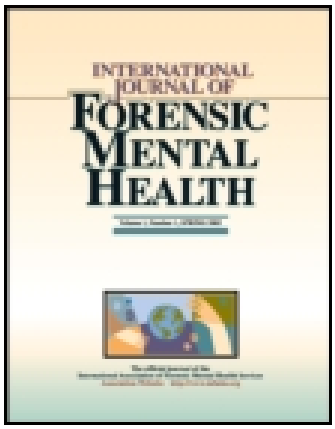

International Journal of Forensic Mental Health
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ufmh20

Collaborative Treatment Planning Using the Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Concordance of Therapist Evaluation and Client Self-Evaluation

Michael P. Lasher[a], Robert J. McGrath[a], Doug Wilson[a] & Georgia F. Cumming[a]
[a] Vermont Department of Corrections, Waterbury, Vermont, USA
Published online: 31 Mar 2015.

To cite this article: Michael P. Lasher, Robert J. McGrath, Doug Wilson & Georgia F. Cumming (2015) Collaborative Treatment Planning Using the Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Concordance of Therapist Evaluation and Client Self-Evaluation, International Journal of Forensic Mental Health, 14:1, 1-9

To link to this article: http://dx.doi.org/10.1080/14999013.2014.974087

INTERNATIONAL JOURNAL OF FORENSIC MENTAL HEALTH, 14: 1–9, 2015

ISSN: 1499-9013 print / 1932-9903 online
DOI: 10.1080/14999013.2014.974087

Routledge
Taylor & Francis Group

Collaborative Treatment Planning Using the Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Concordance of Therapist Evaluation and Client Self-Evaluation

Michael P. Lasher, Robert J. McGrath, Doug Wilson, and Georgia F. Cumming
Vermont Department of Corrections, Waterbury, Vermont, USA

Contemporary views on sex offender treatment planning suggest that therapist and client collaboration may enhance treatment outcomes, but little research has examined this topic. The use of dynamic risk assessment instruments may provide insight into what criminogenic needs should be focused on for clients in sex offender treatment. Here, we compared therapist and client ($N = 80$) mid-treatment assessment scores on the Sex Offender Treatment Intervention and Progress Scale (SOTIPS) in a prison-based treatment program for adult male sex offenders. The SOTIPS is 16-item dynamic assessment measure of treatment needs and progress that was designed to be scored at intake and thereafter as often as every six months on a four-point scale. Data were collected following the completion of SOTIPS as a group treatment exercise. Although there were significant overall differences in SOTIPS scores between therapists and clients ($t(79) = 6.50$, $p < .001$; $d = 1.03$), SOTIPS scores related to criminality, social stability and supports, and sexual deviance showed substantial (ICC = 0.75, $p < .001$), moderate (ICC = 0.59, $p < .001$), and fair (ICC = 0.23, $p < .05$) correlations, respectfully. Using partial correlations, controlling for static risk showed no impact on these relationships. Findings are discussed in light of the experiences of therapists. This study suggests that the SOTIPS provides a useful framework for therapists to engage clients in a collaborative process of indentifying clients' strengths, potential treatment needs, and treatment progress.

Keywords: sex offender, prison treatment, collaborative treatment planning, SOTIPS

Treatment services found most effective for offenders, including sex offenders, are those that follow the principles of risk, need, and responsivity (RNR; Andrews & Bonta, 2010; Andrews, Bonta, & Hoge, 1990; Hanson, Bourgon, Helmus, & Hodgson, 2009). Simply stated, correctional programs that target offenders who are at moderate or higher risk to reoffend (risk), modify offender characteristics that are closely linked to reoffending (criminogenic needs), and use treatment methods that engage offenders and are matched to their learning styles and abilities (responsivity) show the largest reductions in reoffending. The primary focus of this article centers on the last two of these practices, the need and responsivity principles.

Sex offenders' criminogenic needs have been well detailed over the last decade in a series of meta-analyses (Hanson & Morton-Bourgon, 2004, 2005; Mann, Hanson, & Thornton, 2010). They include potentially changeable risk factors such as pro-offending attitudes, deviant sexual interests, impulsivity, intimacy deficits, poor problem solving, and antisocial peers. Sex offender assessment instruments composed of these and other dynamic risk factors can be used to assess dynamic risk, identify relevant treatment targets, and measure treatment progress. Relevant instruments include the Violence Risk Scale–Sexual Offender Version (VRS-SO; Olver, Wong, Nicholaichuk, & Gordon, 2007), structured risk assessment model (SRA; Thornton, 2002), Stable-2007 and Acute-2007 (Hanson,

The views expressed are those of the authors and not necessarily those of the Vermont Department of Corrections.

Michael P. Lasher is now affiliated with the Department of Psychology, East Tennessee State University.

Address correspondence to Michael P. Lasher, 504 S. Center Street #3, Johnson City, TN 37604, USA. E-mail: lasher.michael.p@gmail.com

Harris, Scott, & Helmus, 2007), and the Sex Offender Treatment Intervention and Progress Scale (SOTIPS; McGrath, Cumming, & Lasher, 2013), which is examined in the present study.

The SOTIPS is an updated version of the Sex Offender Treatment Needs and Progress Scale (McGrath & Cumming, 2001; 2003), which has been used by almost one-fifth (18.5%) of community sex offender treatment programs in the United States (McGrath et al., 2010). In the SOTIPS development study, six of 22 SOTNPS items were deleted due to a weak statistical association with sexual recidivism and weak support in contemporary risk assessment literature, resulting in the 16-item SOTIPS measure (McGrath, Lasher, & Cumming, 2011, 2012).

