to score the STABLE-2007 item as intended.

References

Allison, P. D. (1984). Event history analysis: Regression for longitudinal event data. Beverly Hills, CA: Sage.

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders: DSM-IV-TR (4th ed.). Washington, DC: American Psychiatric Association.

Babchishin, K. M., Nunes, K. L., & Hermann, C. A. (2013). The validity of Implicit Association Test (IAT) measures of sexual attraction to children: A meta-analysis. Archives of Sexual Behavior, 42, 487-499. doi:10.1007/s10508-012-0022-8

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating Pearson, Kendall and Spearman correlations. Psychometrika, 65, 23-28. doi:10.1007/BF02294183

Canales, D. D., Olver, M. E., & Wong, S. C. P. (2009). Construct validity of the Violence Risk Scale—Sexual Offender Version for measuring sexual deviance. Sexual Abuse: A Journal of Research and Treatment, 21, 474-492. doi:10.1177/1079063209344990

Craissati, J., Bierer, K., & South, R. (2011). Risk, reconviction and "sexually risky behavior" in sex offenders. Journal of Sexual Aggression, 17, 153-165. doi:10.1080/13552600.2010.490306

Hanson, R. K., Harris, A. J., Scott, T.-L., & Helmus, L. (2007). Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project (Corrections Research User Report No. 2007-05). Ottawa, ON: Public Safety and Emergency Preparedness Canada. Retrieved from http://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ssssng-rsk-sxl-ffndrs/index-eng.aspx

Hanson, R. K., & Morton-Bourgon, K. E. (2004). Predictors of sexual recidivism: An updated meta-analysis (Corrections Research User Report No. 2004-02). Ottawa, ON: Public

Safety and Emergency Preparedness Canada. Retrieved from http://www.static99.org/pdfdocs/hansonandmortonbourgon2004.pdf

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. Psychological Assessment, 21, 1-21. doi:10.1037/a0014421

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. Law and Human Behavior, 24, 119-136. doi:10.1023/A:1005482921333

Hanson, R. K., & Thornton, D. (2003). Notes on the development of Static-2002. (Corrections Research User Report No. 2003-01). Ottawa, ON: Department of the Solicitor General of Canada. Retrieved from http://www.static99.org/pdfdocs/HansonThornton2003.pdf

Harris, G. T., Lalumière, M. L., Seto, M. C., Rice, M. E., & Chaplin, T. C. (2012). Explaining the erectile responses of rapists to rape stories: The contributions of sexual activity, non-consent, and violence with injury. Archives of Sexual Behavior, 41, 221-229. doi:10.1007/s10508-012-9940-8

Helmus, L., Thornton, D., Hanson, R. K., & Babchishin, K. M. (2012). Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: Revised age weights. Sexual Abuse: A Journal of Research and Treatment, 24, 64-101. doi:10.1177/1079063211409951

Kingston, D. A., Firestone, P., Moulden, H. M., & Bradford, J. M. (2007). The utility of the diagnosis of pedophilia: A comparison of various classification procedures. Archives of Sexual Behavior, 36, 423-436. doi:10.1007/s10508-006-9091-x

Moulden, H. M., Firestone, P., Kingston, D., & Bradford, J. (2009). Recidivism in pedophiles: an investigation using different diagnostic methods. Journal of Forensic Psychiatry & Psychology, 20, 680-701. doi:10.1080/14789940903174055

Nunes, K. L., & Babchishin, K. M. (2012). Construct validity of Stable-2000 and Stable-2007 scores. Sexual Abuse: A Journal of Research and Treatment, 24, 29-45. doi:10.1177/1079063211404921

Phenix, A., Doren, D., Helmus, L., Hanson, R. K., & Thornton, D. (2008). Coding rules for Static-2002. Ottawa, ON: Public Safety Canada. Retrieved from http://www.static99.org/pdfdocs/static2002codingrules.pdf.

Rice, M., & Harris, G. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. Law and Human Behavior, 29, 615-620. doi:10.1007/s10979-005-6832-7

Seto, M. C. (2008). Pedophilia and sexual offending against children: Theory, assessment, and intervention. Washington, DC, US: American Psychological Association.

Seto, M. C., Harris, G. T., Rice, M. E., & Barbaree, H. E. (2004). The Screening Scale for Pedophilic Interests predicts recidivism among adult sex offenders with child victims. Archives of Sexual Behavior, 33, 455-466. doi:10.1023/B:ASEB.0000037426.55935.9c

Seto, M. C., & Lalumière, M. L. (2001). A brief screening scale to identify pedophilic interests among child molesters. Sexual Abuse: A Journal of Research and Treatment, 13, 15-25. doi:10.1177/107906320101300103

Snowden, R. J., Craig, R. L., & Gray, N. S. (2011). Indirect behavioral measures of cognition among sexual offenders. Journal of Sex Research, 48, 192 - 217. doi:10.1080/00224499.2011.557750

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological Science Can Improve Diagnostic Decisions. Psychological Science in the Public Interest, 1(1), 1-26. doi:10.1111/1529-1006.001

Table 1

Kendall's Tau Correlations Between SSPI and Measures of Antisociality/General Criminality

