

**RISKY BUSINESS: A REGIONAL COMPARISON OF THE LEVELS OF RISK AND
SERVICE NEEDS OF SEXUALLY OFFENDING YOUTH**

by

Tara McKenzie Schoenfeld

B.A. Honours, Okanagan University College, 2003

A THESIS SUBMITTED IN PARTIAL FULLFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

The College of Graduate Studies

(Interdisciplinary Studies)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

April 2008

© Tara McKenzie Schoenfeld, 2008

ABSTRACT

Considerable attention has focussed on identifying individual factors associated with, or predictive of, sexual offending (e.g., Efta-Breitbach & Freeman, 2004). In light of these individual factors, clinicians and researchers have developed standardized instruments for assessing the risk posed by sexually offending youth. Two such instruments are the Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II; Prentky & Righthand, 2003) and the Estimate of Risk of Adolescent Sexual Offence Recidivism Version 2.0 (ERASOR-II; Worling & Curwen, 2001). In addition to individual factors, research on crime has demonstrated that structural factors within the community may be important determinants of sexual and non-sexual offending (e.g., McCarthy, 1991; Ouimet, 1999; Shaw & McKay, 1942; Wirth, 1938). Therefore, the purpose of this study was twofold: (a) to compare the psychometric properties of two newly developed risk assessment instruments (i.e., J-SOAP-II and ERASOR-II) and (b) to use the better instrument to compare the levels of risk posed by sexually offending youth in 3 neighbouring, but diverse communities. Using file information, the J-SOAP-II and ERASOR-II were scored on 84 adolescent males between the ages of 11 and 20 years who had committed a sexual offence and received treatment at Youth Forensic Psychiatric Services (YFPS) in the Greater Vancouver Area (GVA; $n = 30$), Central Okanagan (CO; $n = 26$), and Thompson Nicola region (TN; $n = 28$). Calculations of interrater reliability and item-total correlations indicated that the J-SOAP-II was a better assessment instrument for this sample of offenders. Consequently, further regional analysis of risk was conducted using the J-SOAP-II data. Results indicated that although there were no regional differences among the severity and history of sexual offending, TN youth generally had a greater number of risk factors than did youth in CO and GVA. Specifically, youth in TN were found to be higher risk in the areas of intervention, general problem behaviour,

and family/environment dynamics. These results suggest that to better understand youth who commit sexual offences and to provide appropriate prevention and intervention strategies for individual offenders and their communities, youth should not be evaluated in isolation from their social and community context. Recommendations for practice are discussed.

TABLE OF CONTENTS

ABSTRACT.....	ii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
ACKNOWLEDGEMENTS.....	viii
CO-AUTHORSHIP STATEMENT.....	ix
1 INTRODUCTION: INDIVIDUAL AND COMMUNITY RISK OF SEXUALLY OFFENDING YOUTH.....	1
1.1 Assessing Risk to Reoffend	3
1.2 Communities and Crime	7
1.3 Purpose of Study	9
References.....	11
2 A COMPARISON OF THE JUVENILE SEX OFFENDER ASSESSMENT PROTOCOL- II (J-SOAP-II) AND THE ESTIMATE OF RISK OF ADOLESCENT SEXUAL OFFENCE RECIDIVISM VERSION 2.0 (ERASOR-II).....	17
2.1 Introduction.....	17
2.1.1 Assessing Risk to Reoffend	18
2.1.2 Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II).....	21
2.1.3 Estimate of Risk of Adolescent Sexual Offense Recidivism Version 2.0 (ERASOR-II)	23
2.1.4 Present Investigation.....	25
2.2 Method.....	26
2.2.1 Participants.....	26
2.2.2 Procedure	26
2.3 Results.....	27
2.3.1 Interrater Reliability.....	27
2.3.2 Structure of the J-SOAP-II.....	30
2.3.3 Structure of the ERASOR-II.....	34
2.3.4 Concurrent Validity	41
2.4 Discussion.....	41
References.....	50
3 A MULTILEVEL ANALYSIS OF SEXUALLY OFFENDING YOUTH.....	55
3.1 Introduction.....	55
3.1.1 Individual-level Research: Identifying and Assessing Risk to Reoffend	57
3.1.2 Ecological Research: Evaluating the Association Between Communities and Crime.....	58
3.1.3 Purpose of Study	63
3.2 Method.....	64
3.2.1 Participants.....	64
3.2.2 Classifying Regions – Independent Variable.....	66
3.2.3 Categorizing Risk Factors – Dependent Variable.....	68
3.2.4 Background Information.....	71
3.2.5 Procedure	71
3.3 Results.....	72

3.3.1	Region and Risk	72
3.3.2	YSOTP Initiation, Duration, and Completion	76
3.4	Discussion	76
	References	86
4	CONCLUSION: RECOMMENDATIONS FOR PRACTICE	92
	References	103
	APPENDIX	105

LIST OF TABLES

Table 2-1 Interrater Agreement on Individual J-SOAP-II Items	29
Table 2-2 Interrater Agreement on Individual ERASOR-II Items	30
Table 2-3 Internal Consistency of J-SOAP-II Scales	31
Table 2-4 Internal Consistency and Endorsement Patterns of Individual ERASOR-II Items	36
Table 2-5 Internal Consistency of ERASOR-II Scales	38
Table 3-1 Average Age and Ethnicity of Offenders in Each Region	66
Table 3-2 Location of Residence at Assessment	68
Table 3-3 Mean and Standard Deviation of J-SOAP-II Scale Scores	73

LIST OF FIGURES

Figure 2-1 Confirmatory Factor Analysis of J-SOAP-II Scales 33

Figure 2-2 Exploratory Factor Analysis of ERASOR-II Scales 40

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to all the people who made the completion of this thesis possible. A special thanks to my supervisor Dr. Michael Woodworth for introducing me to forensic research and for providing much needed support, input, and guidance along the way. Heartfelt thanks to my supervisory committee members Dr. Peter Urmutzer and Dr. Jan Cioe who have provided invaluable advice and support throughout my graduate and undergraduate education. Thank you to all of the supportive and accommodating staff at Youth Forensic Psychiatric Services. In particular, thank you to Moreen Tremblay, Dr. Heather Gretton, and Steve Sigmond for their assistance, encouragement, and excitement throughout this project. A very special thanks to Janice, Gordana, Mary and Caroline who quickly and diligently searched for, ordered, and shipped countless files for me. Thank you, as well, to all of the research assistants at YFPS in Burnaby for training and assistance with coding.

Thank you to my friends and family for their ongoing support and encouragement throughout my many years of graduate and undergraduate education. To my close friends, Jonathan and Aarin, I am so very grateful for too many reasons to list. Most importantly, without you, finishing would not have been possible. And to my partner and best friend Randall, your unwavering support and encouragement has guided me through the most trying of times. I am tremendously grateful for you.

CO-AUTHORSHIP STATEMENT

This thesis was prepared under the supervision and guidance of Dr. Michael Woodworth and with the assistance of a number of individuals named in the acknowledgements page. As a result, some sections of this thesis are co-authored. Consequently, this statement is required to illustrate that I have adequately contributed to the design and implementation of the research program and the preparation of this thesis. The details of my contribution were as follows. I designed the research program by selecting the particular instruments and variables to be coded. I implemented the program by coding files in Burnaby in exchange for risk assessment training and the Burnaby data. In addition, I conducted all of the initial coding of the Kamloops, Kelowna, and Penticton files. With the assistance of Jonathan Brown, I conducted the data analysis for manuscript one. I conducted all of the data analysis for manuscript two. With respect to manuscript preparation, I conducted the literature reviews, formulated the conceptual framework of each of the manuscripts, and wrote the thesis.

1 INTRODUCTION: INDIVIDUAL AND COMMUNITY RISK OF SEXUALLY OFFENDING YOUTH

There is a growing concern among mental health and justice professionals regarding the suitable strategies for preventing and managing adolescent sexual offending (Harris, Rice, & Quinsey, 1998; Moore, Franey, & Geffner, 2004). Until recently, the study of sexual offending has largely focussed on adult males as they have long been assumed to commit the majority of sex crimes. Recent research, however, has indicated that male adolescents are responsible for one in every five sexual assaults against individuals 12 years and older (Maguire & Pastore, 2002 as cited in Moore et al., 2004). With respect to victims under the age of 12, male adolescents account for one in every two incidents of male child sexual victimization and one in three incidents of female child victimization (Zonana, Abel, Bradford, Hoge, & Metzner, 1998 as cited in Moore et al.; Ryan, 1999). In addition, as much as 50% of all adult sexual offenders have reported some sexually deviant behaviour or interest during their youth (Abel, Mittleman, Becker, Rathner, & Rouleau, 1988; Abel & Rouleau, 1990). Given the prevalence of adolescent sexual offending and the possibility of this offending progressing into adulthood, researchers and clinicians have begun to devise assessment, prevention, and intervention techniques that are suitable for youth who commit sexual crimes.

Mental health and justice professionals generally agree that sexual offending is a serious crime with often devastating results. However, a general consensus is lacking about the most appropriate and effective strategies for managing both adult and adolescent sexual offending. There is considerable public pressure to assign sexual offenders' uniform, harsh, and lengthy punishment because of the common misconception that all sexual offenders are similar, untreatable, and incurable (Fedoroff & Moran, 1997). It is evident, however, that simply

detaining all offenders indefinitely would be too costly and inappropriate for many offenders. This is particularly the case for adolescents who commit sexual offences in Canada because harsh and lengthy punishment conflicts with the philosophy of the *Youth Criminal Justice Act* (YCJA). The YCJA was established to promote rehabilitation and to reduce an over-reliance on incarceration for young offenders (Department of Justice Canada, n.d.). In addition, there is little, if any, empirical evidence to support the notion that most sexual offenders are incurable and prone to recidivism (Hall, 1995; Schwartz & Cellini, 1995). Therefore, to adhere to the guidelines proposed by the YCJA and to follow evidence-based practice, clinicians and justice professionals have a responsibility to provide intervention strategies that are suitable for individual offenders, are directed at rehabilitation rather than incarceration, and have the ultimate goal of preventing further harm to the community.

To help achieve these goals, this paper evaluates two main areas of research that are largely lacking for sexually offending youth. The first section of this paper illustrates the vital role that risk assessments play in managing young offenders within the justice system. It is also argued that offenders could benefit from the consideration of individual risk, as well as risk inherent within the structural factors of communities. The second section reports on two studies that have distinct objectives for assessing the risk of sexually offending youth. The first study evaluated some of the psychometric properties of two newly developed risk assessment instruments designed specifically for sexually offending youth. The second study was a multilevel analysis of sexually offending youth by assessing the individual risk factors of offenders, as well as the community context in which they live. The final section of this paper provides recommendations for using this information in practice.

1.1 Assessing Risk to Reoffend

A key issue in assessing risk is to distinguish between different types of assessments (Douglas & Skeem, 2005 as cited in Tolman & Rotzien, 2007). Indeed, there is more to assessing risk than simply predicting future acts of violence. For instance, Heilbrun (1997) argues that there are two distinct approaches to risk assessment depending on the legal context in which they are requested. The first is the prediction approach which focuses on quantifying an offender's level of risk for one-time assessments, such as for sentencing, with little interest in preventing and managing the offending behaviour. The second is the risk management approach, which assesses risk with the goal of providing interventions that will ultimately reduce the risk for future violence. Tolman and Rotzien (2007), however, argue that these approaches are no longer distinct. Contemporary risk assessments have already begun to incorporate prevention and management principles into the prediction context of assessments.

Epperson, Ralston, Fowers, DeWitt, and Gore (2006) outline a number of ways these contemporary risk assessments have proven useful in managing offenders' within the justice system. First, they argue that the absence of risk assessments produces a "one-size-fits-all" approach to risk management. This approach can result in interventions that are too intense and costly for low risk offenders and that are unable to reduce the threat posed by high risk offenders (Epperson et al, 2006). Second, assessing each offender's risk level can reduce the potential for mismatched services. Research has shown that intense treatment can actually have detrimental consequences for some low risk offenders (Boxer, Guerra, Huesmann, & Morales, 2005; Gifford-Smith, Dodge, Dishion, & McCord, 2005; Mager, Milich, Harris, & Howard, 2005). In some cases, for some very low risk offenders, a basic psychoeducational program may be all that is needed to reduce the potential for committing future sexual crimes (Epperson et al.). When

more intense treatment is required, risk assessments can provide information about particular areas to target in treatment. For example, the treatment for an offender who commits sexual crimes under the influence of substances and feels guilt and remorse for the offending behaviour needs a much different focus than the treatment for an offender who feels no guilt and remorse and generally lacks any empathy for his victims. Matching services to the risk posed by offenders can result in more successful risk reduction than providing services that are inappropriate and irrelevant for particular offenders.

Further, risk assessments can help avoid the misuse of resources, which are generally limited within the justice system (Epperson et al., 2006). For instance, risk assessments can help differentiate between the high risk offenders who are in need of intensive treatment and low risk offenders who would equally benefit from receiving fewer resources and lower levels of treatment. It seems quite inappropriate to waste valuable resources on an offender whose risk to reoffend is already so low it cannot be reduced much lower (Epperson et al.). Resources saved with low risk offenders could be better invested in more intensive services for high risk offenders potentially resulting in greater community safety.