In the initial evaluation of the SOTIPS, combined SOTIPS and Static-99R (Helmus Thornton, Hanson, & Babchishin, 2012) total scores predicted sexual recidivism better than either instrument alone (McGrath et al., 2012). SOTIPS also has shown increased accuracy in risk predictions when used in conjunction with the Vermont Assessment of Sex Offender Risk-2 (VASOR-2; McGrath, Hoke, & Lasher, 2013; McGrath et al., 2014). In addition to its usefulness as a risk instrument, the SOTIPS is viewed as a useful tool for identifying treatment needs and measuring client treatment progress, although replication studies are still needed.

Even though clinician identification and measurement of criminogenic needs are necessary activities for effective treatment planning, client agreement about the relevance of treatment plans and engagement in the treatment process is important as well, and is consistent with the responsivity principle. Across a range of psychological interventions, treatment has been found more effective when clients and clinicians have the same treatment goals (Owen, Duncan, Anker, & Sparks, 2012; Tryon & Winograd, 2011; Veysey, 2008) including in sex offender treatment (Abel et al., 1988). Higher levels of treatment engagement in sex offender treatment programs have been positively correlated with changes in pre- and post-treatment test scores on a variety of treatment targets such as insight in sex offending patterns and management of deviant sexual thoughts (Beech & Hamilton-Giachritsis, 2005; Levenson & Macgowan, 2004). Similarly, men in a California prison program who were assessed as having "got it," that is, embraced and made progress in achieving program treatment goals, reoffended less often than completers who did not "get it" (Marques et al., 2005). Client satisfaction surveys in sex offender treatment programs have found that clients are commonly in agreement with the treatment targets in those programs (Levenson, Macgowan, Morin, & Cotter, 2009; Levenson & Prescott, 2009), which were consistent with empirically-supported criminogenic needs (Mann et al., 2010).

Shingler and Mann (2006) note that clinician and client collaboration typically has not been viewed as a necessary endeavor in offender treatment programs. Early approaches towards criminal justice clients emphasized authoritarian and confrontational approaches (Miller & Rollnick, 2002), and early writers in the sex offender field recommended this style (e.g., Salter, 1988). Contemporary methods (e.g., Proulx, Tardif, Lamoureux, & Lussier, 2000; Shingler & Mann, 2006) argue for a collaborative approach to risk assessment and treatment planning in sex offender programs, including educating clients about risk factors, assessment, and the treatment planning process. To our knowledge, however, sex offender treatment programs have not used a collaborative assessment and treatment planning approach in which the client scores himself on a dynamic risk assessment measure and discusses his self-identified strengths and treatment needs with his therapist and peers in a treatment group.

In the present study, clients in a prison sex offender treatment program scored themselves at mid-treatment on a dynamic risk instrument, the SOTIPS (McGrath et al., 2013). The goals were for each client to self-assess his dynamic strengths and risk factors, identify progress that he had made in treatment to date, and help update his treatment plan. Each client discussed his self-assessment in group treatment sessions and received feedback from his peers and therapists.

The present study examined the level of agreement between clients' and their therapists' SOTIPS scores at a mid-treatment evaluation. The influence of static risk on the relationship between clients' and therapists' SOTIPS scores was also examined. These findings are discussed in light of therapists' experiences using the SOTIPS as a collaborative treatment planning activity.

METHOD

Sample

Client participants were 80 incarcerated adult male sex offenders enrolled in a prison-based sex offender treatment program. Therapist participants were six master's level clinicians who provided treatment to clients in the program. All clients enrolled in the treatment program between 2009 and 2013 were included in the study if they were in enrolled in a core treatment group that was co-facilitated by two therapists who each had been certified to score the SOTIPS. Clients in the program's special track for individuals with developmental disabilities also were excluded in the study.

All clients were convicted of at least one Category "A" sexual offense as defined in the Static-99 coding manual (Harris, Phenix, Hanson, & Thornton, 2003). A Category "A" sex offense is illegal sexual behavior committed against an identifiable child or non-consenting adult victim. Thus, it includes contact sex offenses such as sexual

assault, attempted sexual assault, and child molestation, as well as non-contact sex offenses such as exhibitionism, voyeurism, and Internet luring. Individuals whose only sexual offense was a Static-99R Category "B" offense were excluded from the study. A Static-99R Category "B" offense concerns sexual behavior that was illegal, but the parties were consenting, or no identifiable victim was involved. Examples of these types of excluded offenses are consenting sex with an adult in a public place, soliciting a prostitute, and possessing child pornography.

Table 1 details sample characteristics. In terms of primary offender type, clients who committed contact sexual offenses against extrafamilial children age 15 and younger were classified as child molesters. Those who committed contact sexual offenses against victims age 16 or older were classified as rapists. Clients who sexually assaulted their biological children or stepchildren were classified as incest offenders. Non-contact sex offenders committed offenses such as exhibitionism, voyeurism, and obscene phone calls.

Program and Setting

The program in which client participants were enrolled was the Vermont Treatment Program for Sexual Abusers (VTPSA) prison program operated by the Vermont Department of Corrections (McGrath, Cumming, Livingston, & Hoke, 2003). All sentenced sex offenders in Vermont are incarcerated in state prisons because there are no county jails. Client participants, depending on their risk and needs, received between approximately 225 and 375 hours of cognitive-behavioral group treatment, augmented with periodic individual and family treatment sessions, over the course of 14 to 24 months. All clients attended two 90-minute core groups per week that focused on treatment engagement, problem identification, social skills practice, and release planning. Depending on treatment need, clients also attended one or more 12- to 16- week specialized groups that targeted areas such as offense supportive attitudes, offense-related sexual interests, emotion management, and relationships skills. Ancillary mental health, education, recreation, and employment services also were available to program participants.

