Running head: SEXUAL INTEREST PROFILES

Indirect Measures of Sexual Interest in Child Sex Offenders: A Multi-Method Approach

Rainer Banse and Alexander F. Schmidt

University of Bonn

Jane Clarbour

University of York

Abstract

Although there is strong meta-analytical evidence that deviant sexual interest in children is a major risk factor for recidivism in child-sex offenders, the assessment of deviant sexual interest with self-report or phallometric measures is problematic. As an alternative approach for assessment, the Explicit and Implicit Sexual Interest Profile (EISIP) is introduced that features direct self-report and indirect latency-based measures (Implicit Association Tests and viewing time measures) of sexual interest in adults and children. The reliability and validity of the EISIP was investigated using a selected sample of child sex offenders (n = 38), offender (n = 37) as well as non-offender (n = 38) controls. Among the indirect measures, viewing time measures showed higher reliability, convergent, and criterion validity than the IATs. However, the IATs independently accounted for criterion variance in multivariate analyses. The combined indirect measures showed good discriminative validity between child-sex offenders and controls.

(146 words)

Key words:

Indirect Measures of Sexual Interest in Child Sex Offenders: A Multi-Method Approach

It is a common view – not only among forensic researchers – that sexually deviant behaviors result from a favor of these over socially accepted sexual activity. This so called sexual preference hypothesis (Freund & Blanchard, 1989) is corroborated by meta-analytic evidence consistently showing that deviant sexual interest (e.g., in sex with children) is one of the strongest risk factors for reoffending with effect sizes around d = .30 (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005). Recent aetiological models of sex offending (e.g., Ward & Beech, 2006; Ward & Siegert, 2002) postulate that among numerous clinical problems such as emotional dysregulation, social difficulties, and cognitive distortions, deviant sexual interest is a primary causal factor for sex offences.

Assessing deviant sexual interest

Albeit "assessing the nature of the individual's deviant sexual interests is often the centerpiece of a sex offender evaluation" (Lanyon, 2001, p. 257) the assessment of enduring sexual preference is fraught with difficulties, mainly due to the problematic psychometric properties of the most commonly used measures (Kalmus & Beech, 2005). This has led to general skepticism about the utility of assessing deviant sexual interest at all (Marshall & Fernandez, 2003). Because neither the legal (based on sexual offences) nor the clinical approach (diagnosis of pedophilia) allow for a valid inference on deviant sexual interest (Marshall, 2007), conceptually more valid assessment tools are needed. Up to now a range of quantitative, psychometric measures have been developed that can be divided into direct and indirect measures.

Direct measures of sexual interest rely on the self-report of deviant sexual preferences in questionnaires or card-sort procedures. Their validity is jeopardized by impression management and deliberate faking. The general problem of transparency in

direct measures is all the more critical if disclosure of personal information is highly embarrassing, socially undesirable, or has legal implications, as it is commonly the case in forensic contexts (Kalmus & Beech, 2005).

To tackle the issues of socially desirable responding and simulation/dissimulation, great effort has been invested in developing indirect measures of sexual interest. These are supposed to be less susceptible to deliberate manipulation because test subjects are a) less aware of the nature of the measure or the measurement principle, and/or b) because the expression of the measured construct cannot be deliberately controlled. Based on their methodological rationale, indirect measures either rely on physiological measures of sexual arousal, or response latency measures reflecting information processing (in a wide sense).

Among the physiological measures penile plethysmography (PPG) or phallometry is by far the most researched method and commonly used. An index of deviant sexual interest derived from PPG assessment has repeatedly been shown to predict sexual recidivism (Hanson et al., 1998; 2005). However, detailed reviews of the PPG literature come to the conclusion that the main problems of PPG research and application result from a) a lack of standardization of the procedures and stimulus materials, b) low retest reliability, c) low specificity or discriminant validity, d) low response rates, and e) high fakeability (e.g., Kalmus & Beech, 2005; Laws, 2003; Marshall & Fernandez, 2000; Murphy & Barbaree, 1994).

In the last decade, a whole range of latency-based measures have been developed in the very active area of implicit social cognition research. In recent reviews of indirect measures these approaches have been classified as "attentional methodologies" (e.g., Gress & Laws, 2009; Kalmus & Beech, 2005). However, for many indirect paradigms the underlying processes are either unknown or are still subject of debate (e.g., Imhoff, Schmidt, Nordsiek, Luzar, Young, & Banse, in press). Therefore we propose to use the

label *latency-based measures* for this class of instruments that refers to the nature of the dependent variable and does not require any theoretical assumptions about underlying processes.

Latency-based measures are usually based on rating, sorting, or detection tasks that involve different classes of pictorial stimuli of potential sexual interest (e.g., adult women, adult men, and children of either sex). The measurement rationale of indirect paradigms either relies on the fact that certain categories are more strongly associated with the concept of sex or sexual arousal than others (the Implicit Association Test or IAT; e.g., Gray, Brown, MacCulloch, Smith, & Snowden, 2005), or that task irrelevant sexually preferred stimuli function as distractors that interfere with a primary task, such as the Choice Reaction Time Task (e.g., Mokros, Dombert, Osterheider, Zappalà, & Santtila, in press), the Emotional Stroop Task (Smith & Waterman, 2004) or the Attentional Blink Task (Beech, Kalmus, Tipper, Baudouin, Flak, & Humphreys, 2008).

Viewing time (VT) measures – first described by Rosenzweig (1942) – exploit the fact that photographs of sexually attractive individuals are inspected longer than sexually less attractive individuals (e.g., Gress, 2005). In a typical viewing time task, participants are asked to rate the sexual attractiveness of targets of both sexes and different age groups. In addition to the explicit rating, the viewing or inspection time of each stimulus picture is unobtrusively recorded, averaged for each target category, and used as an indirect indicator of sexual interest. Viewing time measures have been used in forensic contexts to successfully differentiate between child sex offenders and non-offenders (Laws & Gress, 2004; Harris; Rice, Quinsey, & Chaplin, 1996) and between different types of sex offenders (Abel, Huffmann, Warberg, & Holland, 1998; Worling, 2006).