To protect the public's safety, it is advisable to assess the risk of adult and adolescent offenders who have committed violent crimes and therefore, have the potential to commit future acts of violence. Moreover, assessing the risk of sexually offending youth may be particularly helpful in reducing threat to the public while simultaneously providing services that adhere to the guidelines of the YCJA. Research has shown that a relatively small group of sexually offending youth commit repeat offences after there has been formal intervention and that most of those who do reoffend tend to do so by committing nonsexual crimes (Righthand & Welch, 2001). If mismatched treatment can actually have a detrimental effect on some offenders, treatment

modalities not guided by risk assessments may be inappropriate for some offenders paradoxically increasing the risk for recidivism. Moreover, treatment guided by risk assessments may be more likely to target areas of need, such as antisocial attitudes and nonsexual offending behaviours, which could ultimately reduce the risk to commit future nonsexual crimes. In addition, risk assessments should be used with this population to adequately identify and provide intensive services for the small group of offenders who are likely to reoffend sexually.

Given the importance of assessing risk, one would expect that many standardized assessment instruments would have been developed for sexually offending youth. Research regarding this population, however, remains limited and as a result there are few instruments developed specifically for this population. While there are a number of tools devised to assess risk for general violence among youth (e.g., PCL:YV, Forth, Kosson, & Hare, 2003; YLS/CMI, Hoge & Andrews, 1996; SAVRY, Borum, Bartel, & Forth, 2002) and sexual violence among adults (e.g., SORAG, Quinsey, Rice, & Harris, 1995; STATIC-99, Hanson & Thornton, 2000), there are few risk assessment instruments that have been developed specifically for youth who commit sexual crimes.

Developing instruments that accurately assess the risk of young offenders has proven challenging because the assessment of youth requires a number of special considerations. First, adolescence is a period of considerable change. Youth during this developmental stage are still growing and maturing as they progress toward adulthood. As such, risk assessment instruments for this population should consider the dynamic nature of young offenders and include items that reflect this development. Second, youth differ considerably from adults in that they are often still heavily reliant on their families. Considering that an offender's family can act as either a risk or protective factor (e.g., Fagan, Van Horn, Hawkins, & Arthur, 2007; Hall-Lande, Eisenberg,

Christenson, & Neumark-Sztainer, 2007; Laub & Lauritsen, 1994), it is crucial that items pertaining to family dynamics are included in assessment instruments for this population. Third, youth who commit sexual crimes are more likely to reoffend by committing nonsexual than sexual offences (Caldwell, 2002; Hanson & Bussière, 1998). Risk assessment instruments that consider a variety of factors, including items pertaining to general offending and antisocial behaviour, are needed to adequately assess for future problem behaviour. To the best of our knowledge there are only two risk assessment instruments that have been developed specifically for youth who have committed sexual crimes and that also take into account all of these special considerations.

The Juvenile Sex Offender Assessment Protocol–II (J-SOAP–II; Prentky & Righthand, 2003) and the Estimate of Risk of Adolescent Sexual Offence Recidivism Version 2.0 (ERASOR-II; Worling & Curwen, 2001) are comprised of items that have been found to be predictive of sexual offending. In addition, they incorporate in their items the unique and dynamic characteristics of adolescents who sexually offend.

Results of research measuring the psychometric properties of the J-SOAP-II and the ERASOR-II has been promising (e.g., Righthand et al., 2005; Worling, 2004); however, both instruments have only recently been developed and as a result have only undergone preliminary testing. Both of these instruments would, therefore, benefit from further evaluation of their psychometric properties. For instance, Doren (2006) argues that research measuring the predictive validity of the ERASOR-II has had mixed results (Morton & Bourgon, 2003) and it has only been minimally tested for its interrater reliability (Worling, 2004) and for its concurrent validity (Bourgon, 2002). The J-SOAP-II, conversely, has undergone more empirical testing than the ERASOR-II, but is still limited by inadequate testing of construct and predictive validity.

Given that risk-assessment instruments can aid in delivering interventions that reduce recidivism, it is crucial that instruments that have been developed specifically for sexually offending youth undergo further evaluation.

1.2 Communities and Crime

Intervention strategies for sexually offending youth have traditionally focused on the individual risk factors and characteristics of offenders. Research on communities and crime, however, has demonstrated that structural factors can also influence crime rates and offending patterns (Jacob, 2006; Ouimet, 2000; Sampson, 1997). Ecological researchers have long argued the necessity of evaluating crime at the structural-level because variables such as population size, density (e.g., population per square kilometre), socioeconomic status (SES), cultural heterogeneity, and residential mobility have been associated with both violent and non-violent crime rates (Osgood & Chambers, 2000; Sampson, 1997; Sampson & Groves, 1989; Shaw & McKay, 1942; Wirth, 1938). In addition, prevention and intervention are often offered at the community level, through the media of schools, courts, and mental health centers. To provide effective services for sexually offending youth it may be beneficial to consider structural risk factors, as well as the risk factors inherent in the individual.

Research investigating the association between structural factors and crime suggests that crime rates and offending patterns tend to vary with the size and composition of the community. For example, the investigation of national statistics for Canada, the United States, and Europe demonstrated that serious crime rates tended to increase with the size of the city (Nettler, 1978 as cited in Hartnagel & Lee, 1990). Hartnagel and Lee (1990) found that larger sized cities had higher crime rates, but the effect became nonsignificant when other factors were considered (e.g., socioeconomic status, residential instability, heterogeneity). Osgood and Chambers (2000)

demonstrated that in addition to population size, juvenile violence was also associated with rates of residential instability, family disruption, and ethnic heterogeneity. The fact that crime rates vary by region suggests that there may be community factors that are heightening the risk of its residents.

The theoretical underpinning of ecological research is that structural factors influence the organization of communities and community members. For instance, as population size increases, greater variation occurs among the community's residents, resulting in poor communication between community members and segregation according to race, ethnicity, and socioeconomic status (Jacob, 2006). Social disorganization and individual alienation are consequences of increased population because there is a lack of commonality among community members. Consequently, there is a lack of social networks and supports resulting in a community being unable to regulate itself through informal social control, such as through social networks and shared socialization practices (Osgood & Chambers, 2000). A lack of informal control leads to an over-reliance on formal means of social control and subsequently increased crime and arrest rates (Schulenberg, 2003).

Although it is generally understood that community composition is associated with crime rates and offending patterns, very little research has directly investigated the relation between community factors and the individual risk of its residents. If, in fact, some communities are at higher risk for crime than others, it can be assumed that communities can be either high- or low-risk social contexts that can produce or attract individuals that are either high or low risk to offend. Wikstrom and Loeber (2000) tested this assumption by evaluating the association between individual risk factors, neighbourhood characteristics, and youth crime. It was found that youth in low SES areas were more likely to have risk factors of poor parental monitoring,

poor school motivation, delinquent peers, and lack of guilt than were youth in high SES areas. Similarly, Gottfredson, McNeil III, and Gottfredson (1991) evaluated the self-reported delinquency of youth in diverse social areas. The results suggested that residents of communities characterized by weakened family units and social disorganization were more likely to report individual difficulties, such as negative peer influences and less commitment to school than were youth residing in more organized communities. The studies above suggest that to better understand individual offenders, it is also necessary to understand the social pressures and constraints of the community in which they live. Further, if it is the intention of clinicians and justice professionals to reduce the risk posed by offenders, as well as to maintain risk reduction when offenders are returned to the community, it is necessary to consider both the risk of the individual, as well as risk inherent in that community.

1.3 Purpose of Study

The overall purpose of this study was to provide information that could assist service providers in assessing and treating sexually offending youth. To do this, this study first evaluated some of the psychometric properties of the J-SOAP-II and the ERASOR-II, two newly developed risk assessment instruments for sexually offending youth. Both of these instruments have been tested only minimally for concurrent validity, interrater reliability, and for internal consistency of the items. Moreover, no studies to date have confirmed the construction of the scales on these instruments. The first goal of this study, therefore, was to test the concurrent validity, interrater reliability, internal consistency, and factor structure of the J-SOAP-II and the ERASOR-II.

The second goal of this study was to provide a multilevel analysis of sexually offending youth by combining individual-level and ecological perspectives on crime. Much of the research

on sexual offending has focussed on identifying individual risk factors associated with, or predictive of, recidivism (e.g., Efta-Breitbach & Freeman, 2004; Righthand & Welch, 2004). However, as outlined above, ecological research on communities and crime has demonstrated that the composition of communities can also influence crime rates and offending patterns (e.g., McCarthy, 1991; Ouimet, 1999; Shaw & McKay, 1942; Wirth, 1938). While it is crucial to assess individual risk factors, results of ecological research suggest that youth should not be evaluated in isolation from their social and community contexts. Therefore, this multilevel analysis of crime was conducted by simultaneously evaluating the individual risk factors of sexually offending youth and the characteristics of the community in which they live. It was expected that youth in the more urbanized and heterogeneous communities would possess a greater number of risk factors than would youth from smaller and less urbanized communities. The results of this study will assist service providers in implementing prevention and intervention strategies that are appropriate for particular offenders and their respective communities.

References

- Abel, G. G., Mittleman, M., Becker, J. V., Rathner, J., & Rouleau, J. L. (1988). Predicting child molesters' response to treatment. *Annals of the New York Academy of Sciences*, 528, 223-234.
- Abel, G. G., & Rouleau, J. J. (1990). The nature and extent of sexual assault. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault* (pp. 9-21). New York: Plenum Press.
- Borum, R., Bartel, F., & Forth, A. (2002). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)*. Tampa, FL: University of South Florida, Department of Mental Health Law and Policy.
- Boxer, P., Guerra, N. G., Huesmann, L. R., & Morales, J. (2005). Proximal peer-level effects of a small-group selected prevention on aggression in elementary school children: An investigation of the peer contagion hypothesis. *Journal of Abnormal Child Psychology*, 33(3), 325-338.
- Caldwell, M. F. (2002). What we do not know about juvenile sexual reoffense risk? *Child Maltreatment*, 7(4), 291-302.
- Department of Justice Canada (n.d.). The Youth Criminal Justice Act: Summary and background. Retrieved November 24, 2007, from <http://www.justice.gc.ca/en/ps/yj/ycja/explan.html>
- Doren, D. (2006). Assessing juveniles' risk within the civil commitment context. In D. S. Prescott (Eds.), *Risk assessment of youth who have sexually abused: Theory, controversy, and emerging strategies* (pp. 100-117). Oklahoma City, OK: Wood & Barnes Publishing.

- Efta-Breitbach, J., & Freeman, K. A. (2004). Recidivism and resilience in juvenile sexual offenders: An analysis of the literature. *Journal of Child Sexual Abuse, 13*(3), 257-279.
- Epperson, D. L., Ralston, C. A., Fowers, D., DeWitt, J., Gore, K. S. (2006). Actuarial risk assessment with juveniles who offend sexually: Development of the Juvenile Sexual Offence Recidivism Risk Assessment Tool – II (JSORRAT-II). In D. S. Prescott (Ed.), *Risk assessment of youth who have sexually abused: Theory, controversy, and emerging strategies* (pp. 100–117). Oklahoma City, OK: Wood & Barnes Publishing.
- Fagan, A. A., Van Horn, M. L., Hawkins, J. D., & Arthur, M. (2007). Using community and family risk and protective factors for community-based prevention planning. *Journal of Community Psychology, 35*(4), 535-555.
- Fedoroff, J. P., & Moran, B. (1997). Myths and misconceptions about sex offenders. *Canadian Journal of Human Sexuality, 6*(4), 263-276.
- Forth, A. E., Kosson, D., & Hare, R. D. (2003). *The Psychopathy Checklist: Youth Version*. Toronto, Canada: Multi-Health Systems.
- Gifford-Smith, M., Dodge, K. A., Dishion, T. J., & McCord, J. (2005). Peer influence in children and adolescents: Crossing the bridge from developmental to intervention science. *Journal of Abnormal Child Psychology, 33*(3), 255-265.
- Gottfredson, D. C., McNeil III, R. J., & Gottfredson, G. D. (1991). Social area influences on delinquency: A multilevel analysis. *Journal of Research in Crime and Delinquency, 28*(2), 197-226.
- Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*(5), 802-809.

- Hall-Lande, J. A., Eisenberg, M. E., Christenson, S. L., & Neumark-Sztainer, D. (2007). Social isolation, psychological health, and protective factors in adolescence. *Adolescence*, 42(166), 265-286.
- Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66(2), 348-362.
- Hanson, R. K., & Thornton, D. (2000). Improving risk assessment for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, 24(1), 119-136.
- Hartnagel, T. F., & Lee, W. (1990). Urban crime in Canada. *Canadian Journal of Criminology*, 32, 591-606.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1998). Appraisal and management of risk in sexual aggressors: Implications for criminal justice policy. *Psychology, Public Policy, and Law*, 4(1), 73-115.
- Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behaviour*, 21(4), 347-359.
- Hoge, R. D., & Andrews, D. A. (1996). *The Youth Level of Service/Case Management Inventory and Manual*. Ottawa, Canada: Carleton University, Department of Psychology.
- Jacob, J. C. (2006). Male and female youth crime in Canadian communities: Assessing the applicability of social disorganization theory. *Canadian Journal of Criminology and Criminal Justice/Revue Canadienne de Criminologie et de Justice Penale*, 48(1), 31-60.
- Laub, J. H., & Lauritsen, J. L. (1994). The precursors of criminal offending across the life course. *Federal Probation*, 58(3), 51-57.