TABLE 1
Client Sample Characteristics ($N = 80$)

	Mean (*SD*)	% (*n*)
Age (*SD*)	37.7 (11.0)	
Years education (*SD*)	11.8 (5.4)	
Ethnicity white		96.25% (77)
Primary offender type		
Rapist		20.00% (16)
Male child molester		17.50% (14)
Female child molester		45.00% (36)
Incest offender		8.75% (7)
Non-contact offender		8.75% (7)
Months incarcerated (*SD*)	41.5 (28.4)	
Months in treatment (*SD*)	11.8 (5.4)	
Mean group size (*SD*)	7.2 (1.5)	
Previous sex offender treatment		47.50% (38)
Static risk scores		
Static-99R	3.7 (2.5)	
% low risk		17.50% (14)
% moderate-low risk		28.75% (27)
% moderate-high risk		28.75% (27)
% high risk		25.00% (20)
VASOR-2	9.9 (3.7)	
% low risk		12.50% (10)
% moderate-low risk		23.75% (19)
% moderate-high risk		31.25% (25)
% high risk		32.50% (26)

Measure

Sex Offender Treatment Intervention and Progress Scale (SOTIPS)

The SOTIPS is a provider-administered measure designed to score clients on 16 dynamic risk factors (see Table 2) at intake and thereafter as often as every six months on a four-point scale ranging from *minimal to no need for improvement* to *very considerable need for improvement*. Total scores range from 0 to 48 points and are organized into three risk groups; low (0 to 10), moderate (11 to 20), and high (21 to 48). Also shown in Table 2, the 16 SOTIPS items can be divided into three factors; sexual deviance, criminality, and social stability and supports.

The SOTIPS can be scored reliably (ICC = .77, $p <$ .001) and has shown moderate predictive accuracy on its own for sexual (AUC = .70, $p < .001$) and other violent (AUC = .66, $p < .001$) offending, as well as in combination with the Static-99R (sexual AUC = .74, $p < .001$; violent AUC = .70, $p < .001$) and VASOR-2 (sexual AUC = .77, $p < .001$; violent AUC = .69, $p < .001$ (McGrath, Cumming, & Lasher, 2013; McGrath et al., 2011). In the present study, interrater reliability between four co-therapists who scored 47 cases showed good overall agreement (ICC = .89, $p < .001$).

Static-99R

The Static-99R is a 10-item actuarial instrument designed to assess the recidivism risk of adult males known to have committed at least one sexual offense (Helmus et al., 2012). Items are identical to the Static-99, with the exception of updated age weights. The 10 items pertain to sexual and nonsexual offense history, victim characteristics, and offender demographics. Total scores range from −3 to 12 points and are organized into four risk groups; low (−3 to 1), moderate-low (2 to 3), moderate-high (4 to 5) and high (6 to 12). A meta-analysis of 22 studies found a moderate relationship (AUC = .70, $p < .001$) between

TABLE 2
Therapist and Client Mean SOTIPS Scores and Treatment Goal Endorsement ($N = 80$)

	Therapist Score[1]			Client Score[1]						
Item/Factor/Scale	Mean	(SD)	Identified Goal[2]	Mean	(SD)	Identified Goal[2]	Therapist Only Identified Goal[2]	Client Only Identified Goal[2]	Mutually Identified Goal [2]	Mutually Identified Strength[3]
Sexual Deviance Factor	5.08	(1.84)		3.16	(1.56)					
Offense Responsibility	0.69	(0.61)	61.3%	0.33	(0.47)	32.5%	31.3%	2.5%	30.0%	36.3%
Sexual Behavior	0.03	(0.16)	2.5%	0.21	(0.41)	21.3%	1.3%	20.0%	1.3%	77.5%
Sexual Attitudes	0.91	(0.53)	81.3%	0.37	(0.51)	36.3%	50.0%	5.0%	31.3%	13.8%
Sexual Interests	0.77	(0.60)	68.8%	0.34	(0.50)	32.5%	41.3%	5.0%	27.5%	26.3%
Sexual Risk Management	1.52	(0.69)	93.8%	0.84	(0.70)	67.5%	28.9%	2.5%	65.0%	3.8%
Stage of Change	1.15	(0.36)	100.0%	1.07	(0.27)	100.0%	0%	0%	100.0%	0%
Criminality Factor	5.66	(2.97)		5.09	(2.63)					
Criminal Behavior	0.84	(0.65)	70.0%	0.89	(0.53)	80.0%	10.0%	20.0%	60.0%	10.0%
Criminal Attitudes	1.18	(0.71)	83.8%	0.90	(0.57)	78.8%	12.5%	7.5%	71.3%	8.8%
Cooperation with Treatment	1.03	(1.10)	58.8%	0.79	(0.99)	51.3%	18.8%	11.3%	40.0%	30.0%
Cooperation with Supervision	1.75	(1.50)	60.0%	1.78	(1.50)	60.0%	5.0%	5.0%	55.0%	35.0%
Impulsivity	0.88	(0.66)	72.5%	0.74	(0.63)	63.8%	18.8%	10.0%	53.8%	17.5%
Sociality Stability & Supports Factor	4.05	(2.07)		3.34	(2.02)					
Emotion Management	1.09	(0.66)	83.8%	1.05	(0.59)	86.3%	5.0%	7.5%	78.8%	8.8%
Problem Solving	1.12	(0.75)	81.3%	0.79	(0.65)	67.5%	17.5%	3.8%	63.8%	15.0%
Employment	0.64	(0.93)	40.0%	0.45	(0.79)	31.3%	15.0%	6.3%	25.0%	53.8%
Residence	0.39	(0.72)	28.8%	0.39	(0.74)	27.5%	10.0%	8.9%	18.8%	62.5%
Social Influences	0.80	(0.66)	66.3%	0.65	(0.53)	62.5%	23.8%	20.0%	42.5%	13.8%
Total Score	14.79	(5.02)		11.59	(4.30)					