Current study

Despite the growing interest in indirect measures of sexual interest and the recent development of a range of conceptually different latency-based measures, there seems to be very little research comparing the reliability and criterion validity of different latency-based measures of sexual interest. To the best of our knowledge, all published studies include comparisons between and among different combinations of direct measures, behavioral offence data, and/or PPG measures (e.g., Abel, Jordan, Hand, Holland, & Phipps, 2001; Letourneau; 2002), but as yet there are no published empirical studies comparing different latency-based measures of sexual interest in forensic samples. However, empirical data of this kind are required to investigate to what extent latency-based measures could improve the assessment of sexual preference over and above direct measures.

In order to address these issues, the general objective of the present study was to introduce and validate the *Explicit and Implicit Sexual Interest Profile* (EISIP) featuring four direct self-report measures of sexual interest, three different IATs, and four VT measures. It was the first aim of the present study to investigate the reliability and convergent validity of the direct and indirect measures of the EISIP. Second, as a more direct test of the discriminant validity of the EISIP, it was explored to what extent the different measures contributed to a discrimination between child sex-offenders and controls. Third, the incremental validity of the measures featured in the EISIP was investigated by contrasting the rates of correct classification and ROC-analyses. Fourth and finally, all EISIP measures were validated against the Screening Scale for Pedophilic Interests (Seto & Lalumière, 2001).

Although not all child sex offenders show deviant sexual preferences and not all men showing deviant sexual preferences do actually offend against children (Seto, 2008), we expected that child sex offenders will show stronger sexual interest in children on direct and indirect measures than controls. Moreover, sex offenders' preferences should be related

to the gender profile of their victims – girl-only and boy-only offenders would be expected to show relatively stronger sexual interest in targets corresponding to the gender of their victims. Offenders with mixed victims would be expected to show stronger sexual interest in both male and female children than control groups. Most male controls would be expected to show exclusively sexual interest in women, a few in men, but virtually none in children. Regarding child molesters' sexual interest in adults as compared to interest in children, no clear-cut predictions could be made, because child sex offenders quite commonly also show some level of, or even exclusively, sexual interest in adults (e.g., Worling, 2006).

Method

Sample. Inclusion criteria for all participants were: age over 21, white ethnic origin (to exclude possible interactions between ethnic origin of participants and target stimuli), IQ of at least 80 (in case of offenders IQ was extracted from prison files of prior psychological assessments; controls were assumed to have IQs above the threshold due to their level of functioning as prison personnel, professionals or students), and general reading ability.

The sample consisted of $N = 113$ participants (38 child sex offenders, 75 controls). Offenders were recruited in four different prison establishments in the UK. The local prison records were used to identify and contact all child sex offenders in each of the four institutions who met the inclusion criteria. Offender and non-offender control groups of equal size were recruited. Potential participants were approached and subsequently informed that the study investigated the usefulness of new measures of sexual preferences. Participation was voluntary. Participants were informed that there were no negative consequences for not participating or withdrawing consent in the course of the study. *Child sex offender groups* (boy victims only $n = 14$, girl victims only $n = 16$, mixed victims $n =$

8) consisted of child sex offenders (most, but not all, repeat offenders) convicted of hands-on offences with extra-familiar victims below 12 years of age (clearly prepubescent children). Prior to this study all child sex offenders had completed either the Rolling, Core, or Extended Sex Offender Treatment Programme of HM Prison Service (SOTP; Beech, Oliver, Fisher, & Beckett, 2005). *Offender controls* (*n* = 37) had convictions for non-sex related offences only. *Non-offending controls* (*n* = 38) consisted of a community sample including prison officers, other prison personnel and men from the community. The age of child sex offenders ranged from 28 to 74 years, for the controls from 21 to 63 years. Controls were on average younger than child-sex offenders, F (4, 108) = 10.16, p < .001 (first row of Table 4).

Direct EISIP measures. The Explicit Sexual Interest Questionnaire (ESIQ, Table 1) was developed for the purposes of the present study. The ESIQ features two subscales assessing sexual behavior (e.g., "I have enjoyed orally stimulating a man/woman/boy/girl.") and sexual fantasy (e.g., "I have daydreamed of having sex with a man/woman/boy/girl") with 5 items each. The ten fantasy and behavior items are combined with the four types of targets (man, woman, boy, or girl). Items were responded to in a dichotomous yes/no format. Corresponding scale frequency scores are calculated from the amount of "yes" responses divided by the number of scale items. With only 40 items it is highly economic and conceptually directly comparable with the indirect measures assessing sexual interest in men, women, and prepubescent boys and girls.

Indirect EISIP measures. Four different viewing time measures regarding sexual interest in men, women, girls and boys were used. Participants were asked to rate the sexual attractiveness of target stimuli on a 5-point Likert-scale ranging from 1 ("sexually unexciting") to 5 ("sexually very exciting") without time constraints. The 20 stimulus pictures were presented on the PC monitor until the response was given. The viewing time

was recorded unobtrusively. To optimize the assessment of individual differences (i.e.,
person and group effects) as opposed to main effects, the same random order of stimuli was
used for all participants. Individual response latencies were truncated at 10,000 ms (i.e.,
latencies longer than 10,000 ms were recoded to that value), and averaged across the five
stimuli belonging to each target category. No suspiciously short response latencies were
observed (all latencies $\geq$ 440 ms).

Three different IATs with the object categories Man-Woman, Girl-Woman, and
Boy-Man, and the attribute categories sexually exciting-unexciting were used. These were
derived from Ahlers et al.'s (2006) suggestion to disentangle sexual orientation (hetero- vs.
homesexual) and sexual age preference (children vs. adults). The former is measured with
the Men-Women IAT and the latter with both the Girls-Women and the Boys-Men IATs.
Each IAT consisted of five blocks following standard procedures as described in
Greenwald, McGhee, and Schwartz (1998). The stimulus pictures were the same as in the
VT task. The first block of 40 trials comprised a discrimination task of 10 words that had to
be classified as *sexually exciting* (erotic, exciting, lustful, sensual, orgasm) or *unexciting*
(dull, bland, indifferent, unexciting, boring). In the second block of 40 trials, 10 pictures
had to be assigned to the categories man and woman (girl vs. woman, boy vs. man,
respectively) by pressing the left or right response key, respectively. In the third block, both
tasks were mixed in alternating order. Four practice trials preceded 80 test trials. The left
response key had to be pressed for items belonging to the categories "Man" or "sexually
unexciting", the right response key for items for "Woman" or "sexually exciting" (and
accordingly for the Boys-Men and Girls-Women IATs). The fourth block of 40 trials was
similar to the second, but the key assignment was reversed. In the fifth and final block of
4+80 trials, both tasks were again combined. In this final block the left response key had to
be pressed for items relating to the categories "Man" or "sexually exciting", and the right

response key for items relating to "Woman" or "sexually unexciting". Incorrect responses were indicated by an error message throughout all blocks. The IAT was scored by calculating the difference between the mean response latencies of the critical third and fifth block, divided by the pooled standard deviation of response latencies (Greenwald, Nosek, & Banaji, 2003). Only trials with a correct answer were used, error trials were discarded. To optimize the measurement of individual differences, and not main effects, the order of the two combined IAT blocks was kept constant for all participants.