- Mager, W., Milich, R., Harris, M. J., & Howard, A. (2005). Intervention groups for adolescents with conduct problems: Is aggregation harmful or helpful? *Journal of Abnormal Child Psychology*, 33(3), 349-362.
- McCarthy, B. R. (1991). Social structure, crime, and social control: An examination of factors influencing rates and probabilities of arrest. *Journal of Criminal Justice*, 19(1), 19-30.
- Moore, T., Franey, K. C., & Geffner, R. (2004). Introduction: Assessment and treatment of youth who sexually offend: An overview. *Journal of Child Sexual Abuse*, 13(3), 1-13.
- Osgood, D. W., & Chambers, J. M. (2000). Social disorganization outside the metropolis: An analysis of rural violence. *Criminology*, 38(1), 81-115.
- Ouimet, M. (1999). Crime in Canada and in the United States: A comparative analysis. *La Revue Canadienne de Sociologie et d'Anthropologie/The Canadian Review of Sociology and Anthropology*, 36(3), 389-408.
- Ouimet, M. (2000). Aggregation bias in ecological research: How social disorganization and criminal opportunities shape the spatial distribution of juvenile delinquency in Montreal. *Canadian Journal of Criminology*, 42(2), 135-156.
- Prentky, R., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II) manual*. Washington, District of Columbia, US: US Department of Justice, Office of Justice Programs; Office of Juvenile Justice and Delinquency Prevention.
- Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence*, 10(1), 85-105.
- Righthand, S., Prentky, R., Knight, R., Carpenter, E., Hecker, J. E., & Nangle, D. (2005). Factor structure and validation of the juvenile sex offender assessment protocol (J-SOAP). *Sexual Abuse: Journal of Research and Treatment*, 17(1), 13-30.

- Righthand, S., & Welch, C. (2001). *Youths who have sexually offended: A review of the professional literature*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.
- Righthand, S., & Welch, C. (2004). Characteristics of youth who sexually offend. *Journal of Child Sexual Abuse, 13*(3), 15-32.
- Ryan, G. (1999). Treatment of sexually abusive youth: The evolving consensus. *Journal of Interpersonal Violence, 14*(4), 422-436.
- Sampson, R. J. (1997). Collective regulation of adolescent misbehaviour: Validation results from eighty Chicago neighbourhoods. *Journal of Adolescent Research, 12*(2), 227-244.
- Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology, 94*(4), 774-802.
- Schulenberg, J. L. (2003). The social context of police discretion with young offenders: An ecological analysis. *Canadian Journal of Criminology and Criminal Justice/Revue Canadienne de Criminologie et de justice Penale, 45*(2), 127-157.
- Schwartz, B. K., & Cellini, H. R. (1995). In B. K. Schwartz, & H. R. Cellini (Eds.), *Female sex offenders*. Kingston, NJ, US: Civic Research Institute.
- Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. Chicago, IL: University of Chicago Press.
- Tolman, A. O., & Rotzien, A. L. (2007). Conducting risk evaluations for future violence: Ethical practice is possible. *Professional Psychology: Research and Practice, 38*(1), 71-79.
- Wikstrom, P. H., & Loeber, R. (2000). Do disadvantaged neighbourhoods cause well-adjusted children to become adolescent delinquents? A study of male juvenile serious offending,

- individual risk and protective factors, and neighbourhood context. *Criminology*, 38(4), 1109-1142.
- Wirth, L. (1938). Urbanization as a way of life. *The American Journal of Sociology*, 44, 1-24.
- Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR): Preliminary psychometric data. *Sexual Abuse: Journal of Research and Treatment*, 16(3), 235-254.
- Worling, J. R. & Curwen, T. (2001). *Estimate of Risk of Adolescent Sexual Offence Recidivism (ERASOR). Version 2.0*. Toronto, Ontario, Canada: Safe-T Program, Thistlethorn Regional Centre for Children and Adolescents, Ontario Ministry of Community and Social Services.

2 A COMPARISON OF THE JUVENILE SEX OFFENDER ASSESSMENT PROTOCOL- II (J-SOAP-II) AND THE ESTIMATE OF RISK OF ADOLESCENT SEXUAL OFFENCE RECIDIVISM VERSION 2.0 (ERASOR-II)¹

2.1 *Introduction*

Conducting accurate and informative risk assessments for sexually offending youth has proven to be a challenging task for researchers, clinicians, and justice professionals alike. This is primarily because there are limited empirical studies that have focussed on risk assessments for this population (Prentky, Harris, Frizell, & Righthand, 2000; Rasmussen, 2004). However, to reduce further harm, important decisions must be made about how best to manage and treat sexually offending youth. Risk assessments can assist with decisions about many critical issues such as sentencing length, level of community access, and type of treatment delivery. Although risk assessments are vital in reducing the threat to public safety, there are currently no validated instruments for assessing young sexual offenders' risk for recidivism.

Standardized instruments for assessing risk have traditionally been developed for adults under the assumption that adult males commit the majority of serious sex crimes. Recent research, however, has demonstrated that in the United States, nearly half (43%) of sexual offences against children ages 6 and younger are committed by adolescents and children (National Center for Juvenile Justice, 1999 as cited in Rasmussen, 2004). In addition, as many as 50% of all adult sexual offenders have reported some sexually deviant behaviour or interest during their youth (Abel, Mittleman, Becker, Rathner, & Rouleau, 1988; Abel & Rouleau, 1990).

¹ A version of this chapter will be submitted for publication. Schoenfeld, T. M., Brown, J. E., Woodworth, M., & Gretton, H. A comparison of the Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II) and the Estimate of Risk of Adolescent Sexual Offence Recidivism Version 2.0 (ERASOR-II)

Given the prevalence of sexual crimes committed by youth and the potential for young offenders to commit sexual crimes in adulthood, researchers and clinicians have begun to devise assessment strategies that are appropriate for youth and that attempt to accurately predict an offender's potential to commit sexual crimes in the future.

2.1.1 Assessing Risk to Reoffend

A number of strategies have been used to assess an offender's risk for recidivism. However, controversy remains about which strategies most accurately predict future offending. It is generally held that clinical judgment, in which a risk rating is formulated based on anecdotal evidence, experience, and professional opinion, is less effective than a more structured and guided approach as seen in actuarial assessments (Bonta, 1996; Campbell, 2003; Hilton, Harris, & Rice, 2006; Quinsey, Harris, Rice, & Cormier, 2006). Assessments using clinical judgment do not require a systematic and standardized evaluation of risk factors and as a result, risk ratings are often based on subjective decisions (Harris & Rice, 2003). Consequently, risk ratings based on clinical judgment are often difficult to support, challenge, and compare and are often unreliable (Bonta, 1996).

Unlike unstructured clinical judgments, actuarial assessments require the systematic evaluation of a fixed number of empirically supported risk factors (Hilton et al., 2006). Actuarial tools can be advantageous because they provide scoring guidelines, cut-offs for risk levels (i.e., low, medium, or high risk for reoffending), and conclude a probabilistic estimate of risk over a fixed time period (Grann, Belfrage, & Tengstrom, 2000; Litwack, 2001; Worling, 2004). Moreover, to be considered an actuarial assessment an instrument must first undergo rigorous empirical testing and demonstrate good psychometric properties. The advantages of the actuarial assessment over less standardized approaches, such as clinical judgment, include empirical

support for the risk factors, better agreement between raters, and more accurate predictions of recidivism (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2004; Harris & Rice, 2003; Hilton et al., 2006). While there are a number of actuarial instruments for assessing the risk of adult sexual offenders (e.g., SORAG, STATIC-99), there are currently no actuarial risk assessments designed specifically to predict future sexual offending of adolescents (Doren, 2006; National Adolescent Perpetrator Network, 1993).

Although there have been risk assessment instruments that have been developed for adolescents, none of these can be considered actuarial instruments because of insufficient testing of their psychometric properties. Without empirically justified scoring guidelines, it is not possible to provide cut-off scores that categorize an offender as either low, medium, or high risk to reoffend (Epperson, Ralston, Fowers, DeWitt, & Gore, 2006; Prentky & Righthand, 2003). In addition, some question the use of statistically derived cut-off scores because these rely on a set number of risk factors based on isolated group norms while ignoring specific contextual and dynamic risk factors (Campbell, 2000 as cited in Fabian, 2006; Litwack, 2001). Moreover, there has been surprisingly little published research on adolescent sexual reoffending. As such, there is not enough information on this population to adequately develop, refine, and test actuarial instruments that assess risk for sexual reoffending (Worling, 2004).

To compensate for the lack of actuarial assessment instruments for sexually offending youth, Hanson (1998) argues that an appropriate assessment strategy is to identify and systematically review relevant risk factors for this population (Worling, 2004). In what Hanson terms “*empirically guided clinical judgment*”, identified risk factors can be used to inform decisions regarding the level of risk posed by young offenders. In contrast to actuarial assessments, empirically guided clinical assessments do not have fixed rules for tallying scores

or predetermined cut-off scores for risk levels. These instruments do, however, require assessors to base their predictions on a fixed list of factors that have been suggested by empirical evidence and professional opinion to be predictive of reoffending (Worling & Curwen, 2001). Although the final risk rating is largely based on clinical judgement, the scoring guidelines and systematic evaluation of empirically supported risk factors can result in more accurate predictions and better agreement between assessors than less standardized procedures, as seen with assessments based solely on clinical judgement (Worling, 2004).

Despite the benefits of assessing risk using a structured and empirically guided instrument, most of the instruments developed for youth are designed to assess risk of antisociality and psychopathy (Righthand et al., 2005). The Psychopathy Checklist: Youth Version (PCL:YV; Forth, Kosson, & Hare, 2003), the Youth Level of Service/Case Management Inventory and manual (YLS/CMI; Hoge & Andrews, 1996), and the Structured Assessment for Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2002) are examples of instruments designed to assess risk for general violence in youth. None of these instruments, however, are intended to specifically assess risk for reoffending sexually. To the best of our knowledge, the only risk assessment instruments designed specifically for sexually offending youth who are currently undergoing empirical testing are the Juvenile Sexual Offence Recidivism Risk Assessment Tool-II (JSORRAT-II; Epperson, et al., 2006), the Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II-II; Prentky & Righthand, 2003) and the Estimate of Risk of Adolescent Sexual Offence Recidivism Version 2.0 (ERASOR-II; Worling & Curwen, 2001). All three of these instruments have the potential to inform about the risk posed by sexually offending youth. However, the J-SOAP-II and the ERASOR-II have undergone the most empirical testing and as a result are commonly used by clinicians and justice professionals

(McGrath, Cumming, & Burchard, 2003). This study compares some of the psychometric properties of the J-SOAP-II and the ERASOR-II to provide clinicians with better statistical justification for choosing between these instruments.

2.1.2 *Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II)*

The original version of the J-SOAP (Prentky et al., 2000) was developed after extensive reviews of the literature involving both adult and adolescent studies. From these reviews, the developers constructed a 23-item risk assessment instrument originally intended to be an actuarial tool. Since its development, and in its pursuit to be an actuarial instrument, the J-SOAP-II has undergone a considerable amount of empirical testing. Accordingly, many of the original items have been modified, added, and omitted. The current version of the J-SOAP-II (Prentky & Righthand, 2003) is a 28-item evaluation divided into four scales: scale 1 - *sexual drive/preoccupation*, scale 2 - *impulsive/antisocial behaviour*, scale 3 - *intervention*, and scale 4 - *community stability/adjustment*. Scales 1 and 2 assess static, or historic, risk factors, such as index offence characteristics, previous non-sexual problem behaviours, and personal sexual and physical abuse history. Scales 3 and 4 assess dynamic factors, or alterable factors, such as treatment progress and psychosocial functioning. Items are scored from 0 to 2 indicating an apparent absence to clear presence of the risk factor. Final scoring involves summing the scores on each scale and then adding the scales together and dividing by the total possible score for the proportion of risk rating. Although the J-SOAP-II is similar to an actuarial instrument in that it provides scoring guidelines and allows for a systematic evaluation of empirically supported risk factors, it has not yet undergone sufficient empirical testing to provide cutoff scores for risk levels (e.g., low, medium, high risk; Righthand et al., 2005) and to make confident predictions of sexual recidivism (Prentky & Righthand, 2003).

Preliminary testing of the psychometric properties of the J-SOAP-II has shown that this tool has promise as an assessment instrument for sexually offending youth. For instance, in a recent study, Righthand et al. (2005) evaluated the concurrent validity of the J-SOAP-II in two steps. First, they examined the relationship between scales 1 and 2 and criminal history variables, such as the total number of nonsexual offences, total number of sexual offences, number of sexual assault victims, and the degree of aggression displayed during sexual activities. The results indicated that scale 1 was not linearly correlated with total number of offences and scale 2 was not significantly correlated with total number of sex offences. There were significant correlations between scale 1 and scale 2 and all other criminal history variables. Next, they calculated the correlation of scores on scale 2, which is comprised of items pertaining to nonsexual offending and other problem behaviours, with total scores on the YLS/CMI (Hoge & Andrews, 1996). The YLS/CMI is a valid assessment instrument for predicting risk for general violence among youth. There was a strong correlation between scale 2 and the YLS/CMI total score ($r = .81$). These tests of scale 1 and 2 provide evidence for the concurrent validity of the J-SOAP-II. Righthand et al. (2005) also tested the discriminate validity of the J-SOAP-II by comparing risk ratings for sexually offending youth who were given either community or residential placements. It was found that youth residing in the community scored lower on scales 1, 2, 3, and 4 than did youth in residential settings. In addition to adequate concurrent and discriminate validity, the J-SOAP-II has also been shown to have good to excellent interrater reliability (Prentky et al., 2000; Righthand et al., 2005).

Despite these encouraging results, the J-SOAP-II has undergone insufficient testing to be considered a valid measure of risk. Specifically, the predictive validity of the instrument has been difficult to test due to small sample sizes and low sexual recidivism rates (Prentky et al.,

2000); however, recent research suggests that the J-SOAP-II may have some utility for predicting sexual (Martinez, Flores, & Rosenfeld, 2007) and nonsexual offending (Viljoen et al., 2008). In addition, based on the results of factor analysis, the original items of version one of the J-SOAP were categorized into four scales (Righthand et al., 2005). However, confirmatory factor analysis has yet to be conducted on the revised items comprising the J-SOAP-II. Therefore, while evaluation of the J-SOAP-II has produced encouraging results, more psychometric testing is needed before this instrument can be used with confidence.