Note. [1]Scores on SOTIPS (Sex Offender Treatment Intervention and Progress Scale) are 0 = minimal or no need for improvement, 1 = some need for improvement, 2 = considerable need for improvement, and 3 = very considerable need for improvement.
[2]Identified Goal defined as a SOTIPS score of 1, 2, or 3, indicating a need for improvement.
[3]Identified Strength defined as a SOTIPS score of 0.

Static-99R and sexual recidivism (Helmus, Hanson et al., 2012).

Vermont Assessment of Sex Offender Risk-2 (VASOR-2)

The VASOR-2 is an actuarial instrument designed to assess sexual recidivism risk and offense severity of adult males known to have been convicted of committing at least one sexual offense. The 12-item reoffense risk scale total scores range from 0 to 22 points and are organized into four risk groups: low (0 to 5), moderate-low (6 to 8), moderate-high (9 to 11), and high (12 to 22). The reoffense risk scale has shown good interrater reliability (ICC = .88, $p < .001$) and moderate predictive ability for sexual recidivism (AUC = .74, $p < .001$; McGrath et al., 2014). The VASOR-2 Severity Checklist is a checklist which does not yield a total score but inventories offense severity characteristics (McGrath, Hoke, & Lasher, 2013).

Procedure

When a client was approximately midway through the prison treatment program ($M = 11.9$ months), he completed a SOTIPS evaluation in his treatment group, which included self-scoring, therapist-scoring, and feedback from peers in the same treatment group. This provided opportunities to educate clients about dynamic risk and protective factors and how SOTIPS scores were used to identify treatment targets, develop treatment plans, and measure treatment progress, which was a factor in determining successful program completion. Groups typically had two therapists present, except in seven cases where a co-therapist was absent. The modal group size was eight clients. For consistency, therapists provided the target client and peers an adapted version of the SOTIPS manual to use for scoring the SOTIPS in group sessions. Using the manual, therapists would explain each SOTIPS item in turn and participants would score it, and then the next item would be explained and scored, and so on and so forth. This manual retained language relevant to scoring the scale in a residential setting, and did not contain SOTIPS scoring criteria pertinent to clients on community supervision.

Although clients were involved in scoring themselves on the SOTIPS at mid-treatment, clients were not involved in scoring themselves during the initial assessment before entering the program or the final assessment at the end of the program. Mental health clinicians completed these pre- and post-treatment SOTIPS assessments independently.

Data Analysis

Client and therapist SOTIPS scores were examined for significant similarities and differences. Statistical similarities were examined using intraclass correlations, and differences were examined using paired-sample *t*-tests. Analyses were conducted for individual items, total scores, and three factors scores (sexual deviance, criminality, and social stability and supports) identified in the SOTIPS development sample analyses (McGrath et al., 2011). Descriptive statistics examined the degree to which therapists and clients agreed that a SOTIPS item represented a treatment goal. Partial correlations were conducted on items and composite score to control for the impact of static risk as measured by the Static-99R and VASOR-2.

The SOTIPS Manual (McGrath, Cumming, & Lasher, 2013) includes scoring instructions that should have constrained participants from selecting certain scores on five items (Sexual Attitudes, Criminal and Rule-breaking Attitudes, Stage of Change, Cooperation with Treatment, and Cooperation with Supervision), and the degree to which participants followed these guidelines was examined.

RESULTS

Therapist and Client Interrater Reliability

Table 2 reports therapist and client SOTIPS mean individual, factor, and total scores. Additionally, the percentage of cases where therapists and clients independently identified each SOTIPS item as a treatment goal is reported. The percentages of cases that showed concordance between therapists and clients that a SOTIPS item met the criteria for a client goal (i.e., SOTIPS score = 1, 2, or 3) or a client strength (i.e., SOTIPS score = 0) are also reported. To illustrate, for the Sexual Offense Responsibility item, 49 (61.3%) therapists' scores were 1 or above, 26 (32.5%) of clients' scores were 1 or above, and in 24 (30%) cases both therapists and clients scored this item 1 or above indicating agreement that it was a potential treatment need for which there was at least "some need for improvement."

Table 3 statistically compares and contrasts therapist and client SOTIPS scores. On average, *t*-tests showed that therapists scored half of the SOTIPS items significantly higher than clients, and the total and factor scores were significantly higher among therapists' scores, indicating that therapists assessed greater levels of need for clients than clients did for themselves. Only on the Sexual Behavior item did clients score themselves higher, on average, than therapists.