Other measures. To control for an influence of socially desirable responding (SD), the *Balanced Inventory of Desirable Responding* (BIDR; Paulhus, 1998) was included. This questionnaire has proven to be valid in offender samples (Kroner & Weekes, 1996). As a short screening tool of pedophilic sexual interest the Screening Scale for Pedophilic Interests (SSPI; Seto & Lalumière, 2001) was used. It is an actuarial four-item scale summarizing offenders' sexual victim characteristics (any male victims, more than one victim, any victims < 12 years, any unrelated victims). This index has been shown to predict phallometrically assessed sexual arousal to children and serious violent or sexual reoffending in adult male child sex offenders (Seto, Harris, Rice, & Barbaree, 2004) and was used as a validation measure.

Materials. The pictures of target persons used for the indirect measures were selected from the Not-Real-People picture set A and B (Pacific Psychological Assessment Corporation, 2004). The picture set features categories of sexual maturation for individuals (all white Caucasian) according to Tanner (1978), ranging from Tanner categories 1 to 3 (prepubescent children), Tanner category 4 (adolescents) to Tanner category 5 (adults). For men, women, boys and girls, five pictures showing the head and full body in swimming clothes were used. Pictures of children were taken from Tanner categories 1 to 3, and pictures of adults from Tanner category 5. No pictures of adolescents (Tanner 4) were used.

Procedure. All experimental protocols and data collection methods were approved by the Department of Psychology's Ethics Committee at the University of York, and the prison establishments. All assessments were run on an IBM-compatible laptop in individual sessions in a separate and quiet room. To familiarize participants with the 20 stimulus pictures, these were presented at the beginning of the session one by one. Then the participants worked through the different tasks in the following order: Men-Women IAT, Girls-Women-IAT, Boys-Men IAT, Viewing Time Task, BIDR, and the ESIQ explicit sexual interest questionnaires. After completing the assessments, each participant was thanked and debriefed.

Results

Psychometric Data

Reliabilities. Psychometric data of the self-report ESIQ are presented in Table 1. Internal consistencies of the eight subscales and the four combined scales of the ESIQ were satisfactory ($.86 < \alpha < .97$). For the purposes of the present study only the combined sexual interest scales will be reported. The VT-indices based on truncated raw data showed satisfactory reliabilities. With the exception of viewing time for young girls ($\alpha = .77$) internal consistency was .85 or better (Table 2). An exploration of the scoring technique used by Gress (2005) based on intra-individual z-transformation (i.e., ipsatization) diminished the Cronbach's alphas of the VT-indices to $.30 < \alpha < .66$ due to an inflation of small intrapersonal differences that reduced interpersonal variance.

The reliability results for the IATs were mixed. After discarding the IAT scores of nine participants with error rates $\geq 35\%$ in at least one combined block, the Man-Woman and Girl-Woman IATs showed nearly satisfactory reliability ($\alpha = .79$). The Boys-Men IAT showed an unsatisfactory low α of .65. This weak reliability coefficient is likely to be due to a lack of variability in the sexual preference of boys over men, because the sample did

not include a sizeable proportion of homosexual men who were not child sex offenders. This means that virtually all participants were either sexually attracted to men and boys, or to neither of them. In consequence, the IAT as a relative measure of sexual interest for one target category over the other could not detect substantial individual differences in the Boy-Man IAT, and hence, reliability was low. The BIDR total score showed a satisfactory internal consistency of $\alpha = .84$.

Convergent and discriminant validity of sexual interest measures. Correlations between all sexual interest measures were calculated to investigate convergent validity (Table 3). To enable comparisons with the three relative IAT measures, conceptually analogous difference scores were also calculated for the ESIQ and viewing time measures.

All intra-method intercorrelations of the ESIQ showed coefficients in the expected directions. The correlations between three out of four VT-indices and the corresponding explicit ESIQ measures were significant and in the expected directions. Only VT for women did not correlate with the explicit measure. The correlations outside the main diagonal were low or negative with the exception of high correlations between VT-men and VT-boys that are again due to the fact that virtually all participants with a sexual preference for boys also showed a sexual preference for men and vice versa. A demonstration of discriminant validity of these VT-scales (and the corresponding explicit scales) would require a sizeable proportion of homosexual offender or non-offender controls. In this sample only 2 (2.7%) men in the control groups reported more sexual interest in men than in women in the corresponding ESIQ categories compared to 12 (31.6%) child sex offenders. Nine out of 12 (75%) gay child sex offenders had boy victims only, the remaining 3 (25%) had abused boys and girls. Interestingly, the VT-boy correlated with the VT-girl (indicating interest in children in general), whereas VT-men did correlate with VT-

boys, but not significantly with VT-girl, indicating that homosexual orientation as such was more strongly related to sexual interest in boys than in girls in this sample.

The correlations of the three IATs and the explicit ESIQ correlations indicate a substantial convergent validity for the Men-Women IAT and the Girls-Women IAT and the corresponding explicit difference scores. However, no significant correlation was obtained for the Boys-Men IAT. No unexpected correlations between the IATs and the ESIQ scales emerged, thus providing evidence for the discriminant validity of the measures. Regarding the convergent validity of the VT and IAT measures, a significant correlation emerged only for the Girls-Women IAT.