2.1.3 Estimate of Risk of Adolescent Sexual Offense Recidivism Version 2.0 (ERASOR-II)

Items were selected for inclusion in the ERASOR Version 1.2 (Worling & Curwen, 2000) using three sources of information: published studies of adolescent sexual recidivism, published studies of adult sexual recidivism, and published guidelines and checklists regarding the clinical judgment of risk and/or protective factors of sexually offending youth (Worling, 2004). The instrument was then peer reviewed, field tested, and revised to create the ERASOR-II (Worling & Curwen, 2001). The ERASOR-II is a 25-item checklist divided into five scales: scale 1 - *sexual interests, attitudes, and behaviours*, scale 2 - *historical sexual assaults*, scale 3 - *psychosocial functioning*, scale 4 - *family/environmental functioning*, and scale 5 - *treatment*. Scoring involves indicating whether the factor is *present*, *partially/possibly present*, *not present*, or *unknown*. The assessor is not to tally scale scores or final risk scores, nor are there scoring guidelines or suggestions provided. Rather, the assessor determines whether the youth is low, medium, or high risk to reoffend based on the number and/or the type of risk factors present. For instance, if a young offender has more risk factors identified as present than not present, the assessor may deem him to be high risk to reoffend sexually. However, it is also possible to deem an offender high risk to reoffend when there are only a few risk factors indicated as present. For

example, if an offender has a number of risk factors pertaining to general antisocial behaviour and there is only one instance of sexually inappropriate behaviour, the offender may be considered low or medium risk to reoffend sexually. Conversely, an offender that has only a few risk factors, but they all pertain to and suggest a pattern of inappropriate sexual behaviour, the offender may be deemed high risk to reoffend. Therefore, although the ERASOR-II allows for a systematic evaluation of empirically supported risk factors, the final risk rating remains heavily reliant on clinical judgment.

The ERASOR-II has not undergone as rigorous empirical testing as the J-SOAP-II. Doren (2006) argues that the ERASOR-II is in need of further testing because research measuring the predictive validity has produced mixed results (Morton & Bourgon, 2003) and testing of the concurrent validity is minimal (Bourgon, 2002). However, some of the psychometric testing that has been conducted has provided support for the reliability and composition of the tool. For example, Worling (2004) found the interrater reliability of the ERASOR-II to be excellent for the final risk rating ($ICC = .92$; e.g., low, medium, high) and the individual items ranged from good to excellent for all but one item (all ICCs ranged from .57 to .97). In addition, Worling reported adequate item-total correlations ($r \geq .25$; $p < .01$) for 21 of the 25 risk factors indicating that these items contributed well to the total score of the ERASOR-II (i.e., total number of risk factors scored as present). Worling also found that scores on the ERASOR-II significantly discriminated between youth who had or had not been previously sanctioned by an adult for a prior sexual offence. Although these results provided some evidence for the utility the ERASOR-II as a risk assessment, more research is needed. One area of the ERASOR-II that is lacking is statistical justification for the categorization of items into the five scales. It remains unclear exactly how these scales were derived. For instance, to the best of our

knowledge there are no published studies that report statistical rationale for combining the items into these scales; similarly, there is no empirical evidence cited for combining the items in this fashion. The lack of statistical or empirical rationale for the ERASOR-II scales makes the purpose of these scales confusing. Given the potential for clinicians to describe youth according to these scales (e.g., the offender was high risk in the area of psychosocial functioning), it is imperative that some rationale be provided for the categorization of the items in this manner.

2.1.4 Present Investigation

The purpose of this study was to further evaluate some psychometric properties of the J-SOAP-II and the ERASOR-II. First, this study evaluated consistency between raters by testing interrater reliability. While previous research on the first version of the J-SOAP indicated the instrument was generally reliable, no published studies to date have reported on the interrater reliability for the individual items on the revised second version. Further, this study sought to confirm Worling's finding that the final risk rating of the ERASOR-II was highly reliable even though it is largely based on clinical judgment. Second, item-total correlations were calculated and factor analyses were conducted to provide support for the inclusion of individual items and to provide statistical justification for the categorization of the items on each of the instruments. This was especially necessary given that these tests have not yet been conducted on the revised version of the J-SOAP-II and there is currently no known statistical rationale for the scales of the ERASOR-II. Lastly, the concurrent validity of the ERASOR-II was assessed by comparing its total score to the total score of the more validated J-SOAP-II instrument. If both of these instruments are intended to assess risk, they should conclude a similar level of risk for each offender.

2.2 Method

2.2.1 Participants

Risk assessments were conducted on 84 male adolescents who were assessed and treated between 1997 and 2006 at Youth Forensic Psychiatric Services (YFPS) in British Columbia. The adolescents ranged in age from 11 to 20 years ($M = 16.12$, $SD = 1.79$). All had been convicted of a sexual offence and/or acknowledged committing a sexual offence. The majority of the sample was Caucasian (69.9%), followed by Aboriginal (17.8%), Asian (5.5%), Indo-Canadian (2.7%), and 4.1% were coded as Mixed or Other. Only youth with closed files who had completed or had been discharged from treatment were included in the study.

2.2.2 Procedure

Trained researchers scored the J-SOAP-II and ERASOR-II from extensive file information on youth treated in one of four YFPS clinics: Burnaby ($n = 30$), Kelowna ($n = 22$), Kamloops ($n = 28$), and Penticton ($n = 4$). Files generally contained psychological assessments, school records, police records, treatment notes, and progress and discharge reports. Given the resource and regional limitations of this study, there was only one researcher who could do initial coding on the files from Kelowna, Kamloops, and Penticton. Conversely, there were multiple coders in Burnaby. Therefore, one researcher scored first the ERASOR-II followed by the J-SOAP-II on the Kelowna, Kamloops, and Penticton files. The instruments were scored in this order to reduce the potential for a biased rating on the ERASOR-II *final risk rating*, which is heavily reliant on clinical judgement. Researchers in Burnaby, however, had different raters score the J-SOAP-II and the ERASOR-II on each file. Reliability coding was done on a random

sample of files ($n = 18$) from the Kelowna, Kamloops, Penticton, and Burnaby clinics by a variety of trained raters employed at the Burnaby clinic.

2.3 Results

2.3.1 Interrater Reliability

Interrater agreement on the J-SOAP-II and the ERASOR-II was calculated for a random sample of cases (20%). Intraclass correlation coefficients (ICC) were calculated for the J-SOAP-II *proportion of risk*, ERASOR-II *total score*, and the ERASOR-II *final risk rating*. Considering that not all of the cases were scored by the same two raters, a two-way random effects model was selected. In addition, the more conservative test of absolute agreement was used because this test measures the potential differences between each rater on each case rather than examining patterns of scores as calculated by the consistency test (McGraw & Wong, 1996). The more conservative test of single measure ICCs are reported because further tests were calculated using single ratings rather than an average rating as calculated by the average measure. The ICC tables, however, display both the single measure and the average measure for each of the items on the J-SOAP-II and the ERASOR-II.

All three variables demonstrated good to excellent levels of interrater agreement (for interpretation guidelines on the strength of agreement for ICCs, see Cicchetti et al., 2006). The J-SOAP-II *total score* had the strongest agreement between raters ($n = 16$, $ICC = .943$). The ERASOR-II *total score* (i.e., total number of risk factors scored as present or partially/possibly present) also demonstrated excellent interrater agreement ($n = 18$, $ICC = .790$). The score that was most reliant on clinical judgment (i.e., ERASOR-II *final risk rating*), however, was found to be the least reliable measure ($n = 17$, $ICC = .636$). Despite the lower agreement between raters on this measure, a Spearman's correlation coefficient for ranked data indicated that the

ERASOR-II *total score* was highly correlated with the ERASOR-II *final risk rating* ($r_s = .751, p < .001$) suggesting that the number of risk factors identified as present or partially present is positively associated with the *final risk rating* determined by the rater.

To further evaluate the interrater reliability of the instruments, all of the risk factors were examined individually (see Table 1 and Table 2). The items of the J-SOAP-II had good to excellent reliability with individual ICCs ranging from .25 to 1.0. While the ERASOR-II had a number of items that were highly reliable, the range of individual ICCs was greater, ranging from -.16 to 1.0. These results suggest that the J-SOAP-II was generally a more reliable measure for this sample.

Table 2-1 Interrater Agreement on Individual J-SOAP-II Items

J-SOAPII risk factors (<i>n</i> = 16)	ICC (single rating)	ICC (average rating)
<i>Sexual drive/preoccupation</i>		
1. Prior legally charged sex offences	~	~
2. Number of sexual abuse victims	.88	.93
3. Male child victim	.91	.95
4. Duration of sex offence history	.93	.96
5. Degree of planning	.41	.58
6. Sexualized aggression	.73	.84
7. Sexual drive and preoccupation	.76	.87
8. Sexual victimization history	1.00	1.00
<i>Impulsive/antisocial behaviour</i>		
9. Caregiver consistency	.87	.93
10. Pervasive anger	.73	.85
11. School behaviour problems	.87	.93
12. History of conduct disorder	.91	.95
13. Juvenile antisocial behaviour	.67	.80
14. Ever charged/arrested before age 16	.74	.85
15. Multiple types of offences	.35	.52
16. History of physical assault and/or exposure	.84	.91
<i>Intervention</i>		
17. Accepting responsibility	.53	.69
18. Internal motivation for change	.50	.67
19. Understands risk factors	.78	.87
20. Empathy	.33	.49
21. Remorse and guilt	.34	.50
22. Cognitive distortions	.60	.75
23. Quality of peer relationships	.57	.73
<i>Community stability/adjustment</i>		
24. Management of sexual urges and desire	.25	.40
25. Management of anger	.52	.68
26. Stability of current living situation	.70	.82
27. Stability in school	.79	.88
28. Evidence of positive support systems	.81	.89
Total Proportion of Risk	.94	.97

~ ICCs could not be calculated because the item(s) had zero variance.

Table 2-2 Interrater Agreement on Individual ERASOR-II Items

ERASOR-II risk factors (<i>n</i> = 18)	ICC (single rating)	ICC (average rating)
<i>Sexual interests, attitudes, and behaviours</i>		
1. Deviant sexual interests	.50	.67
2. Obsessive sexual interests/Preoccupation with sexual thoughts	.34	.51
3. Attitudes supportive of sexual offending	~	~
4. Unwillingness to alter deviant sexual interests/attitudes	~	~
<i>Historical sexual assaults</i>		
5. Ever sexually assaulted 2 or more victims	.87	.93
6. Ever sexually assaulted same victim 2 or more times	.97	.98
7. Prior adult sanctions for sexual assault(s)	.82	.90
8. Threats of, or use of, violence/weapons during sexual offence	.57	.73
9. Ever sexually assaulted a child	.83	.91
10. Ever sexually assaulted a stranger	.95	.98
11. Indiscriminate choice of victims	.80	.89
12. Ever sexually assaulted a male victim	1.00	1.00
13. Diverse sexual assault behaviours	.57	.72
<i>Psychosocial functioning</i>		
14. Antisocial interpersonal orientation	.16	.27
15. Lack of intimate peer relationships / Social isolation	.49	.66
16. Negative peer associations and influences	.56	.72
17. Interpersonal aggression	.59	.74
18. Recent escalation in anger or negative affect	-.16	-.37
19. Poor self-regulation of affect and behaviour	.42	.59
<i>Family/Environment functioning</i>		
20. High-stress family environment	.53	.69
21. Problematic parent-offender relationships/Parental rejection	.51	.68
22. Parent(s) not supporting sexual-offence-specific assessment/treatment	.65	.79
23. Environment supporting opportunities to reoffend sexually	.59	.74
<i>Treatment</i>		
24. No development or practice of realistic prevention plans/strategies	.77	.87
25. Incomplete sexual-offence-specific treatment	.97	.99
Overall Risk Rating	.64	.79

~ ICCs could not be calculated because the item(s) had zero variance .

2.3.2 Structure of the J-SOAP-II

Internal consistency. The first step to establishing the structure of the J-SOAP-II was to look at internal consistency using two measures: Cronbach's alpha and item-total correlations (ITC). Refer to Table 3 for computed values. From the results reported in the table it can be seen that overall the scales were relatively consistent. For example, alpha values ranged between .69 and .90. Generally, it is accepted that values greater than .7 are considered highly reliable. The

individual items and corresponding ITCs are also highly consistent. Many values were greater than .40, which Leech, Morgan, and Barrett (2005) argue is characteristic of a good scale component. However, three items in the *sexual drive/preoccupation scale* (prior legally charged sex offenses, sexualized aggression, and sexual victimization history) had low ITCs (i.e., < .3) raising concern about their inclusion. This scale also had the lowest overall alpha ($\alpha = .69$).