Table 3 also shows Cohen's *d* statistics, which further examine differences between therapist and client scores. Cohen (1988) defined effect sizes as small ($0.2 < d < 0.5$), medium ($0.5 < d < 0.8$) and large ($d > 0.8$). Based on these qualifications, the differences between therapist and client scores showed a large difference within the Sexual Deviance factor ($d = 1.28$), a medium difference within the Social Stability and Supports factor ($d = 0.55$), and a small difference within the Criminality factor ($d = 0.41$). The total SOTIPS score showed a large difference between therapist and client scores ($d = 1.03$).

The intraclass correlation statistic was used to examine the degree of similarity in scores between a single client and his therapist. Following Landis and Koch (1977), a "slight" relationship was defined as correlations below 0.20, a "fair" relationship from 0.21 to 0.40, a "moderate" from 0.41 to 0.60, "substantial" from 0.61 to 0.80, and an "almost perfect" from 0.81 to 1.00. Eleven (68.8%) SOTIPS individual items showed significant correlations in therapist and client scores. Seven (63.6%) of these significantly correlated items showed a moderate degree of agreement between therapists and clients and only one (9.1%), Cooperation with Supervision, showed good (i.e., almost perfect) agreement. Criminality, social stability and supports, and sexual deviance factor scores showed substantial, moderate, and fair correlations, respectfully. Total scores also showed significant correlations but were only moderately consistent between therapists and clients.

Impact of Static Risk on Therapist and Client Interrater Reliability

Partial correlations examined therapist and client SOTIPS scores when controlling for Static-99R and VASOR-2 risk levels. A comparison of these statistics to the zero-order interrater statistics indicates that there was very little impact on both items and scales when controlling for static risk levels. The average difference between zero-order and partial correlations for individual items was 0.02 ($SD = 0.02$) and for factor and full scores was 0.03 ($SD = 0.02$). The one exception was that for the Sexual Risk Management item, the significance of inter-rater agreement diminished from ICC = 0.20, $p = .03$ to $pr = 0.21$, $p = .07$ when controlling for either Static-99R or VASOR-2 risk level. This loss of significance appears to be due to the added static risk variable in the analysis, however this is overshadowed by the fact that the actual degree of agreement on this item was relatively poor.

Scoring Error Analysis

With respect to how well clients followed SOTIPS scoring rules that should have constrained them from selecting certain scores on five items (Sexual Attitudes, Criminal and Rule-breaking Attitudes, Stage of Change, Cooperation with Treatment, and Cooperation with Supervision), four clients (5%) scored two items incorrectly and 20 clients (25%) scored one item incorrectly.

TABLE 3
Comparisons of Therapist and Client SOTIPS Scores ($N = 80$)

Item/Factor/Scale	Mean Score Difference (t) [1]	Effect Size (d)	Effect Size 95% CI	ICC
Sexual Deviance Factor	8.08***	1.28	0.93 – 1.62	0.23*
Offense Responsibility	5.60***	0.89	0.56 – 1.21	0.43***
Sexual Behavior	−3.96***	0.63	0.31 – 0.94	0.08
Sexual Attitudes	6.94***	1.10	0.76 – 1.43	0.12
Sexual Interests	5.52***	0.87	0.55 – 1.20	0.17
Sexual Risk Management	6.99***	1.11	0.77 – 1.44	0.20*
Stage of Change	1.62	0.26	−0.06 – 0.57	0.14
Criminality Factor	2.57*	0.41	0.09 – 0.72	0.75***
Criminal Behavior	−0.66	0.10	−0.21 – 0.41	0.35***
Criminal Attitudes	3.75***	0.59	0.28 – 0.91	0.48***
Cooperation with Treatment	2.28*	0.36	0.05 – 0.67	0.61***
Cooperation with Supervision	−0.26	0.04	−0.27 – 0.35	0.83***
Impulsivity	1.95	0.31	0.00 – 0.62	0.53***
Sociality Stability & Supports Factor	3.45***	0.55	0.23 – 0.86	0.59***
Emotion Management	0.52	0.08	−0.22 – 0.39	0.47***
Problem Solving	4.91***	0.78	0.46 – 1.10	0.62***
Employment	2.10*	0.33	0.02 – 0.64	0.58***
Residence	0.00	0.00	−0.31 – 0.31	0.50***
Social Influences	1.59	0.25	−0.06 – 0.56	0.01
Total Score	6.50***	1.03	0.69 – 1.36	0.56***

Note. [1]df = 79.
*$p < .05$. **$p < .01$. ***$p < .001$.

DISCUSSION

The present study compared the SOTIPS self-evaluation scores of clients in a prison sex offender treatment program with the SOTIPS scores of the clients' therapists. SOTIPS total scores indicate a quantified level of dynamic risk, and individual item scores can serve as a proxy for what risk factors clients and their therapists see as important treatment targets. Following previous research which has demonstrated an association of positive treatment outcomes when clients and therapists agree on treatment targets (e.g., Owen et al., 2012; Tryon & Winograd, 2011), we hoped to see a good level of agreement between clients and their therapists on SOTIPS scores.

The present findings indicate that, based on the interrater analysis of total SOTIPS scores, therapists and clients show moderate agreement in overall treatment goals, as well as dynamic risk. This suggests these clients are somewhat cognizant of their overall dynamic risk and treatment need. Nevertheless, analyses of mean total scores revealed that therapists score clients significantly higher than clients score themselves. Furthermore, this relationship appears to be independent of clients' level of static risk. Examining differences in scores broken down by the three major SOTIPS factors assists in explaining some of these findings.

Criminality

The criminality factor showed the greatest degree of agreement between therapists and clients. The highest degree of agreement was on the items Cooperation with Treatment and Cooperation with Supervision. This is not surprising given that most of these items have relatively concrete scoring criteria.