SD as measured by the averaged BIDR showed only a negative correlation with self-reported interest in women ($r = -.22$). Participants with high SD scores showed a tendency to downplay explicit sexual interest in women (Table 3).

Correlations of all EISIP measures and the SSPI pedophilic interests score were consistently significant for sexual interest in men and men over women. The correlation levels up to .60 (Table 3), were surprisingly strong taken into consideration that the SSPI is only defined for child sex offenders. Therefore, the variability in this sample was quite restricted ($M = 3.9$; $SD = 1.2$; range 2-5, for group levels see Table 4) due to selection criteria and small sample size of the child sex offenders. In the trade-off between criterion group purity in a known-group validation approach and group heterogeneity needed to optimize the test of convergent validity we opted for a clearly selected child sex offender sample. Due to this only two of the four actuarial SSPI-items showed variability in this study. Therefore, to calculate the internal consistency of the SSPI under these circumstances is meaningless. Nevertheless, the VT Boys-Men measure showed good convergent validity with the SSPI ($r = -.49$). All other correlations had the expected sign.

All EISIP measures of homosexual interest in men indicated pedophilic preferences. However, across methods, all three Boys-Men difference scores were not at all, or even negatively associated with the SSPI score. It has to be noted though, that the association between homosexual orientation and deviance is due to the distribution of gay participants in the sample, thus confounding age preference and sexual orientation.

Group Differences and Classification

Group differences across sexual interest variables. To test whether sexual interest variables differentiated between child sex offenders and control groups, a series of oneway ANOVAs was conducted (Table 4). For all ESIQ measures except the Boys-Men difference measure, $F < 1$, a significant group effect was found, $3.54 < F < 10.67, p < .05$. All but one of the VT categories showed group effects $2.82 < F < 10.05, p < .05$ and for the IATs, the Men-Women, Girls-Women, and averaged Children-Adults IATs produced group effects, $2.91 < F < 6.19, p < .05$, but not the Boys-Men IAT, $F(4, 102) = 1.01$.

Post-hoc comparisons confirmed the hypothesis that in the ESIQ scales both control groups showed strong sexual interest in women only. Child sex offenders' interest for children was significantly more pronounced than controls', but interestingly, child molesters also reported high levels of interest in adults – especially in women. As expected, boy-only sex offenders showed the strongest interest in boys and hardly any interest in girls; this finding was reversed for girl-only offenders and the difference was statistically significant. The mixed victim group reported interest in girls and boys. Offenders with any boy victims showed strong explicit sexual interest in both men and boys, differing significantly from the other three groups. Sex offenders with any girl victims reported high explicit interest in women, not significantly differing from the controls, but from the boy-only offenders.

VT data showed that there were no group differences in VT for women. Boy-only sex offenders produced the highest VTs for men and boys, but only differed significantly from the offender controls. Offenders with any girl victims had the highest latencies in both adult categories, indicating primary interest in adult sex objects. No significant differences were observed between offender groups with girl only and mixed victims. In the IAT results the only group differences emerged for the Girl-Women and the averaged Children-Adults IATs. In these measures child sex offenders with any boy victims differed significantly from non-offending control groups.

Univariate discrimination. A series of ROC-analyses was conducted to test the criterion validity of all single measures used in this study (Table 2). Comparison groups were all controls or offender controls versus the groups of all child sex offenders or boy- or girl-only victim groups. As implied by the effect sizes reported in Table 2, nearly all measures discriminated above chance levels between child sex offenders and all controls. Only the explicit Boys-Men difference score and the VT Children-Adults score did not show significant criterion validity. VT measures in general had good classificatory power in discriminating child sex offender groups from offender controls. The IATs performed generally less well. Only the Children-Adult IAT consistently discriminated between controls and all offender groups with boy victims. Particularly high *AUCs* were found for the VT categories of Men and Boys ($.78 < AUC < .90$) indicating that long VTs for boys and men were associated with prior sex offending against children in general – even for the girl victims only group. The explicit measures discriminated well, too, especially for the categories involving child-related items, indicating a relatively open self-report of deviant sexual interest in the sample.

Multivariate analyses. Sex offenders showed higher levels of general sexual interest than controls across the full range of ESIQ and VT measures as evidenced by

significant group main effects in the corresponding 5 (Group) x 4 (Sexual interest categories) mixed model MANOVAs, $F(4, 108) = 15.83$, $p < .001$, $\eta_p^2 = .37$ and $F(4, 107) = 8.79$, $p < .001$, $\eta_p^2 = .25$, respectively. Regarding the IAT measures, the same significant group main effect emerged for the sex offenders in an omnibus 5 (Group) x 3 (IAT measures) mixed model MANOVA, $F(4, 99) = 6.24$, $p = .000$, $\eta_p^2 = .20$. These results suggest that sex offenders generally have stronger sexual interest or in case of the IATs stronger homosexual or child preferences across all the EISIP measures.

Binary logistic regression and ROC analyses. To test whether a profile of conceptually different direct and indirect measures of sexual interest would generate incremental validity, we conducted block-wise binary logistic regression analyses combining groups of measures to classify controls versus child sex offenders. All regressors were *z*-standardized. The resulting probability estimates from multivariate binary logistic regressions were used as test-scores in corresponding ROC-analyses (controls vs. child sex offenders).

The IAT measures alone explained the smallest fraction of variance in the sample, but criterion validity was acceptable (Table 5). VT measures alone performed better. The combination of all indirect measures explained 55% of variance and showed a very good criterion validity of $AUC = .88$. The explicit ESIQ scales accounted for 63% of the variance and showed the same criterion validity (.88) as all indirect measures taken together. The whole profile including all explicit and indirect predictors explained 75% of variance and showed excellent discriminative power ($AUC = .95$).

To ascertain that sample specific characteristics did not produce an artificial overestimation, all binary logistic regressions were cross-validated via the leave-one-out-method (Efron, 1983). This statistical procedure allows for estimating the size of sample-dependent estimation errors in the binary logistic regression function – the so called

optimism. Optimism rates were low ranging between 2.0% for all indirect measures and 3.5% for the explicit ESIQ self-reports only (Table 5).