	Total SSPI Score				Approximated SSPI Score			
Item	r	95% CI		n	r	95% CI		n
Static-99R: Index non-sexual violence	.04	-.02	.11	365	.03	-.03	.10	410
Static-99R: Prior non-sexual violence	.02	-.04	.09	365	-.01	-.07	.06	410
Static-99R: Prior sentencing dates	.01	-.06	.07	365	-.01	-.08	.05	410
Static-2002R: Prior sentencing occasions	**.11**	**.04**	**.18**	**349**	**.08**	**.02**	**.15**	**387**
Static-2002R: Breach of conditional release	**.08**	**.01**	**.15**	349	.05	-.02	.12	387
STABLE: Cooperation with supervision	.07	-.01	.14	303	.05	-.02	.12	336
STABLE: Impulsive acts	-.01	-.08	.07	303	.02	-.05	.09	336
STABLE: Poor cognitive problem-solving	**.09**	**.02**	**.16**	**303**	**.08**	**.01**	**.15**	**336**
STABLE: Negative emotionality/hostility	**.17**	**.10**	**.24**	**303**	**.18**	**.10**	**.24**	**336**
Risk Matrix 2000: Break and Enter convictions	-.02	-.08	.05	347	-.04	-.10	.03	385

Note. Bold values denote significant correlations ($p < .05$)

Table 2

Predictive Accuracy of Deviant Sexual Interest Items and Scales, and Risk Scales

		Sexual Offense Recidivism			Sexual Recidivism with Breaches		
	N	AUC	95% CI		AUC	95% CI	
SSPI Items and Total Scores							
SSPI male victim	410	.559	.456	.662	.565	.475	.656
SSPI 2+ victims	410	**.604**	**.506**	**.702**	**.635**	**.550**	**.720**
SSPI unrelated victim	410	.589	.496	.681	**.617**	**.538**	**.696**
SSPI victim age 11 or younger	365	.522	.422	.621	.516	.427	.604
SSPI Total	365	**.621**	**.522**	**.719**	**.641**	**.557**	**.726**
SSPI Approximated Total	410	**.635**	**.541**	**.728**	**.661**	**.582**	**.740**
Other Deviant Sexual Interest Items/Subscales							
Static-99R unrelated victim	410	**.629**	**.545**	**.712**	**.637**	**.564**	**.710**
Static-99R male victim	410	.566	.463	.669	.569	.478	.659
Static-2002R 2+ victims <12, at least one unrelated	387	.596	.492	.701	**.607**	**.514**	**.699**
Static-2002R Deviant Sexual Interest Subscale Score	387	**.634**	**.527**	**.740**	**.631**	**.538**	**.725**
Stable-2000 Deviant Sexual Interests	336	.575	.459	.690	.599	.498	.700
Stable-2007 Deviant Sexual Interests	336	**.615**	**.504**	**.725**	**.627**	**.533**	**.721**
Risk Scale Total Scores							
STABLE-2007	336	**.684**	**.581**	**.788**	**.709**	**.622**	**.796**
Static-99R	410	**.769**	**.688**	**.851**	**.771**	**.700**	**.843**
Static-2002R	387	**.773**	**.697**	**.848**	**.780**	**.715**	**.846**

Note. Bolded values denote AUCs that are statistically significant ($p < .05$).

Table 3

Delong Tests Comparing the SSPI to Other Risk Scales

Comparison	n	Difference in AUCs	95% CI		χ^2	p
Sex Recidivism						
SSPI – Static-2002R	349	-.143	-.236	-.050	9.16	.002
SSPI – Static-99R	365	-.146	-.239	-.053	9.52	.002
SSPI – STABLE-2007	303	-.061	-.168	.047	1.23	.267
Sex Recidivism Including Breaches						
SSPI – Static-2002R	349	-.133	-.212	-.053	10.75	.001
SSPI – Static-99R	365	-.125	-.206	-.044	9.22	.002
SSPI – STABLE-2007	303	-.048	-.146	.051	0.90	.343

Table 4

Incremental Validity of SSPI with Risk Scales

		Sexual Recidivism					Sexual Recidivism Including Breaches				
	N	Hazard Ratio	95% CI		Wald	p	Hazard Ratio	95% CI		Wald	p
Static-99R	365	1.473	1.268	1.711	25.78	<.001	1.427	1.254	1.623	29.23	<.001
SSPI Complete Score		.959	.740	1.243	0.10	.750	1.045	.837	1.305	0.15	.696
Static-2002R	349	1.397	1.200	1.626	18.53	<.001	1.396	1.225	1.591	24.97	<.001
SSPI Complete Score		.921	.696	1.219	0.33	.566	.969	.763	1.231	0.07	.798
STABLE-2007	303	1.105	1.029	1.187	7.53	.006	1.114	1.047	1.185	11.75	.001
SSPI Complete Score		1.161	.894	1.506	1.26	.263	1.252	1.001	1.566	3.86	.049

Table 5

Sample Sizes and Recidivism Rates for Each Total Score on the SSPI

SSPI Score	N recid	N total	Recid Rate (%)
0	1	38	2.6
1	10	128	7.8
2	8	80	10.0
3	5	62	8.1
4	5	29	17.2
5	6	28	21.4
Total	35	365	9.6

Note. The overall recidivism rate in this table differs from what is presented in the results section because this table was restricted to cases with no missing information on the SSPI (n = 365).

Table 6

Predictive Accuracy of STABLE-2007 with the SSPI Item Substituting for Deviant Sexual Interests

		Sexual Offense Recidivism			Sexual Recidivism with Breaches		
Predictor	N	AUC	95% CI		AUC	95% CI	
SSPI 3-point item	303	**.609**	**.503**	**.715**	**.622**	**.529**	**.715**
Original STABLE-2007	303	**.677**	**.572**	**.783**	**.702**	**.612**	**.793**
STABLE-2007 with SSPI substitute	303	**.685**	**.582**	**.788**	**.708**	**.619**	**.797**