Table 2-3 Internal Consistency of J-SOAP-II Scales

Items	Corrected ITC	Alpha if item deleted	Subscale alpha
<i>Sexual drive/preoccupation</i>			.69
1. Prior legally charged sex offenses	.16	.70	
2. Number of sexual abuse victims	.70**	.58	
3. Male child victim	.30*	.69	
4. Duration of sex offence history	.54**	.63	
5. Degree of planning	.50**	.65	
6. Sexualized aggression	.19	.70	
7. Sexual drive and preoccupation	.48**	.64	
8. Sexual victimization history	.19	.70	
<i>Impulsive/antisocial behaviour</i>			.90
9. Caregiver consistency	.47**	.90	
10. Pervasive anger	.81**	.87	
11. School behaviour problems	.71**	.88	
12. History of conduct disorder	.80**	.87	
13. Juvenile antisocial behaviour	.77**	.87	
14. Ever charged/arrested before age 16	.62**	.89	
15. Multiple types of offences	.66**	.89	
16. History of physical assault and/or exposure	.62**	.89	
<i>Intervention</i>			.86
17. Accepting responsibility	.61**	.84	
18. Internal motivation for change	.71**	.83	
19. Understands risk factors	.64**	.84	
20. Empathy	.62**	.84	
21. Remorse and guilt	.76**	.82	
22. Cognitive distortions	.52**	.86	
23. Quality of peer relationships	.55**	.85	
<i>Community stability/adjustment</i>			.76
24. Management of sexual urges and desire	.35*	.77	
25. Management of anger	.61**	.68	
26. Stability of current living situation	.57**	.70	
27. Stability in school	.59**	.69	
28. Evidence of positive support systems	.51**	.72	

* $p < .05$. ** $p < .01$.

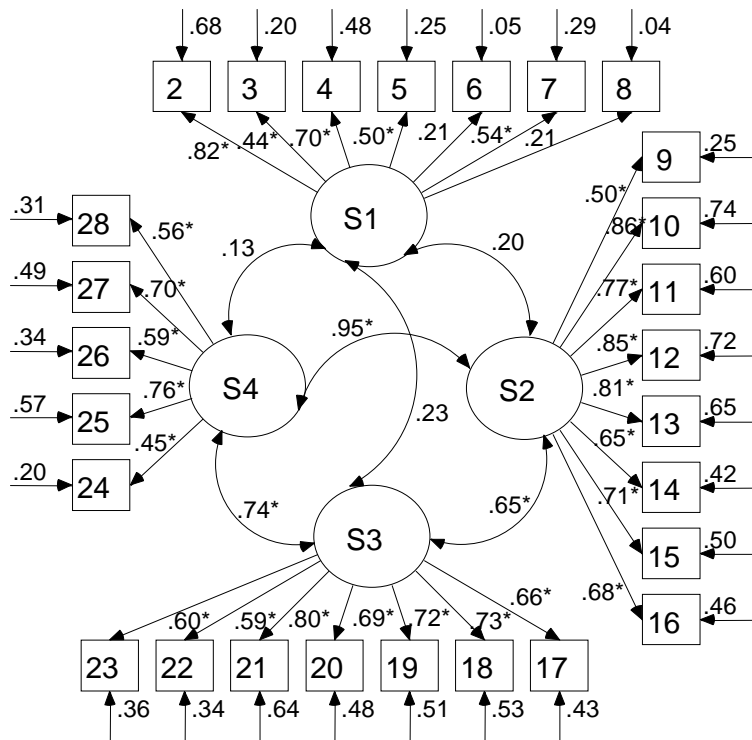
Confirmatory factor analysis (CFA). CFA was conducted using structural equation modeling (SEM) with Amos 4.0 software. Prior to conducting the CFA, each variable was assessed for univariate and multivariate outliers and normality. One case exceeded a z score of

3.29 ($p < .001$) on the variable *prior legally charged sex offenses*, indicating that it was an extreme case. Only one youth had been previously charged with a sexual offence (*prior legally charged sex offenses*). Due to a lack of variance in scores on the variable *prior legally charged sex offenses*, a determinant of zero would have been obtained resulting in a not-positive, definite covariance matrix, which would have made the solution unattainable. Consequently, it was necessary to remove the outlier variable from the model before further analyses. In addition to univariate outliers, multivariate outliers were examined. According to Mahalanobis distance there were no multivariate outliers, $\chi^2(N = 28) = 56.89, p = .001$. However, calculations computed through Amos showed many of the variables to deviate from normality. Given that the distributions are not normal and the data are ordinal with few categories, transformations were not a viable option. In addition, it is inappropriate to conduct regular *SEM* procedures on nonnormal data using more standardized estimations, such as maximum likelihood (Bollen & Stine, 1992). Consequently, the more appropriate procedure of bootstrapping was used instead. A requirement of the bootstrapping procedure is that there are no missing data; therefore, all missing data were replaced with the mean. The conservative procedure of missing data substitution (Tabachnick & Fidell, 2001) was chosen because it is not desirable to omit cases when the sample size is small. In addition, the mean of the distribution does not change and it allows for final scores to be calculated on the J-SOAP-II and ERASOR-II even when some of the items are omitted because of a lack of information. Although the bootstrapping procedure is sensitive to missing data, the advantage of this procedure is that it does not require assumptions about the shape of the population, just that the sample distribution would be similar to the population distribution (Preacher, Rucker, & Hayes, 2007). Two thousand bootstrap samples were estimated to create a sufficient pseudo population with which to compare the model. Model

fit was assessed using the Bollen-Stine estimation procedure because this method is recommended as a conservative test for small sample sizes and nonnormal data (Bollen & Stine, 1992).

When the model that was created to simulate the scales of J-SOAP-II was run (see Figure 1), none of the samples that were generated were discarded due to singularity or other statistical reasons. Results from the analyses yielded a Bollen-Stine value $p = .035$, which failed to indicate that the model was a good fit. Further, the Bias Corrected (BC) confidence level was reported, which corrects for skewness in bootstrap samples (Arbuckle & Wothke, 1999). Figure 1 displays the standardized and significant parameter estimates.

Figure 2-1 Confirmatory Factor Analysis of J-SOAP-II Scales



All except two variables (*sexual aggression* and *sexual victimization history*) were significant predictors ($p < .001$) of the latent constructs (scale 1, 2, 3, and 4). Consistent with the

ITC calculations, the variables *sexualized aggression* and *sexual victimization history* were not good predictors of scale 1, only accounting for 5% and 4% of the variance respectively.

Examination of the correlations between the latent constructs provided further explanation for the lack of model fit. All scales were significantly correlated with each other except for scale 1. According to this model, scale 1 (*sexual drive/preoccupation*) is not correlated with scale 2 (*impulsive/antisocial behaviour*), scale 3 (*intervention*), or scale 4 (*community stability/adjustment*). Therefore, while the majority of the risk items load significantly on their respective scale, the items on scale 1 that are not a good fit may be contributing to the overall lack of model fit.

2.3.3 Structure of the ERASOR-II

Internal consistency. In the ERASOR-II manual (Worling & Curwen, 2001) and a subsequent article (Worling, 2004), the developers present the instrument as a single construct measure. Although the items are categorized, the instrument is not intended to be evaluated by its scales, but rather all of the items compile to represent an overall risk to reoffend. Therefore, the first step in examining the structure of the ERASOR-II was to evaluate Cronbach's alpha and ITCs for the overall risk items. All items were simultaneously entered into the analysis, resulting in a good overall Cronbach, $\alpha = .71$. Examination of the individual items, however, indicated that few are consistent. See Table 4 for individual ITCs and the percentage of endorsed risk factors. For instance, only four of the items (*threats of or use of violence/weapons during sex offence*, *antisocial interpersonal orientation*, *interpersonal aggression*, and *poor self-regulation or affect of behaviour*) had ITCs greater than .40, which is an indication of a good instrument (Leech et al., 2005). ITCs of all other items ranged from only .03 to .38, with the lowest two items being *diverse sexual assault behaviours* (ITC = .03) and *ever sexually assaulted a child* (ITC = .07).

Note, however, that the percentages of endorsed items indicated that the distributions on most of the variables are skewed. That is, this sample generally had many fewer risk factors identified as present than not present. Given that these ITCs are much lower than those reported by Worling (2004), it must be considered that the nonnormality of this small sample could be reducing the consistency of the items of the overall instrument.

Table 2-4 Internal Consistency and Endorsement Patterns of Individual ERASOR-II Items

ERASOR-II risk factors (n = 84) Cronbach's $\alpha = .71$	ITC	% Present	% Possibly or Partially Present	% Not Present	% Unknown
<i>Sexual interests, attitudes, and behaviours</i>					
1. Deviant sexual interests	.13	10.7	8.3	72.6	8.3
2. Obsessive sexual interests/Preoccupation with sexual thoughts	.29*	7.1	10.7	70.2	11.9
3. Attitudes supportive of sexual offending	.15	0.0	14.3	82.1	3.6
4. Unwillingness to alter deviant sexual interests/attitudes	~	6.0	7.1	86.3	4.8
<i>Historical sexual assaults</i>					
5. Ever sexually assaulted 2 or more victims	.37**	42.9	9.5	45.2	2.4
6. Ever sexually assaulted same victim 2 or more times	.12	57.1	6.0	33.3	3.6
7. Prior adult sanctions for sexual assault(s)	.31*	23.8	8.3	67.9	0.0
8. Threats of, or use of, violence/weapons during sexual offence	.40**	17.9	7.1	73.8	1.2
9. Ever sexually assaulted a child	.07	81.0	0.0	17.9	1.2
10. Ever sexually assaulted a stranger	.11	7.1	2.4	89.3	1.2
11. Indiscriminate choice of victims	.29*	23.8	7.1	67.9	1.2
12. Ever sexually assaulted a male victim	.32*	34.5	4.8	59.5	1.2
13. Diverse sexual assault behaviours	.03	42.9	13.1	42.9	1.2
<i>Psychosocial functioning</i>					
14. Antisocial interpersonal orientation	.45**	23.8	19.0	54.8	2.4
15. Lack of intimate peer relationships / Social isolation	.12	27.4	22.6	50.0	0.0
16. Negative peer associations and influences	.26**	19.0	21.4	54.8	4.8
17. Interpersonal aggression	.51**	20.2	14.3	63.1	2.4
18. Recent escalation in anger or negative affect	.37**	6.0	9.5	81.0	3.6
19. Poor self-regulation of affect and behaviour	.47**	17.9	32.1	45.2	4.8
<i>Family/Environment functioning</i>					
20. High-stress family environment	.16	35.7	20.2	41.7	2.4
21. Problematic parent-offender relationships/Parental rejection	.27*	29.8	19.0	51.2	0.0
22. Parent(s) not supporting sexual-offence-specific assessment/treatment	.20	21.4	17.9	54.8	6.0
23. Environment supporting opportunities to reoffend sexually	.20	25.0	26.2	33.3	15.5
<i>Treatment</i>					
24. No development or practice of realistic prevention plans/strategies	.38**	21.4	7.1	70.2	1.2
25. Incomplete sexual-offence-specific treatment	.31*	20.2	4.8	73.8	1.2

* $p < .05$. ** $p < .01$.

~ Item was not included in analysis because it had zero variance resulting in the covariance matrix being close to zero and indeterminable.

In contrast with our results, Worling (2004) examined the structure of the ERASOR-II and concluded that the instrument had good internal consistency because the ITCs were adequate for 21 out of 25 of the risk factors ($r \geq .25$; $p < .01$). However, it is not clear whether these items were examined individually or according to the predetermined categories. Considering there were discrepancies between Worling's results and our own, we also calculated the item analyses according to the scales by entering each scale separately. Table 5 displays the ITCs and Cronbach's alphas for each scale. Separate analysis for each category yielded better and more consistent results than did the previous analyses on individual items. Despite these improved results, less than half (i.e., 12 out of 25 items) had ITCs greater than .40. However, in contrast to previous analyses, these results were more similar to Worling's in that 21 out of 25 of the items had ITC's of .25 or greater ($p < .05$), but only 14 items had ITCs greater than .32 ($p < .01$). In addition, some of the items remained very low. For instance, *ever sexually assaulted a stranger* and *lack of intimate peer relationships/social isolation* had ITCs of only -.12 and .09 respectively. While most of the scales had some low ITCs, the overall scale alpha remained high for all but scale 4 (scales 1, 2, 3, and 5 had alpha's ranging from .61 to .93). Scale 4 (*family/environment functioning*) had a low scale alpha of .47 and the ITCs for the individual items on this scale ranged from only .19 to .36 suggesting inconsistencies among the items. Although ITC calculations for the ERASOR-II improved when the individual items were analyzed according to their respective categories, this improvement was not enough to suggest a reliable and consistent instrument overall.

Table 2-5 Internal Consistency of ERASOR-II Scales

Items	Corrected ITC	Alpha if item deleted	Subscale alpha
<i>Sexual interests, attitudes, and behaviours</i>			.61
1. Deviant sexual interests	.28*	.67	
2. Obsessive sexual interests/Preoccupation with sexual thoughts	.51**	.44	
3. Attitudes supportive of sexual offending	.58**	.50	
4. Unwillingness to alter deviant sexual interests/attitudes	.36**	.55	
<i>Historical sexual assaults</i>			.69
5. Ever sexually assaulted 2 or more victims	.62**	.60	
6. Ever sexually assaulted same victim 2 or more times	.32*	.67	
7. Prior adult sanctions for sexual assault(s)	.55**	.62	
8. Threats of, or use of, violence/weapons during sexual offence	.13	.70	
9. Ever sexually assaulted a child	.22*	.69	
10. Ever sexually assaulted a stranger	-.12	.72	
11. Indiscriminate choice of victims	.59**	.62	
12. Ever sexually assaulted a male victim	.40**	.65	
13. Diverse sexual assault behaviours	.41**	.65	
<i>Psychosocial functioning</i>			.71
14. Antisocial interpersonal orientation	.68**	.60	
15. Lack of intimate peer relationships / Social isolation	.09	.79	
16. Negative peer associations and influences	.32*	.71	
17. Interpersonal aggression	.70**	.59	
18. Recent escalation in anger or negative affect	.34*	.70	
19. Poor self-regulation of affect and behaviour	.67**	.60	
<i>Family/Environment functioning</i>			.47
20. High-stress family environment	.29*	.38	
21. Problematic parent-offender relationships/Parental rejection	.36**	.31	
22. Parent(s) not supporting sexual-offence-specific assessment/treatment	.25*	.42	
23. Environment supporting opportunities to reoffend sexually	.19	.47	
<i>Treatment</i>			.93
24. No development or practice of realistic prevention plans/strategies	.87**	~	
25. Incomplete sexual-offence-specific treatment	.87**	~	

* $p < .05$. ** $p < .01$.