<div align="center">Discussion</div>

The results of this study provide first preliminary evidence that the EISIP test battery is a reliable and valid measure of sexual interest in child sex offenders, non sex-related offenders, and non-offenders. Using an univariate approach, specific IATs and VT measures of the EISIP have been found to discriminate child sex offenders from controls nearly as well as the explicit measures in this largely non-denying sample. This finding is in line with previous research that has mostly considered single indirect measures only (e.g., Abel et al., 1998; Gray et al., 2005; Gress, 2005; Harris et al., 1996). Abel et al. (2001) further compared VT measures with phallometric assessments and found both indirect measures to be valid. However, the present study is the first to compare the criterion validity of different latency-based indirect measures of sexual interest using a forensic sample.

In the present study, the VT measures outperformed the IATs not only in terms of discriminatory power but also with respect to reliability, convergent, and discriminant validity. With regard to their psychometric properties, the VT measures reached the level of reliability and validity that is otherwise only known for direct assessment methods such as questionnaires. The reliability and validity of the IATs can be considered as moderate to satisfactory at best. However, it has to be noted that the sample characteristics of the present study may have obscured the psychometric quality of the Boys-Men IAT which performed less well than the other two IATs. Given the relative nature of this measure, this particular IAT can only perform well if participants substantially differ in their relative sexual preference for men and boys. The present sample showed only very little variability in this respect. The critical test of the sensitivity of the Boys-Men IAT can only be

conducted with a community sample with a sizeable proportion of homosexual men. This has still to be examined in future studies. Importantly, even though the IATs were psychometrically less satisfactory than the VT-measures, they still provided additional information and showed a significant correlation with the SSPI scores of pedophile interest.

The newly developed explicit self-report questionnaire ESIQ was found to be psychometrically sound. Besides the good psychometric quality and economy, the ESIQ has the advantage of providing a direct assessment of sexual interest that is conceptually directly comparable to the information provided by the VT measures (using absolute ESIQ-scores) and the IAT measures (using relative ESIQ difference scores).

Importantly, by combining conceptually different measurement approaches to profiles of sexual interest, the predictive validity could be increased. Taken together, the indirect measures used in this study explained nearly the same amount of between-group variance as the self-report measures. The combination of all direct and indirect measures could further increase the prediction to the level of nearly perfect criterion validity ($AUC =$ .95) and the very large amount of 75% of explained variance. The evidence suggests that the different measures tap into slightly different domains of sexual interest with specific variance (e.g., based on more automatic or more controlled processes).

This differentiation is beneficial because one has to be very cautious when interpreting effects of indirect and/or implicit measures at single-case level. The scores of indirect measures may be influenced by numerous personal and contextual factors (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). Therefore it is questionable – if not dubious – to interpret any single indirect measurement as an absolute index of a specific psychological attribute. Of course, this note of caution applies to direct measures as well. A solution to this problem lies in the general diagnostic principle of convergence.

Conclusions about deviant or non-deviant sexual interest can be drawn with greater confidence if they are based on several conceptually different, convergent, and valid, indirect measures (De Houwer et al., 2009).

Limitations of the study

Some limitations of this study have to be taken into account regarding the conclusions that can be drawn from the results. First, the limited number of child sex offenders does not allow for strong inferences about differences between child sex offender subgroups with different victim characteristics. Second, as already mentioned, this study lacks a balanced proportion of homosexual men in the control sample. Hence, offender status is confounded with sexual orientation. This may lead to underestimating the criterion validity of deviant sexual interest measures. It may also result in overestimating the validity of measures of sexual orientation for discriminating between offenders and controls. However, an additional series of of hierarchical multiple regressions revealed that measures of sexual interest in children (mainly the Boys-Men and Girls-Women IATs) independently accounted for criterion variance. This finding corroborates that the criterion validity of the EISIP was not only based on the homosexual-heterosexual discrimination. Nevertheless, future studies should use balanced control samples.

Implications for the practical use of the EISIP

The results of the present study suggest that the *Explicit and Implicit Sexual Interest Profile* (EISIP) could be useful as an additional diagnostic tool in several respects: First, the EISIP is an economical (with respect to both assessment time and monetary costs) computer-based assessment tool to assess socially problematic and/or inappropriate sexual interest. Participants took 35 minutes on average to complete the EISIP test battery. The assessment is non-intrusive and ethically acceptable due to the use of sexually non-explicit, clothed stimuli not showing identifiable, real individuals. From a practical point of view the

EISIP appears to be an attractive and affordable complement to be used alongside established PPG procedures. Incremental validity might be gained from converging data of conceptually different measurement paradigms that tap into differential domains of deviant sexual interest. First evidence into this direction is indicated by the associations with the phallomerically validated SSPI scores across relevant EISIP categories. Hence, comparing EISIP with PPG data would be an important next research step.

Second, because deviant sexual interest has proven to be one of the best predictors of sexual recidivism (Hanson et al., 1998; 2005) the EISIP may prove useful to assess sexual interest in the context of criminal prognosis – especially in settings where ongoing sex offender risk assessment and -management is mandatory. It could also be used to provide additional diagnostic information in cases where little actuarial and detailed file data are available. In these instances, the EISIP may be useful in detecting deviant sexual preferences that call for more intense monitoring and further dynamic risk assessment strategies. Whether the EISIP measures are valid predictors of reoffending is at this point an open empirical question that requires further research.

Interestingly, the sex offender sample not only showed higher levels of deviant sexual interests than the control groups in this study, but also a generally higher level of socially accepted sexual interests in adults (female and/or male). The first finding is in line with the so-called sexual preference hypothesis (Freund & Blanchard, 1989). The latter result of even stronger sexual interest in adults – at least at group level – is a little more puzzling, although it is not totally unexpected from a clinical perspective. It is rather common among child sex offenders to show adult sexual interest (e.g, Worling, 2006). Seto and Lalumière (2001) found that among the child sex offenders who scored lowest on the SSPI (score 0-1) approximately 80% showed greater sexual arousal for adults than for children; among those who scored highest (SSPI score 5) this pattern was still obtained in

28%. Hence, a large proportion of child sex offenders is genuinely sexually interested in adults. It can be speculated whether sexual preoccupation – empirically a powerful risk factor for sexual recidivism (Hanson, 2006) – is the underlying cause of the observed increased level of sexual interest scores.