Clients scored themselves higher on the Criminal Behavior item than their therapists did, although not to a significant degree. Therapists reported that prison security and other nontreatment staff often provided information that they used to score this item. However, clients often disclosed previously undetected minor rule-breaking behaviors during SOTIPS scoring. The disclosure of rule-breaking behavior may have involved several factors, such as the individual's degree of antisocial orientation, degree of institutionalization, and the cost-benefit analysis involved in disclosing rule-breaking behavior. These factors may be related to trait Neuroticism, which has recently been implicated in the willingness of prisoners to disclose such behaviors (Ferguson, Ireland, & Ireland, 2013).

Conversely, therapists scored clients higher on the Criminal Attitudes item. As shown in Table 3, client scores for Criminal Attitudes were, on average, slightly higher than Criminal Behaviors scores. In general, clients tended to identify problem behaviors more so than the attitudes and beliefs underlying these behaviors. Given that the cognitive-behavioral model is arguably the most empirically supported approach for treating adult male sexual offenders (Hanson et al., 2009; Lösel & Schmucker, 2005), a balanced treatment emphasis on helping clients learn and practice prosocial behaviors as well as ways of thinking is clearly warranted.

Regarding the Cooperation with Treatment item, therapists might have scored clients higher on this item due to poor participation in group therapy sessions. Clients, on the other hand, would assert they deserved more favorable scores due to a focus on either concrete issues like completing assignments on time or positive behaviors following a written warning early in the six-month scoring window. It is worth acknowledging, though, that this item may be prone to inflated scores on the part of the therapist. Therapists may have difficulty separating some of the personality issues an offender presents from the evaluation of the offender's degree of active participation in treatment. This is a reminder that therapists in general may be susceptible to counter-transference with this population, and as Newman (1997) has noted, this can occur among even the most well-intentioned therapists.

Social Stability and Supports

In contrast to the criminality factor, the level of agreement between therapists and clients on social stability and supports items was only moderate in degree. As seen in Table 3, the only item within this factor that showed significant differences between mean scores was Problem Solving. On the other hand, Social Influences showed the lowest degree of agreement between therapists and clients across all SOTIPS items.

The disagreement between therapists and clients on Social Influences may be explained by how this item was discussed in treatment sessions. As clients were subject to greater antisocial influences in prison than would be expected in the community, it was common for therapists to score clients a 1 on this item unless they had demonstrated a significant effort to associate with individuals who were attempting to live a prosocial prison lifestyle as opposed to those who, for example, regularly broke facility and program rules. Of course, some clients may have simply disagreed that their social influences were negative and therefore saw no need to change these influences. Regardless, many clients viewed their social habits as more positive than did their therapists. In cases where therapists believed clients demonstrated an effort to avoid negative influences, these clients would sometimes score themselves higher on the premise that it is impossible to avoid the antisocial influences inherent in prison. Key to these discussions was how individuals with whom a person chooses to associate can have a powerful influence on one's behavior (Mann et al., 2010).

The finding of few differences between therapists and clients on the Emotion Management item was contrary to therapists' common expectation that clients lack emotional awareness. However, this finding supports a previous study showing offenders are not less emotionally aware than the general population (Paxton, 1995). Emotion management issues are often paid extra attention in group therapy when evident. There may be some internalizing of these discussions, resulting in at least a basic understanding of emotional awareness by the time clients are midway through treatment.

Sexual Deviance

This factor showed the greatest level of discrepancy between mean scores and the worst degree of agreement. One notable exception was the inter-rater reliability on Offense Responsibility, which was higher than the factor itself, despite the average therapist scores being significantly higher than client scores. Furthermore, regarding the Stage of Change item, the relationship between therapist scores and client scores was not significant. As can be seen in Table 2, though, all therapists and clients rated this item a score of 1 or above, as SOTIPS scoring rules exclude a score of 0 on this item while an individual is in prison. Therefore it may be that there were too few deviations from the client being rated at the "active stage" on this item (score = 1) to produce statistically significant differences in mean scores.

A surprising finding within this factor concerned the Sexual Behavior item. On average, therapists scored clients lower than clients scored themselves. Clients who scored themselves as having a treatment need in this area often noted that they were breaking facility rules when they masturbated, even though they did so in private. Conversely, therapists commonly scored clients higher on the sexual attitudes item. It appeared particularly challenging for some clients to identify current sexual offense–related attitudes as well as healthy sexual outlets.

It was not surprising that therapists would regularly score clients higher on the Sexual Risk Management item. This item focuses on both understanding and practicing risk management strategies. It was common for clients to have a deficient understanding of risk factors but not engage in any obvious "risky" behavior. Overall, differences in therapist versus client scores on Sexual Risk Management, Sexual Interests, and Sexual Attitudes items suggest clients find it easier to focus on concrete and observable behaviors rather than the thinking underlying those behaviors.

Impact of Static Risk

The diversity of participants' static risk levels (see Table 1) enabled us to test the impact of clients' static risk on their ability to identify treatment goals. Given that sex offender treatment programs that follow the risk principle show reductions in reoffending (Hanson et al., 2009), it is encouraging that higher risk clients appeared as able to reach agreement with their therapist about treatment goals as were lower risk clients.