~ Not computed because there were only two items in the scale.

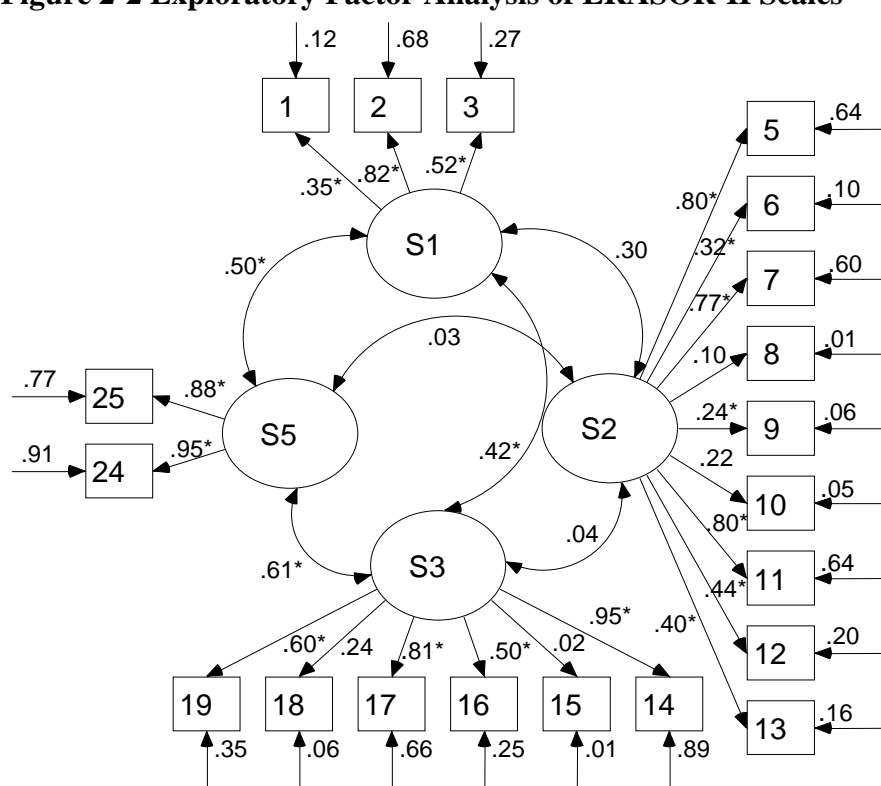
Exploratory factor analysis. Given that the internal consistency of the ERASOR-II was improved when the items were considered as categories rather than individual items, the scales of the instrument were further explored. Specifically, an exploratory factor analysis using Amos 4.0 software was conducted to assess whether the ERASOR-II items loaded onto the latent constructs of their respective categories. Prior to analysis, variables were assessed for outliers and normality. Two variables (*unwillingness to alter deviant sexual interests/attitudes* and *ever sexually assaulted a stranger*) had cases that were outliers because they exceeded a z score of 3.29, $p = .001$. Outlier cases were those individuals who had one or both of the risk factors

scored as present because it was rare to do so. Because these extreme scores reflect the lack of deviancy in this sample overall and they barely exceed the z score of 3.29, they remained in the analysis. In addition to univariate outliers, multivariate outliers were examined. According to Mahalanobis distance, there were no multivariate outliers, $\chi^2(N = 25) = 52.62, p = .001$.

Calculations computed through Amos showed many of the variables deviated from normality. Bootstrapping was used to adjust for nonnormality and for the ordinal nature of the data. To meet the requirements of the bootstrapping procedure, missing data were replaced with the mean. Two thousand bootstrap samples were estimated to create a sufficient pseudo population with which to compare the model and the model fit was assessed using the Bollen-stine procedure (Bollen & Stine, 1992).

A model was created to test the five scales of the ERASOR-II instrument. Initial attempts to run the model were unsuccessful because of negative variance that resulted in standardized estimates being unachievable. As a result, the model was re-evaluated to obtain a better fit. Examination of the ITCs indicated two fundamental problems that could be causing the model to be unsolvable. First, an ITC for item 4 (*unwillingness to alter deviant sexual interests/attitudes*) could not be calculated because it had zero variance resulting in the covariance matrix being close to zero and indeterminable. Second, scale 4 (*family/environmental functioning*) had a low scale alpha indicating that these items did not comprise a good scale. Therefore, the model was redrawn excluding scale 4 and item 4 and it was rerun. Figure 2 illustrates the final model of the ERASOR-II.

Figure 2-2 Exploratory Factor Analysis of ERASOR-II Scales



Examination of the bootstrap summary indicated that 1 bootstrap sample was unused because of a singular covariance matrix and 16 were unused for other statistical reasons. The Bollen-stine calculation indicated that even with the problem items omitted the model was not a good fit ($p = .021$). Refer to Figure 2 for the standardized and significant parameter estimates.

According to the Bias Corrected confidence intervals, all but four variables were significant predictors ($p < .05$) of the latent constructs (scale 1, 2, 3, and 5). Consistent with the ITC calculations, items 8, 10, and 15 did not load on their respective factors. In addition, item 18 was not found to be predictive of its respective scale. All of these factors accounted for only .02 to .24% of the variance of their respective scales. Examination of the correlations between the latent constructs provided further explanation for the lack of model fit. All scales were significantly correlated with each other ($p < .05$) except for scale 2. According to this model, scale 2 (*historical sexual assaults*) is not correlated with scale 1 (*sexual interest, attitudes, and*

behaviours), scale 3 (*psychosocial functioning*), and scale 5 (*treatment*). Therefore, according to these results, the overall lack of model fit was a result of both individual factors and the *historical sexual assault scale*.

2.3.4 Concurrent Validity

To evaluate the concurrent validity of the ERASOR-II, a Spearman's rho was calculated for the ERASOR-II *final risk rating* and the J-SOAP-II *proportion of risk*. The ERASOR-II *final risk rating* was strongly correlated with the J-SOAP-II *proportion of risk* ($r_s = .739, p < .001$). However, considering that the ERASOR-II *final risk rating* was the least reliable measure, we also tested the association between the ERASOR-II *total score* and the J-SOAP-II *proportion of risk*. There was a stronger correlation between the ERASOR-II *total score* and the J-SOAP-II *proportion of risk* ($r = .810, p < .001$) than there was between the ERASOR-II *final risk rating* and the J-SOAP-II *proportion of risk*. These results suggest that when compared to the J-SOAP-II, the ERASOR-II demonstrates better concurrent validity when the least subjective measure of ERASOR-II *total score* is used.

2.4 Discussion

The purpose of this study was to test some of the psychometric properties of both the J-SOAP-II and the ERASOR-II. Specifically, interrater reliability, construct validity, and concurrent validity were examined. Both the J-SOAP-II and the ERASOR-II were shown to have good to excellent agreement between raters. Of the total scores on each risk assessment, the J-SOAP-II *final risk rating* had the strongest agreement between raters. The ERASOR-II *total score* (i.e., number of risk factors scored as present or partially/possibly present) also demonstrated excellent interrater reliability. Not surprisingly, however, the score that was most

reliant on clinical judgment (ERASOR-II *final risk rating*) was found to be the least reliable measure among raters. A correlation between the ERASOR-II *total score* and the ERASOR-II *final risk rating* indicated that the number of risk factors identified as present or partially/possibly present was correlated with the final rating determined by clinical judgment. The lack of agreement between raters on the ERASOR-II *final risk rating* and the strong association between this score and the number of risk factors may indicate that simply tallying the number of risk factors might be a more reliable method of assessing risk using the ERASOR-II. However, with our sample, the risk rating determined by the J-SOAP-II was more reliable than either the ERASOR-II *final risk rating* or the ERASOR-II *total score*.

To further investigate the interrater reliability, however, inter-item correlation coefficients (ICCs) were calculated for each of the individual items that comprise the J-SOAP-II and the ERASOR-II. Given that the J-SOAP-II was shown to be the most reliable overall, it was not surprising that the individual items of this scale were also highly reliable. The ICCs for individual items on this instrument ranged from .25 to 1.00. Although many of the individual items of the ERASOR-II were shown to be highly reliable, there was more variability among the ICCs (-.16 to 1.0) indicating that the individual items were not as consistently reliable as those of the J-SOAP-II.

Although scoring on the J-SOAP-II was highly reliable between raters, the ICCs were not as strong as correlations reported by Prentky et al. (2000). These authors assessed the original version of the J-SOAP and indicated that Pearson correlations for individual items ranged from .59 to .91. They concluded that with the exception of the item *caregiver instability* ($r = .59$), all other items had good to excellent reliability. For the revised version of the J-SOAP-II the item of *caregiver instability* was modified to improve consistency between raters. This study found that

the modified item (*caregiver consistency*) was much improved ($ICC = .87$). By contrast, some of the items that had been modified from the original version of the J-SOAP-II showed a reduction in agreement. For example, the original item of *evidence of empathy, remorse, and guilt* was separated into two items for the second version of the J-SOAP-II. The result, unfortunately, was that these items had a reduction in reliability from an overall Pearson correlation of .84 to low ICCs of only .33 and .34, respectively. However, some of the new items of the J-SOAP-II that were not included in the original were found to be highly reliable between raters (ICCs = .41 to 1.0). Therefore, discrepancies between the results of this study and Prentky et al. could be due to the subsequent modifying, omitting, and adding of items from version one to version two of the J-SOAP.

In addition to the J-SOAP-II, previous research on the ERASOR-II has found more reliable scoring between raters than what was found in our study. Worling (2004) collected risk ratings from 28 clinicians who had conducted comprehensive clinical assessments on 136 male adolescents. In contrast to the findings of this study, Worling reported strong single rating ICCs ranging from .40 to .92. Moreover, the ERASOR-II *final risk rating* of low, medium, high was found to be highly reliable ($ICC = .85$), which is inconsistent with our ICC of only .67 for this measure.

There are a number of possible explanations for the discrepancies between our results and those reported by Worling (2004). First, Worling conducted extensive interviews to determine the levels of risk for each individual, whereas we conducted our assessments using comprehensive file information. In Worling's study, pairs of assessors actually divided their assessments, such that one clinician focused on the offender and the other focused on the family; Worling suggested that having access to exactly the same information, as was the case with our

file review, might actually improve agreement between raters. The results of our study suggest that this is probably not the case because our ratings were based on file review and were not as reliable as Worling's. However, had we based ratings on both interviews and file review we may have increased agreement between raters because it is advisable to conduct risk assessments using multiple sources (Worling & Curwen, 2001). Second, since Worling had more than one rater for each offender, clinicians first completed their ratings independently and subsequently met to discuss their assessments to produce a combined overall risk rating. Although the raters of this study were trained by the same organization, the risk assessments were conducted individually. Third, considering that the assessments in Worling's study were based on clinical interviews rather than file information, assessors may have had the ability to ask specific questions to obtain the necessary information for scoring the ERASOR-II. This study, however, was limited to file information. Consequently, when information was lacking or ambiguous, raters did not have the ability to request clarification or elaboration. As such, the ratings in this study may have been disadvantaged by subjectivity. Therefore, the present study suggests that the ERASOR-II may be a more reliable measure when raters are able to interview offenders, and discuss and agree upon their ratings.

In addition to evaluating the degree of agreement between raters on the J-SOAP-II and the ERASOR-II, the construct validity of both instruments was evaluated using two main measures: item-total correlations and factor analyses. These two tests were conducted on each of the instruments for different reasons. Based on preliminary psychometric testing, the J-SOAP-II has been revised considerably from its original form. However, the J-SOAP-II has yet to undergo the necessary testing of this new construction. For instance, item-total correlations of the modified items comprising the scales have yet to be evaluated. In addition, although the

developers conducted an exploratory factor analysis to derive the scales, confirmatory factor analysis has not been conducted on these modified scales. The ERASOR-II, on the other hand, has not undergone revisions since it was last examined (Worling, 2004). However, this study sought to confirm the internal consistency of the items when scored on file information. In addition, exploratory factor analysis of the ERASOR-II scales was conducted to determine whether the scales proposed by the developer were statistically justified. Considering that the developers have argued that final risk ratings should not be derived from tallying the number of risk items, it is not entirely clear what the purpose of the scales are. For instance, the developers neglected to explain exactly how these scales were derived and how they should be used in the assessment. It is unclear whether these scales are to be used for descriptive purposes or are simply an arbitrary categorization of the items.

The results of this study were consistent with previous research in that calculations of item-total correlations on the J-SOAP-II items indicated that the instrument was relatively internally consistent. Similar to the findings of Prenky et al. (2000), ITCs for individual items and scale alphas were high (significant ITCs ranged from .30 to .81 and scale alpha's ranged from .69 to .90). Two of the items with low ITCs, however, were newly added items (*sexualized aggression* and *sexual victimization history*) and the other was modified (*prior legally charged sex offenses*) raising concern about their inclusion. Despite these low scores, the J-SOAP-II was shown to be relatively consistent overall.