Last but not least, the EISIP may be useful in a therapeutic context such as to confront denying child sex offenders with their deviant sexual interests in order to initiate a process of questioning self-relevant assumptions about their child related sexual preferences. In addition, the EISIP may also be used to identify socially acceptable sexual interest that can be used as a resource in sex offender therapy.

In summary, the present research is a first step on the way of developing a multi-method measuring approach of sexual interest that can be further extended with other, conceptually different measures ranging from phallometry to new latency-based measures of sexual interest. With an increasing number of related but conceptually different measures diagnostic decisions benefit from the aggregation principle that renders the assessment more accurate, and makes it more difficult to fake. This is of great importance since in the age of the internet it is easy to retrieve information about the underlying rationale of any specific measurement method used in sex offender assessment. The availability of a test battery of indirect measures of sexual interest with promising psychometric properties now calls for a thorough evaluation of the usefulness of this approach for child sex offender treatment, management, and the prediction of reoffending especially in less select samples of more denying and treatment refusing child sex offenders.

References

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

Abel, G. G., Huffmann, J., Warberg, B., & Holland, C. L. (1998). Visual reaction time and plethysmography as measures of sexual interest in child molesters. *Sexual Abuse. A Journal of Research and Treatment, 10,* 81-95.

Ahlers, C. J., Schaefer, G. A., & Beier, K. M. (2006). Das Spektrum der Sexualstörungen und ihre Klassifizierbarkeit in DSM-IV und ICD-10 [The spectrum of sexual disorders and their classifiability in ICD-10 and DSM-IV]. *Sexuologie* [Sexuology]*, 12,* 120-152.

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

Beech, A., Oliver, C., Fisher, D., & Beckett, R. (2005). *STEP 4: The Sex Offender Treatment Programme in prison: Adressing the offending behaviour of rapists and sexual murderers*. London: HM Prisonservice. Retrieved November 20, 2009, from http://www.hmprisonservice.gov.uk/assets/documents/100013DBStep_4_SOTP_repo rt_2005.pdf

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135,* 347-368.

Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on crossvalidation. *Journal of the American Statistical Association, 78,* 316-331.

Freund, K., & Blanchard, R. (1989). Phallometric diagnosis of pedophilia. *Journal of Consulting and Clinical Psychology, 57,* 100-105.

Gray, N. S., Brown, A. S., MacCulloch, A. J., Smith, J., & Snowden, R. J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology, 114,* 304-308.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216.

Gress, C. L. Z. (2005). Viewing time measures and sexual interest: Another piece of the puzzle. *Journal of Sexual Aggression, 11,* 117-125.

Gress, C. L. Z., & Laws, D. R. (2009). Measuring sexual deviance: Attention-based measures. In A. R. Beech, L. A. Craig, & K. D. Browne (Eds.). *Assessment and treatment of sexual offenders: A handbook* (pp. 109-128). West Sussex, UK: Wiley.

Hanson, R. K. (2006). Stability and change: Dynamic risk factors for sexual offenders. In W. L. Marshall, Y. L. Fernandez, L. E. Marshall, & G. A. Serran (Eds.). *Sexual offender treatment. Controversial issues* (pp. 18-31). New York, NJ: Wiley

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, *66,* 348-362.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistant sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73,* 1154-1163.

Harris, G. T., Rice, M. R., Quinsey, V. L., & Chaplin, T. C. (1996). Viewing time as a

measure of sexual interest among child molesters and normal heterosexual men.

Behavior Research and Therapy, 34, 389-394.

Imhoff, R., Schmidt, A. F., Nordsiek, U., Luzar, C., Young, A. W. & Banse, R. (in press).

Viewing time effects revisited: Prolonged response latencies for sexually attractive

targets under restricted task conditions. *Archives of Sexual Behavior*.

Kalmus, E., & Beech, A. R. (2005). Forensic assessment of sexual interest: A review.

Aggression and Violent Behavior, *10,* 193-218.

Kroner, D. G. & Weekes, J. R. (1996). Balanced Inventory of Desirable Responding:

Factor structure, reliability, and validity with an offender sample. *Personality and*

Individual Differences, 21, 323-333.

Laws, D. R. (2003). Penile plethysmography: Will we ever get it right?. In T. Ward, D. R.

Laws, & S. M. Hudson (Eds.). *Sexual deviance – Issues and controversies* (pp. 82-

102)*.* London: Sage.

Laws, D. R., & Gress, C. L. Z. (2004). Seeing things differently: The viewing time

alternative to penile plethysmography. *Legal and Criminal Psychology, 9,* 183-196.

Lanyon, R. I. (2001). Psychological assessment procedures in sex offending. *Professional*

Psychology: Research and Practice, 32, 253-260.

Letourneau, E. J. (2002). A comparison of objective measures of sexual arousal and

interest: Visual reaction time and penile plethysmography. *Sexual Abuse: A Journal*

of Research and Treatment, 14, 207-223.

Marshall, W. L. (2007). Diagnostic issues, multiple paraphilias, and comorbid disorders in

sexual offenders: Their incidence and treatment. *Aggression and Violent Behavior*,

12, 16-35.

Marshall, W. L., & Fernandez, Y. L. (2000). Phallometric testing with sexual offenders: Limits to its value. *Clinical Psychology Review, 20,* 807-822.

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

Mokros, A., Dombert, B., Osterheider, M., Zappalà, A., & Santtila, P. (in press). Assessment of pedophilic sexual interest with an attentional choice reaction time task. *Archives of Sexual Behavior.*

Murphy, W. D., & Barbaree, H. E. (1994). *Assessments of sex offenders by measures of erectile response: Psychometric properties and decision making.* Brandon, VT: Safer Society.

Pacific Psychological Assessment Cooperation (2004). *The Not-Real People Stimulus Set for Assessment of sexual interest.* Victoria, BC: Author.

Paulhus, D. L. (1998). *The Balanced Inventory of Desirable Responding.* Toronto/Buffalo: Multi-Health Systems.

Rosenzweig, S. (1942). The photoscope as an objective device for evaluating sexual interest. *Psychosomatic Medicine, 4,* 150-157.

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

Seto, M. C., Harris, G. T., Rice, M. E., & Barbaree, H. E. (2004). The Screening Scale for Pedophilic Interests and recidivism among adult sex offenders with child victims. *Archives of Sexual Behavior, 33,* 455-466.