Comparison to Previous Interrater Reliability Analysis

The initial SOTIPS development study contained data with which to compare interrater reliability information between clients' therapists and probation officers ($N = 320$; McGrath et al., 2011) with interrater reliability information between therapists and clients in the present study. Among individual items, the mean difference in ICC between the initial SOTIPS development study and the current analysis was only 0.18 ($SD = 0.15$). Notable exceptions (that is, one standard deviation above or below the mean difference) included greater differences on Sexual Behavior (2011 ICC = 0.51, current ICC = 0.08), Stage of Change (2011 ICC = 0.40, current ICC = 0.14), and Social Influences (2011 ICC = 0.43, current ICC = 0.01). Few differences were seen on Cooperation with Treatment (2011 ICC = 0.63, current ICC = 0.61) and Impulsivity (2011 ICC = 0.51, current ICC = 0.53).

Not surprisingly, there are differences between the present study and SOTIPS development study sexual deviance items (2011 ICC = 0.68, current ICC = 0.23). Conversely, interrater reliability was similar for the criminality factor (2011 ICC = 0.76, current ICC = 0.75) and the social stability and support factor (2011 ICC = 0.69, current ICC = 0.59). Given that the criminality and social stability and support factors showed few differences on factor and item interrater performance, clients in the present study appeared able to comprehend SOTIPS scoring instructions. Furthermore, the differences between interrater scores on the sexuality factor suggests that the client scores overall reflected a different perception of their treatment need and functioning on these items overall.

Limitations and Future Directions

Some clients appear to have had difficulty conceptualizing some of the SOTIPS scoring rules, particularly abstract rules, and this is a limitation of this study. Although the main goal of these analyses was to examine if therapists and clients agree on treatment targets, we have a limited capability to say why therapists and clients agreed or disagreed on a SOTIPS item. For example, the Employment and Residence items are scored on concrete measures of change (i.e., job changes or residence changes), but also on the degree that clients are satisfied with their status. Therapists often rely on the concrete factors associated with these items because they have access to reports that inform them about occupational and residential stability. However, the question of satisfaction with employment and residence is not as well attended to, and therefore may not be considered fully by either therapists or clients.

In order for therapists and clients to score SOTIPS items similarly, they should possess the same information. The prison environment facilitates therapists' knowledge of clients' behaviors, both positive and negative. However, this strength is attenuated by the degree to which non-treatment prison staff report information to therapists. This was, unfortunately, both inconsistent and unpredictable.

Another study limitation is that it did not include measures of social desirability. Sexual thoughts and behavior are typically much more private and difficult to talk about than other types of thoughts and behavior, and more shame is associated with sexual offending than for most other life problems. Clients may have under-reported some treatment needs due to shame and to be viewed in a favorable light as possible in order to move successfully through treatment and gain early release. Future research in this area would benefit from including measures that might identify and account for these and other potential self-report biases.

Finally, some of the SOTIPS areas may not have been fully addressed before the mid-treatment assessment. For example, some clients may not fully understand their sexual or criminal attitudes, let alone be able to change these attitudes early in treatment. Based on therapist experiences, this may be an issue that clients do not significantly assimilate until some point beyond the mid-treatment SOTIPS assessment. Additionally, we did not systematically collect feedback from clients about their experiences collaboratively completing the SOTIPS. In the future, structured examination of clients' perspectives on using the SOTIPS for collaborative assessment and treatment planning may be useful in being responsive to clients' needs.

Conclusion

Structured, collaborative treatment planning appears to have many benefits, and therapist and client agreement on treatment goals has been associated with positive treatment outcomes (Abel et al., 1988; Owen et al., 2012; Tryon & Winograd, 2011; Veysey, 2008). In the present study, the SOTIPS appeared to be a useful collaborative treatment planning tool, identifying and helping therapists and clients reach agreement on treatment goals. It facilitated open discussion of clients' strengths, potential treatment needs, and treatment progress. Comparison of therapist and client scores allowed for identification of areas of agreement, as well as areas of disagreement. Further research should examine how the collaborative treatment planning process impacts outcomes in treatment, both in terms of progress made in treatment, such as changes in SOTIPS scores, as well as how this impacts the overall treatment goal of reducing recidivism.

REFERENCES

Abel, G. G., Mittelman, M. S., Becker, J. V., Rathner, J., & Rouleau, J. L. (1988). Predicting child molesters' response to treatment. *Annals of the New York Academy of Science*, 528, 223–234. doi:10.1111/j.1749-6632.1988.tb50866.x

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson.

Andrews, D., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation rediscovering psychology. *Criminal Justice and Behavior*, *17*(1), 19–52. doi:10.1177/0093854890017001004

Beech, A. R., & Hamilton-Giachritsis, C. F. (2005). Relationship between therapeutic climate and treatment outcome in group-based sexual offender treatment programs. *Sexual Abuse: A Journal of Research and Treatment*, *17*, 127–140. doi:10.1007/s11194-005-4600-3

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Ferguson, K., Ireland, C. A., & Ireland, J. L. (2013). Developing a self-report measure to assess disclosure strategies in adult male prisoners and its association with personality. *Journal of Forensic Practice*, *15*(2), 97–108. doi:10.1108/14636641311322287

Hanson, R. K., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders. *Criminal Justice and Behavior*, *36*, 865–891. doi:10.1177/0093854809338545

Hanson, R. K., Harris, A. J. R., Scott, T. L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project*. Ottawa, ON: Public Safety Canada.

Hanson, R. K., & Morton-Bourgon, K. E. (2004). *Predictors of sexual recidivism: An updated meta-analysis.* (User Report 2004-02). Ottawa, ON: Public Safety and Emergency Preparedness Canada.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73*, 1154–1163. doi:10.1037/0022-006X.73.6.1154

Harris, A., Phenix, A., Hanson, R. K., & Thornton, D. (2003). *Static-99 coding rules: Revised 2003*. Ottawa, ON: Department of the Solicitor General of Canada.