The confirmatory factor analysis of the J-SOAP-II scales yielded similar results. The results indicated that the model was not a good fit overall, failing to support the scales of the instrument. However, further evaluation of the items and the scales indicated that the same two newly added items (*sexualized aggression* and *sexual victimization history*) were the only non

significant predictors of the latent construct *sexual drive/preoccupation scale*. Further, this scale was the only scale that was not correlated with the others. Therefore, the *sexual drive/preoccupation scale* seems to be reducing the strength of the overall fit of this instrument.

Based on this sample of youth, the ERASOR-II seemed to be less internally consistent than the J-SOAP-II. Inconsistent with previous research, many of the items in the ERASOR-II were shown to have poor internal consistency. For example, Worling (2004) reported that the item-total correlations were acceptable for most of the risk factors, whereas we found that only four of the items had ITCs greater than .40. It should be noted, however, that even if we used the less conservative cutoff of $p < .01$ as used by Worling, we still only would have 12 items that reached significance compared to Worling's 21 out of 25. There are two possible explanations for this discrepancy between our results and Worling's. First, our sample seemed to be less deviant. Few of our offenders were scored as high risk offenders, whereas Worling's sample was more normally distributed with approximately equal numbers of youth scoring high and low on these risk factors. This study might suggest that the ERASOR-II is not a reliable instrument for generally low risk offenders. Second, it is unclear whether Worling assessed the ITCs for each category separately, or for all items simultaneously. If it was the case that items were evaluated as scales not as individual items, then our assessment of the scales should have produced more consistent results. Although ITCs did improve when items were assessed as categories, the improvement was small; only 12 out of 25 items had ITCs greater than .40. In summary, the ERASOR-II did not demonstrate good internal consistency, regardless of whether the instrument was assessed by its scales or individual items.

An exploratory factor analysis was conducted to further evaluate the relevance of the ERASOR-II scales. The model created to test the scales would not run successfully without

removal of the item *unwillingness to alter deviant sexual interests/attitudes* and the scale with the low Cronbach's alpha of .49 (*family/environment functioning*). Once removed, the model ran, but was still not a good fit. Specifically, four variables were inadequate predictors of their respective scales. In addition, all scales were significantly correlated with each other except for the *historical sexual assaults scale*. This scale was uncorrelated with all three other scales of the instrument. The lack of internal consistency and model fit suggest, therefore, that there is little statistical justification for the categorization of the individual items of the ERASOR-II into scales. This finding has important implications for clinicians who use the ERASOR-II to conduct risk assessments on youth. As mentioned above, it is not clear in the manual nor in subsequent research exactly how the scales should be used. Specifically, there is no information about how these scales were derived and whether it is appropriate to describe offenders according to these scales. For example, given that clinicians conducting assessments with the ERASOR-II are to include a description of the offender's risk, a clinician may be tempted to describe an offender as being generally high risk on one or more scales, such as in the area of psychosocial functioning. Our results suggest that this is not an appropriate use of the ERASOR-II. The lack of internal consistency and model fit indicate that it is incorrect to assume that all of the items on a scale are measuring the same underlying concept, such as psychosocial functioning.

It is also important to note that for both the J-SOAP-II and the ERASOR-II, the scale that described sexual offending behaviours and history was not associated with any other scales for either of the instruments. That is, it seems as though risk factors pertaining to sexual offending are unrelated to other risk factors about general offending and other problematic behaviours. This suggests that while the risk assessments are intended to evaluate the likelihood of sexual recidivism, they might actually be measuring two unrelated concepts, such as sexual offending

and nonsexual offending. Interestingly, recent research found the total score of the J-SOAP-II to be predictive of general and sexual recidivism; however, the *sexual drive/preoccupation scale* was not associated with any of the outcome variables (Martinez, Flores, & Rosenfeld, 2007). Therefore, although the developers of the J-SOAP-II and the ERASOR-II expected the accumulation of sexual and nonsexual factors to predict risk for sexual recidivism, this research suggests that these factors may not be as interrelated as previously thought. Further, items pertaining to sexual interests, behaviours, and offending histories may actually not be adding to the predictive validity of these instruments.

One potential limitation of this study is the relatively small sample of youth that were assessed. Eighty-four sexually offending youth were assessed using the ERASOR-II and 82 with the J-SOAP-II. Although this is not unlike previous research on the J-SOAP-II and the ERASOR-II that had samples ranging from 96 to 153, larger samples are needed to accurately generalize these findings. In addition, a number of difficulties arose when attempting to run the model of the ERASOR-II for the exploratory factor analysis. Specifically, the model would not run until the problematic item one and scale four were omitted. Had this study had a larger sample size, these items may not have posed such a problem. Further, unlike previous research, only 20% of the sample was used for the calculations of interrater reliability. Although this is a standard proportion for testing interrater reliability, a larger sample may have resulted in better agreement between raters as was seen in previous studies with larger samples. Regardless of sample size, however, this study confirmed that neither of these instruments has yet to demonstrate sufficient evidence to be considered a valid instrument for assessing risk of sexually offending youth.

Overall, this study found that compared to the ERASOR-II, the J-SOAP-II was a more reliable measure for this sample of sexually offending youth. While the ERASOR-II had a number of strengths, such as interrater reliability and internal consistency for some of the items, the J-SOAP-II had stronger agreement between raters, better internal consistency of the items, and a statistical rationale for some of the scales. Despite some of the limitations of these instruments, both the J-SOAP-II and the ERASOR-II concluded a similar level of risk for each offender. The results of this study, however, indicated that neither of these instruments has yet proven to be a valid measure for assessing risk to reoffend. Rather, at this stage, a more practical and appropriate use of these instruments would be for descriptive purposes. That is, clinicians may use these instruments as a guide for conducting assessment interviews or as a means to review empirically supported risk factors for treatment planning. However, until these instruments can demonstrate adequate validity, statements regarding risk for recidivism should be avoided.

References

- Abel, G. G., Mittleman, M., Becker, J. V., Rathner, J., & Rouleau, J. L. (1988). Predicting child molesters' response to treatment. *Annals of the New York Academy of Sciences*, 528, 223-234.
- Abel, G. G., & Rouleau, J. J. (1990). The nature and extent of sexual assault. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), *Handbook of sexual assault* (pp. 9-21). New York: Plenum Press.
- Arbuckle, J., & Wothke, W. (1999). *AMOS 4.0 user's guide*. Chicago, IL: Smallwaters Corporation.
- Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods and Research*, 21(2), 205-229.
- Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply*. (pp. 18-32). Thousand Oaks, CA: Sage Publications, Inc.
- Borum, R., Bartel, F., & Forth, A. (2002). *Manual for the Structured Assessment of Violence Risk in Youth (SAVRY)*. Tampa, FL: University of South Florida, Department of Mental Health Law and Policy.
- Campbell, T. W. (2003). *Assessing sex offenders: Problems and pitfalls*. Springfield, IL: Charles C Thomas.
- Cicchetti, D., Bronen, R., Spencer, S., Haut, S., Berg, A., Oliver, P., & Tyrer, P. (2006). Rating scales, scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. *Journal of Nervous and Mental Disease*, 194(8), 557-564.

- Doren, D. (2006). Assessing juveniles' risk within the civil commitment context. In D. S. Prescott (Eds.), *Risk assessment of youth who have sexually abused: Theory, controversy, and emerging strategies* (pp. 100–117). Oklahoma City, OK: Wood & Barnes Publishing.
- Epperson, D. L., Ralston, C. A., Fowers, D., DeWitt, J., & Gore, K. S. (2006). Actuarial risk assessment with juveniles who offend sexually: Development of the Juvenile Sexual Offence Recidivism Risk Assessment Tool – II (JSORRAT-II). In D. S. Prescott (Ed.), *Risk assessment of youth who have sexually abused: Theory, controversy, and emerging strategies* (pp. 100–117). Oklahoma City, OK: Wood & Barnes Publishing.
- Fabian, J. M. (2006). A literature review of the utility of selected violence and sexual violence risk assessment instruments. *Journal of Psychiatry & Law*, 34(3), 307-350.
- Forth, A. E., Kosson, D., & Hare, R. D. (2003). The Psychopathy Checklist: Youth Version. Toronto, Canada: Multi-Health Systems.
- Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20. *Criminal Justice and Behaviour*, 27(1), 97-114.
- Hanson, R. K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy, and Law*, 4, 50-72.
- Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66(2), 348-362.
- Hanson, R. K., & Morton-Bourgon, K. E. (2004). Predictors of sexual recidivism: An updated meta-analysis. Retrieved November 25, 2007 from http://ww2.ps-sp.gc.ca/publications/corrections/pdf/200402_e.pdf

- Harris, G. T., & Rice, M. E. (2003). Actuarial assessment of risk among sex offenders. In R.A. Prentky, E.S. Janus, & M.C. Seto, (Eds.), *Sexually coercive behaviour: Understanding and management*. New York: Annals of the New York Academy of Sciences.
- Hilton, N. Z., Harris, G. T., & Rice, M. E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *Counselling Psychologist*, 34(3), 400-409.
- Hoge, R. D., & Andrews, D. A. (1996). *The Youth Level of Service/Case Management Inventory and manual*. Ottawa, Canada: Carleton University, Department of Psychology.
- Leech, N., Barrett, K., & Morgan, G. (2005). *SPSS for intermediate statistics: Use and interpretation*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law*, 7(2), 409-443.
- Martinez, R., Flores, J., & Rosenfeld, B. (2007). Validity of the juvenile sex offender assessment protocol-II (J-SOAP-II) in a sample of urban minority youth. *Criminal Justice and Behavior*, 34(10), 1284-1295.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30-46.
- McGrath, R. J., Cumming, G. F., & Burchard, B. L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 Nationwide Survey*. Brandon, VT: Safer Society.
- National Adolescent Perpetrator Network (1993). The revised report from the National Task Force on Juvenile Sexual Offending. *Juvenile & Family Court Journal*, 44, 5-120.

- Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioural Research*, 42(1), 185-227.
- Prentky, R., Harris, B., Frizell, K., & Righthand, S. (2000). An actuarial procedure for assessing risk with juvenile sex offenders. *Sexual Abuse: Journal of Research and Treatment*, 12(2), 71-93.
- Prentky, R., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II) manual*. Washington, District of Columbia, US: US Department of Justice, Office of Justice Programs; Office of Juvenile Justice and Delinquency Prevention.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC, US: American Psychological Association.
- Rasmussen, L. A. (2004). Differentiating youth who sexually abuse: Applying a multidimensional framework when assessing and treating subtypes. *Journal of Child Sexual Abuse*, 13(3), 57-82.
- Righthand, S., Prentky, R., Knight, R., Carpenter, E., Hecker, J. E., & Nangle, D. (2005). Factor structure and validation of the Juvenile Sex Offender Assessment Protocol (J-SOAP). *Sexual Abuse: Journal of Research and Treatment*, 17(1), 13-30.
- Tabachnick, B. G., Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: A Pearson Education Company.
- Viljoen, J. L., Scalora, M., Cuadra, L., Bader, S., Chávez, V., Ullman, D., et al. (2008). Assessing risk for violence in adolescents who have sexually offended: A comparison of the J-SOAP-II, J-SORRAT-II, and SAVRY. *Criminal Justice and Behavior*, 35(1), 5-23.

Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism

(ERASOR): Preliminary psychometric data. *Sexual Abuse: Journal of Research and Treatment*, 16(3), 235-254.

Worling, J. R., & Curwen, T. (2000). *Estimate of Adolescent Sexual Offense Recidivism (The*

ERASOR) (Version 1.2). Unpublished Manuscript. Toronto, Ontario: Ontario Ministry of Community and Social Services.

Worling, J. R., & Curwen, T. (2001). *Estimate of Risk of Adolescent Sexual Offence Recidivism*

(*ERASOR*). *Version 2.0*. Toronto, Ontario, Canada: Safe-T Program, Thistlethorn

Regional Centre for Children and Adolescents, Ontario Ministry of Community and Social Services.

3 A MULTILEVEL ANALYSIS OF SEXUALLY OFFENDING YOUTH²

3.1 *Introduction*

There is a growing concern among mental health and justice professionals regarding the suitable strategies for preventing and managing adolescent sexual offending (Harris, Rice, & Quinsey, 1998; Moore, Franey, & Geffner, 2004). This is partially based on recent evidence indicating that young offenders are responsible for a large proportion of sexual offences against adults and children. For instance, The National Crime Victimization Survey (NCVS) found that in the United States in 1997, individuals under the age of 18 were involved in 27% of all serious violent victimizations, including 14% of sexual assaults and 27% of aggravated assaults (Snyder & Sickmund, 1999 as cited in Efta-Breitbach & Freeman, 2004). In addition, adolescents are responsible for approximately 60% of all sexual offences committed against children less than 12 years of age (Bourke & Donohue, 1996). Given the prevalence of sexual crimes committed by adolescents, it is not surprising that increasing attention has been devoted to this population.

The study of sexual crimes is a relatively new area, but for decades researchers have attempted to better understand general offenders and crime patterns. Much of this investigation has focussed on identifying individual factors associated with, or predictive of, offending behaviour (e.g., Efta-Breitbach & Freeman, 2004; Righthand & Welch, 2004). However, research on crime has established that the composition of communities can also influence crime rates and offending patterns suggesting there are structural factors associated with crime (Hartnagel & Lee, 1990; Sampson, 1997; Shaw & McKay, 1942; Wikstrom & Loeber, 2000).