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

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

Tanner, J. (1978). *Foetus into man: Physical growth from conception to maturity.* Cambridge, MA: Harvard University Press.

Ward, T., & Beech, T. (2006). An integrated theory of sexual offending. *Aggression and Violent Behavior*, *11,* 44-63.

Ward, T., & Siegert, R. J. (2002). Toward a comprehensive theory of child sexual abuse: a theory knitting perspective. *Psychology, Crime and Law*, *9,* 319-351.

Worling, J. (2006). Assessing sexual arousal with adolescent males who have offended sexually: Self-report and unobtrusively measured viewing time. *Sexual Abuse: A Journal of Research and Treatment, 18,* 383-400.

Author Note

Biographical Sketches

Rainer Banse, PhD, is Professor for Social & Legal Psychology at the Department of Psychology, University of Bonn. His main research interests are indirect assessment methods in social and legal psychology, aggression, personal relationships, and personality psychology.

Alexander F. Schmidt, PhD, is currently a post-doctoral researcher at the Department of Psychology, Social & Legal Psychology, University of Bonn. His research interests include forensic assessment, offender therapy, indirect measures, and self-regulation processes.

Jane Clarbour, PhD, is a Senior Lecturer at the University of York and a chartered Forensic Psychologist. Her research interests are psychometrics, personality and antisocial behaviour, as well as emotional development.

Table 1: Items and psychometric properties of the Explicit Sexual Interest Questionnaire (ESIQ).

	Part-whole corrected item-total correlations			
Item	Man	Woman	Boy	Girl
Behavior				
I have enjoyed orally stimulating a ...	.91	.66	.62	.64
I have sexually caressed a ...	.92	.82	.82	.76
I have sexually penetrated a ... with my tongue or finger	.76	.80	.66	.73
I have sexually touched a ...	.92	.80	.71	.90
I have sexually penetrated a ... with my penis	.81	.82	.68	.62
Cronbach's *α*	**.95**	**.91**	**.86**	**.89**
Fantasy				
I find it erotic if I see a ...'s beautiful chest	.93	.87	.87	.76
I have daydreamed of having sex with a ...	.85	.82	.66	.76
I find it erotic to see a ...'s body through the clothes	.90	.73	.75	.65
I find it erotic to see a ...'s beautiful legs or bottom	.84	.73	.77	.71
I get excited when I imagine that a ... stimulates me	.89	.77	.66	.75
Cronbach's *α*	**.96**	**.91**	**.89**	**.88**
Total Score				
Cronbach's *α*	**.97**	**.94**	**.88**	**.90**

Note. In the instruction it is stated, that "boy" and "girl" refers to prepubescent children below the age of 12.

Table 2. Psychometric properties of social desirability, direct, and indirect measures of sexual preference.

	Reliability (α)	Effect size[a] (r)	Criterion Validity (ROC-analysis)		
			Child Sex Offenders ($n = 38$) All Controls ($n = 75$)	Boy Victims Only ($n = 14$) Offender Controls ($n = 37$)	Girl Victims Only ($n = 16$) Offender Controls ($n = 37$)
Explicit Sexual Interest					
Men	.97	**.41**	**.66**	**.74**	.45
Boys	.88	**.50**	**.72**	**.89**	.50
Women	.94	**-.37**	**.29**	**.24**	.37
Girls	.90	**.45**	**.69**	.49	**.79**
Men-Women	n/a	**.42**	**.74**	**.79**	.59
Boys-Men	n/a	-.16	.42	.36	.55
Girls-Women	n/a	**.59**	**.83**	**.78**	**.83**
Children-Adults (average)	n/a	**.52**	**.71**	**.77**	**.87**
Viewing Times					
Men	.85	**.54**	**.82**	**.89**	**.78**
Boys	.85	**.49**	**.80**	**.90**	**.86**
Women	.86	.08	.56	.63	**.74**
Girls	.77	**.44**	**.76**	**.81**	**.73**
Men-Women	n/a	**.33**	**.72**	**.82**	.46
Boys-Men	n/a	**-.34**	**.29**	**.22**	.48
Girls-Women	n/a	**.20**	**.61**	.51	.43
Children-Adults (average)	n/a	.02	.51	.33	.46
IATs					
Men-Women	.79	.13	.57	.63	.34
Boys-Men	.65	.16	**.62**	.60	.57
Girls-Women	.79	**.37**	.72	.67	.56
Children-Adults (average)	n/a	**.32**	**.71**	**.71**	.60
SSPI	n/a[b]				
BIDR averaged	.85				

Note. [a] Effect size Child Sex Offenders vs. All Controls; [b] see results section; n/a = not applicable; bold coefficients are significant at $p < .05$; values below .50 can be turned into their complement (1-*value below .50*) when testing direction is recoded.

Table 3. Correlations of direct and indirect measures of sexual interest, social desirability, and the Screening Scale for Pedophilic Interests.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
Explicit Sexual Interest																				
1. in Men																				
2. in Women	**-.70**																			
3. in Boys	**.65**	**-.62**																		
4. in Girls	.06	.08	.01																	
5. Men-Women[a]																				
6. Boys-Men[a]																				
7. Girls-Women[a]																				
8. Children-Adults[b]																				
Viewing Times																				
9. Men	**.42**	**-.45**	**.32**	.16	**.47**	**-.31**	**.48**	**.28**												
10. Women	-.16	.17	-.18	.08	-.17	.07	-.09	-.04	**.34**											
11. Boys	**.23**	**-.29**	**.24**	**.27**	**.28**	-.11	**.41**	**.35**	**.79**	**.47**										
12. Girls	.13	-.10	.11	**.51**	.12	-.08	**.40**	**.37**	**.55**	**.48**	**.62**									
13. Men-Women[a]																				
14. Boys-Men[a]																				
15. Girls-Women[a]																				
16. Children-Adults[b]																				
IATs																				
17. Men-Women					**.36**	-.15	**.28**	**.19**					.15	-.11	.19	.12				
18. Boys-Men					.12	-.04	.11	.09					.11	-.15	.14	.06	.13			
19. Girls-Women					**.27**	**-.22**	**.32**	**.19**					**.23**	**-.23**	**.27**	.14	**.44**	.17		
20. Children-Adults[b]					**.26**	-.17	**.29**	.18					**.21**	**-.25**	**.26**	.11	**.40**	**.73**	**.84**	
BIDR averaged	.05	**-.22**	.04	-.05	.14	-.04	.16	.14	.06	-.01	.05	-.04	.05	-.04	-.02	-.03	.16	.09	.07	.11
SSPI	**.56**	**-.50**	**.58**	-.24	**.60**	-.27	.29	.07	**.42**	-.06	.19	.03	**.42**	**-.49**	.09	-.19	**.39**	.12	.17	.21

Note. [a] Difference scores based on single interest measures; [b] Based on differences in mean combined single interest measures; only meaningful correlations are reported; bold sizes are significant at $p < .05$, two-tailed.