Helmus, L., Hanson, R. K., Thornton, D., Babchishin, K. M., & Harris, A. J. (2012). Absolute recidivism rates predicted by Static-99R and Static-2002R sex offender risk assessment tools vary across samples: a meta-analysis. *Criminal Justice and Behavior*, *39*(9), 1148–1171. doi:10.1177/0093854812443648

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*(1), 64–101. doi:10.1177/1079063211409951

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. doi:10.2307/2529310

Levenson, J. S. & Macgowan, M. J. (2004). Engagement, denial, and treatment progress among sex offenders in group therapy. *Sexual Abuse: A Journal of Research and Treatment*, *16*, 49–63. doi:10.1023/B:SEBU.0000006284.33837.d7

Levenson, J. S., Macgowan, M. J., Morin, J. W., & Cotter, L. P. (2009). Perceptions of sex offenders about treatment: Satisfaction and engagement in group therapy. *Sexual Abuse: A Journal of Research and Treatment*, *21*, 35–56. doi:10.1177/1079063208326072

Levenson, J. S., & Prescott, D. S. (2009). Treatment experiences of civilly committed sex offenders a consumer satisfaction survey. *Sexual Abuse: A Journal of Research and Treatment*, *21*(1), 6–20. doi:10.1177/1079063208325205

Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, *1*, 117–146. doi:10.1007/s11292-004-6466-7

Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, *22*, 191–217. doi:10.1177/1079063210366039

Marques, J. K., Weideranders, M., Day, D. M., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, *17*, 79–107. doi:10.1007/s11194-005-1212-x

McGrath, R. J., & Cumming, G. F. (2001, 2003). *The Sex Offender Treatment Needs and Progress Scale – Version 2001; 2003*. Middlebury, VT: Author.

McGrath, R. J., Cumming, G. F., Burchard, B. L., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American survey.* Brandon, VT: Safer Society Press.

McGrath, R. J., Cumming, G. F., & Lasher, M. P. (2013). *The Sex Offender Treatment Intervention and Progress Scale (SOTIPS) Manual – Version 2013*. Middlebury, VT: Author.

McGrath, R. J., Cumming, G., Livingston, J. A., & Hoke, S. E. (2003). Outcome of a treatment program for adult sex offenders: From prison to community. *Journal of Interpersonal Violence*, *18*, 3–17.

McGrath, R. J., Hoke, S. E., & Lasher, M. P. (2013). *Vermont Assessment of Sex Offender Risk –2 (VASOR-2) manual*. Middlebury, VT: Author.

McGrath, R. J., Lasher, M. P., & Cumming, G. F. (2011). *A model of static and dynamic risk assessment* (Document No. 236217). Washington, DC: United States Department of Justice.

McGrath, R. J., Lasher, M. P., & Cumming, G. F. (2012). The Sex Offender Treatment Intervention and Progress Scale (SOTIPS): Psychometric properties and incremental predictive validity with Static-99R. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 431–458. doi:10.1177/1079063211432475

McGrath, R. J., Lasher, M. P., Cumming, G. F., Langton, C. M., & Hoke, S. E. (2014). Development of Vermont Assessment of Sex Offender Risk-2 (VASOR-2) Reoffense Risk Scale. *Sexual Abuse: A Journal of Research and Treatment*, *26*(3), 271–290. doi:10.1177/1079063213486936

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing for change* (2nd ed.). New York, NY: Guilford.

Newman, C. F. (1997). Maintaining professionalism in the face of emotional abuse from clients. *Cognitive and Behavioral Practice*, *4*(1), 1–29. doi:10.1016/S1077-7229(97)80010-7

Olver, M. E., Wong, S. C. P., Nicholaichuk, T. P., & Gordon, A. E. (2007). The validity and reliability of the Violence Risk Scale–Sexual Offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment*, *19*, 318–329. doi: 10.1037/1040-3590.19.3.318

Owen, J., Duncan, B., Anker, M., & Sparks, J. (2012). Initial relationship goal and couple therapy outcomes at post and six-month follow-up. *Journal of Family Psychology*, *26*, 179–186. doi:10.1037/a0026998

Paxton, L. B. (1995). *Narcissism, anger, and depression in the prediction of emotional awareness and empathy in perpetrators of child sexual abuse* (Doctoral dissertation, California School of Professional Psychology, San Diego).

Proulx, J., Tardif, M., Lamoureux, B., & Lussier, P. (2000). How does recidivism risk assessment predict survival? In D. R. Laws, S. M. Hudson and T. Ward (Eds.), *Remaking relapse prevention with sex offenders* (pp. 466–484). Newbury Park, CA: Sage.

Salter, A. C. (1988). *Treating child sex offenders and victims*. Thousand Oaks, CA: Sage.

Shingler, J., & Mann, R. E. (2006). Collaboration in clinical work with sexual offenders: Treatment and risk assessment. In W. L Marshall et al. (Eds.), *Sexual offender treatment: Controversial issues* (pp. 225–239). West Sussex, UK: Wiley.

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment*, *14*, 139–153. doi:10.1023/A:1014620214905

Tryon, G. S. & Winograd, G. (2011). Goal consensus and collaboration. *Psychotherapy*, *48*, 50–57. doi:10.1037/a0022061

Veysey, B. (2008). Rethinking reentry. *The Criminologist*, *33*, 1–5.