Table 4. Mean differences and standard deviations of age, BIDR, SSPI, direct, and indirect measures of sexual preference across child sex offender and control groups.

Measures	Child sex offenders			Offender Controls (*n*= 37)	Non-offender Controls (*n* = 38)	Group effect $F(4, (df_2))$
	Boy victims only (*n* = 14)	Girl Victims only (*n*= 16)	Mixed victims (*n* = 8)			
Age	47.2 AB (12.3)	50.9 A (9.1)	48.4 AB (10.2)	35.2 C (9.1)	41.8 BC (8.8)	10.20*** (108)
SSPI[ac]	4.79 A (0.43)	2.62 B (0.50)	5.00 A (0.00)	n/a	n/a	32.01***
BIDR averaged	3.21 A (0.59)	3.05 A (0.46)	2.98 A (0.41)	2.91 A (0.34)	2.79 A (0.51)	2.38 (108)
Explicit Sexual Interest						
in Men[bc]	0.50 A (0.46)	0.02 B (0.08)	0.45 AB (0.38)	0.05 B (0.18)	0.03 B (0.16)	5.88*** (30.37)
in Boys[ac]	0.33 A (0.23)	0.00 B (0.00)	0.28 AB (0.32)	0.00 B (0.00)	0.01 B (0.02)	73.57***
in Women[bc]	0.51 A (0.45)	0.93 B (0.09)	0.73 AB (0.34)	0.95 B (0.15)	0.94 B (0.18)	3.54* (30.07)
in Girls[bc]	0.01 A (0.05)	0.31 B (0.34)	0.35 AB (0.29)	0.01 A (0.05)	0.02 A (0.07)	5.14** (28.78)
Men-Women[bc]	-0.01 A (0.77)	-0.91 B (0.14)	-0.28 AB (0.61)	-0.90 B (0.32)	-0.91 B (0.33)	6.20*** (30.55)
Boys-Men[bc]	-0.17 A (0.41)	-0.02 A (0.08)	-0.18 A (0.32)	-0.05 A (0.18)	-0.02 A (0.16)	0.98 (30.51)
Girls-Women[bc]	-0.50 A (0.44)	-0.61 A (0.36)	-0.39 A (0.31)	-0.94 B (0.16)	-0.92 B (0.19)	10.67*** (28.05)
Children-Adults[bc] (average)	-0.34 AB (0.25)	-0.32 A (0.16)	-0.28 AB (0.27)	-0.49 B (0.04)	-0.47 B (0.06)	7.14*** (26.90)
Viewing Times (*ms*)						
Men[bc]	4035 A (1933)	2679 A (863)	3459 A (1190)	1826 B (770)	2048 AB (660)	8.47*** (28.69)
Boys[bc]	2656 A (948)	2291 A (830)	2385 AB (945)	1356 B (529)	1815 A (636)	10.05*** (28.90)
Women	3404 A (1774)	3595 A (1181)	3460 A (2085)	2637 A (1362)	3795 A (1624)	2.82* (107)
Girls[bc]	2438 A (872)	2516 AB (1480)	2573 AB (1045)	1604 B (671)	1654 B (553)	4.86** (28.47)
Men-Women	631 A (1448)	-916 BC (1063)	-1 AB (2777)	-811 BC (1174)	-1746 C (1551)	7.59*** (107)
Boys-Men[bc]	-1380 A (1104)	-388 AB (603)	-1074 AB (816)	-469 AB (623)	-233 B (548)	4.78** (29.49)
Girls-Women	-967 A (1161)	-1079 A (1391)	-887 A (2010)	-1033 A (1208)	-2141 A (1310)	4.53** (107)
Children-Adults (average)	-1173 A (918)	-734 A (773)	-980 A (1072)	-751 A (719)	-1187 A (634)	2.19 (107)
IATs						
Men-Women	0.03 A (0.52)	-0.42 A (0.36)	-0.15 A (0.61)	-0.21 A (0.41)	-0.44 A (0.54)	2.91* (99)
Boys-Men	-0.01 A (0.30)	-0.11 A (0.36)	0.02 A (0.16)	-0.13 A (0.28)	-0.19 A (0.37)	1.01 (102)
Girls-Women	-0.06 A (0.36)	-0.20 AB (0.31)	0.01 A (0.46)	-0.31 AB (0.40)	-0.52 B (0.34)	6.19*** (102)
Children-Adults (average)	-0.06 A (0.26)	-0.19 AB (0.28)	0.00 A (0.31)	-0.23 AB (0.25)	-0.35 B (0.27)	5.02*** (106)

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$. Group means with different subscripts in one row are statistically different (Student-Newman-Keuls-Test, $p < .05$). [a] = Non-parametric Kruskal-Wallis-Test (exact χ^2) [b] = Welch-Test [c] = Dunnett's C post-hoc comparisons, $p < .05$. n/a = not applicable.

Table 5. Criterion validity (ROC-analysis) after crossvalidated binary logistic regression (Controls vs. Child Sex Offenders).

Measures	Correct Classifications	Crossvalidated Optimism	*AUC*	Nagelkerke's R^2
Direct + Indirect	90%	2.0%	.95***	.75
Direct	87%	3.5%	.88***	.63
Indirect	85%	2.0%	.88***	.55
Viewing Time	80%	2.7%	.86***	.48
IAT	77%	2.9%	.77***	.23

Note. *** = p < .001; all regression models *p* < .001, all Hosmer-Lemeshow-Tests are non-significant.