These perspectives regarding the causes of crime have resulted in the two prominent, but distinct,

² A version of this chapter will be submitted for publication. Schoenfeld, T. M., Woodworth, M, & Gretton, H. A multilevel analysis of sexually offending youth.

traditions of individual-level and ecological research (Gottfredson, McNeil, & Gottfredson, 1991). Both of these traditions evaluate the correlates of crime; however, individual-level research focuses on the individual factors, such as family dynamics and peers, whereas ecological research evaluates structural factors, such as population size, urbanization, and other community characteristics. Each of these perspectives has made an invaluable contribution to the study of crime, yet few researchers have attempted to combine these traditions (Wikstrom & Loeber, 2000). In addition, no studies known to date have attempted to combine these perspectives to investigate sexual offending. Simultaneously evaluating individual and community factors may result in a more comprehensive understanding of sexual offending. Indeed, it has been argued that the study of crime has more to gain from combining the individual-level and ecological perspectives than from their continued advancement as independent perspectives (Farrington, Sampson, & Wikstrom, 1993; Reiss, 1986; Wikstrom & Loeber, 2000; Wikstrom, 1991).

This study attempts to fill this gap in previous research by evaluating the relationship between community composition and individual risk factors of sexually offending youth. To better understand youth who commit sexual offences and to provide prevention and intervention strategies that are appropriate for individual offenders and their communities, it is important that youth are not evaluated in isolation from their social and community context. Therefore, the purpose of this study was to provide a regional comparison of sexually offending youth by conducting a multilevel analysis that combines individual-level and ecological research methods. A more comprehensive evaluation of sexual offenders may add valuable insight into the potentially inextricable relationship between communities and young offenders.

3.1.1 Individual-level Research: Identifying and Assessing Risk to Reoffend

In 1990, Andrews, Bonta, and Hoge first introduced the risk principle, which states that offenders should receive intervention and treatment services that match their levels of risk to reoffend. Since then, countless studies have focused on identifying individual risk factors that are associated with, or predictive of, sexual offending among both adults and adolescents. Efta-Breitbach and Freeman (2004) conducted a recent review of studies on sexual offending youth and categorized the most commonly cited risk factors in the literature as the following: (a) incomplete treatment, (b) a family history of dysfunction, (c) prior abuse or maltreatment, (d) delinquent relations and peer relations, (e) characteristics of the youths' sexual offences, (f) deviant arousal and/or sexual maladjustment, and (g) mental health. The identification of these common risk factors has encouraged researchers to devise strategies that can systematically assess an individual's level of risk to reoffend.

One such instrument that has been devised specifically for sexually offending youth and that systematically assesses empirically supported risk factors is the Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II; Prentky & Righthand, 2003). The J-SOAP-II is a 28-item evaluation that assesses the static and dynamic risk factors of youth between the ages of 12 and 18 who have been adjudicated or convicted of a sexual offence. Static risk factors are characteristics or circumstances that are unchangeable, such as being victimized sexually and having a history of offending. By contrast, dynamic risk factors are characteristics of the offender that have the potential to change throughout development or can be altered through intervention. Examples of such factors include motivation for change, peer relationships, living situation, and support systems. Therefore, the two fundamental goals of the J-SOAP-II are to

provide an estimation of risk to reoffend and to provide clinicians with information about areas to target in treatment.

3.1.2 Ecological Research: Evaluating the Association Between Communities and Crime

Ecological research examines crime at the structural level (Jacob, 2006). Proponents of this perspective argue that there are a number of factors beyond the individual that can influence and perpetuate offending behaviour. Accordingly, a large body of research has established that community characteristics are related to crime rates and offending patterns (Jacob, 2006; Osgood & Chambers, 2000; Ouimet, 2000; Sampson, 1997; Schulenberg, 2003; Shaw & McKay, 1942). The most commonly cited structural and social factors associated with crime are urbanization, residential mobility, low socio-economic status, family disruption, and ethnic heterogeneity (Jacob, 2006; Ouimet, 2000; Schulenberg, 2003; Shaw & McKay, 1942; Wells & Weisheit, 2004).

Shaw and McKay's (1942) landmark study was the first to examine the linkages between community characteristics and crime by investigating Chicago neighbourhoods. After repeatedly demonstrating that delinquency rates declined drastically beyond the city center, the researchers concluded that these differences in crime rates were due to the physical condition of houses, income levels, demographic stability, and ethnicity composition of neighbourhoods (Jacob, 2006). Shaw and McKay formulated the theory of structural disorganization to explain how these particular structural factors influence crime rates. It was proposed that low economic status, ethnic heterogeneity, and residential mobility increase crime by decreasing the social organization of communities (Gottfredson et al., 1991). That is, these communities lack cohesiveness, which is the ability of community members to recognize their common values and

goals. In the absence of community cohesiveness, communities are unable to regulate themselves through informal social control and thus criminality arises.

Using similar principles as social disorganization theory, urbanization theory (Wirth, 1938) posits that increased crime rates in city centers are primarily the consequence of population size and density. Proponents of this theory argue that as population size increases, greater variation occurs among the community's residents, resulting in poor communication between community members, and segregation according to race, ethnicity, and socioeconomic status. A lack of commonality among community members' results in a lack of social networks and supports; consequently, the community is unable to regulate itself through informal social control. As such, urbanized communities are no longer able to efficiently regulate the behaviour of their children through social networks and shared socialization practices (Osgood & Chambers, 2000). A lack of informal control leads to an over-reliance on formal means of social control and subsequently increased crime and arrest rates (Schulenberg, 2003).

Ecological research has consistently found that crime rates are higher in urban areas for every category of index crime (Weisheit, Falcone, & Wells, 1999). This pattern has been demonstrated in a number of different countries throughout the world. For instance, Van Dijk (1999) evaluated data from the International Crime Victim Survey for 55 countries and found that urbanization was the strongest predictor of serious crime. In addition, examination of US police data indicated that compared to rural counties, violent crime was between 5 and 10 times higher in the largest cities, and property crime was between 4 and 5 times higher in the largest cities (Weisheit & Donnermeyer, 2000 as cited in Wells & Weisheit, 2004). Moreover, individuals are more likely to be victims of violent assaults in large urban centers than in rural communities (Weisheit et al., 1999). Based on a review of the literature, Wikstrom (1998)

concluded that the differences between urban and rural crime rates tended to persist for police recorded crimes (e.g., Wikstrom, 1991), self-reported youth crime (e.g., Christie, Andenaes, & Skirbekk, 1965), and victimization rates (e.g., Sampson, 1986). In addition, there tends to be a greater fear of crime in urban than in rural areas (Wikstrom & Dolmen, n.d. as cited in Wikstrom, 1998).

Despite general consensus that urbanization is related to crime (McCarthy, 1991), much of the ecological research evaluating this association has focussed on neighbourhoods and cities within the United States. For instance, since Shaw and McKay's (1942) landmark study on neighbourhood crime in Chicago, the majority of ecological research has focussed on the largest cities in the United States, such as New York, Baltimore, Boston, and San Diego (Osgood & Chambers, 2000). Consequently, information about the association between communities and criminal behaviour is often inferred from these American studies (Ouimet, 1999). However, Ouimet argues that comparisons between the United States and Canada are not necessarily warranted given that the high crime rates in cities south of the border is often due to the social conditions particular to the United States. Specifically, high crime rates in particular US cities can be attributed to the prevalence of ghettos, or the residential segregation of the poor, and the accessibility of firearms. Although some Canadian cities suffer from social disorganization, Ouimet contends that there are no real ghettos or centralization of the poor.

The few ecological studies that have evaluated the association between urbanization and crime rates in Canada have yielded inconsistent results. For instance, Hartnagel & Lee (1990) investigated statistics on 88 Canadian cities and found that larger sized cities had higher violent and non-violent crime rates; however, these effects became non significant when other factors, such as socioeconomic status were considered. Leonard (1997) examined Canadian crime

statistics and concluded that crime occurred in large and small census metropolitan agglomerations in equal proportions. Shulenberg (2003) found that when the impact of social disorganization was controlled, overall crime rates were actually negatively correlated with population size. Jacob (2006) found that the associations between urbanization variables (i.e., population size and density) and youth crime rates varied by gender. Fitzgerald, Wisener, and Sovoie (2004) examined the distribution of crime in Montreal and found that high crime neighbourhoods were characterized by increased population density, as well as reduced access to resources, and decreased residential stability.

Finally, Statistics Canada recently conducted a comparison of large urban, small urban, and rural crime rates. In comparison to large urban and rural areas, small urban areas actually had the highest overall crime rates and the highest rates for violent crimes, property crimes, and breaking and entering. Rural areas were found to have the lowest overall crime rates and large urban areas reported the highest rates for both robbery and motor vehicle theft, but had the lowest violent crime rates overall. This report concluded that crime is not necessarily a large urban phenomenon, suggesting that factors other than urbanization may be influencing Canadian crime rates and offending patterns (Francisco & Chenier, 2007).

The results of ecological research investigating the association between urbanization and crime rates in Canada have been inconsistent at best. It remains unclear exactly how urbanization affects the residents of a community, particularly those residing in Canadian communities. This is primarily because there is a lack of ecological research on Canadian crime in general. As such, it is evident that further local ecological research, which is sensitive to the unique dynamics of Canadian communities, is much needed.

In addition to studying crime in Canada, it has been argued that ecological research could benefit from studying a greater variety of communities. Although the majority of ecological research focuses on neighbourhoods and communities within metropolitan cities, there is evidence that these theories can be generalized to nonmetropolitan or rural communities. For example, while assessing rural communities Osgood and Chambers (2000) found that there was a curvilinear relationship between juvenile violence and population size, with the smallest communities having the lowest offending rates. These researchers concluded that to truly understand the association between communities and crime, the full variety of communities must be studied. Comparing offending behaviour among large and small urban cities may provide evidence for the applicability of ecological theories in a variety of settings.

In addition to studying a variety of communities, ecological research on crime would also benefit from the evaluation of different types of crime. As mentioned above, social ecological studies of crime have traditionally focused on community correlates of overall crime rates, often only differentiating between violent and non-violent crime, and neglecting to focus on sexual offending specifically. In addition, few studies have evaluated the influence of community factors on aspects of criminal careers (Wikstron & Loeber, 2000), such as an individual's propensity to reoffend. Sexual offenders have unique behavioural, social, and psychological difficulties which could result in an unusual dynamic between offenders and their communities. Moreover, assessing risk to reoffend rather than crime rates and general offending patterns may provide insight into the community's influence on criminal careers. Further, research evaluating the community's influence on an individual's risk of sexual offending may inform about prevention and intervention strategies that are unique and suitable for particular regions.

3.1.3 *Purpose of Study*

The purpose of this study was to provide a multilevel analysis of sexually offending youth by combining ecological and individual-level perspectives on crime. To do this, individual risk factors of sexually offending youth in three neighbouring, but diverse Western Canadian regions of the Greater Vancouver Area (GVA), Central Okanagan (CO), and Thompson Nicola (TN) were compared. These regions vary according to three variables that are commonly evaluated in ecological research (i.e., population size, density, and ethnic heterogeneity) and that measure the degree of urbanization and social disorganization in particular communities. Specifically, the 2006 census metropolitan and agglomeration profiles provided by Statistics Canada (2007a, b, & c) indicated that according to population size and density per square kilometre, the GVA is the largest and most dense region (pop. = 2,116,581; density = 735.6), CO is considerably smaller (pop. = 162,276; density = 55.9) than GVA, but TN is the smallest and least dense of the three (pop. = 122,286; density = 2.7). In addition, GVA is the most ethnically heterogeneous city of the three with 36.9% of the total population being comprised of visible minorities, as opposed to only 3.9% in CO and 5.4% in TN.

It should be noted that unlike traditional ecological research, associations between the ecological variables (i.e., population size, density, and ethnic heterogeneity) and crime were not directly tested. Rather these variables were used for descriptive purposes only, as a means of illustrating the variation among the composition of these communities. Further, unlike traditional ecological research, the dependent variables of this study were the individual risk factors of offenders. Ecological research traditionally uses crime, arrest, or victimization rates to assess the influence of structural factors on crime. This study, rather, was interested in how community composition influences individual offenders and their risk to reoffend. Accordingly, the

dependent variables of this study were scores on the J-SOAP-II. This instrument was used because it is specific to adolescents and it provides a systematic way of assessing characteristics of youth who sexually offend.

This is believed to be the first study to combine individual-level and ecological perspectives to assess sexually offending youth. As mentioned above, ecological researchers propose that crime is generally a consequence of structural factors, such as urbanization and heterogeneity; however, the association between these structural factors and Canadian crime is not well understood. If ecological theories hold true for Canada and larger communities are more at risk for crime, it can be assumed that the residents have a greater number of risk factors for engaging in crime and consequently are more at risk to reoffend. Therefore, to test the generalizability of these ecological theories to Canadian communities, it is hypothesized that sexually offending youth in the more populated, dense, and ethnically heterogenic region of GVA will be highest risk to reoffend and youth in the least urbanized community of TN will be at lowest risk for reoffending. This multilevel analysis will provide a more comprehensive evaluation of sexually offending youth while providing information about the prevention and intervention strategies that are appropriate for offenders in particular regions.

3.2 Method

3.2.1 *Participants*

The sample was comprised of male adolescents between the ages of 11 and 20 who had committed a sexual offence and had received treatment at Youth Forensic Psychiatric Services (YFPS) located in Burnaby, Kamloops, Kelowna, and Penticton, British Columbia between 1997 and 2006. YFPS has a number of mental health outpatient facilities in British Columbia that specialize in assessing and treating youth who have been charged with crimes ranging in severity

from relatively minor offences (e.g., shoplifting) to violent crimes, including sexual offences. YFPS offers a range of treatment strategies including individual, family, and group therapy, with sexually offending youth often receiving a combination of these strategies as part of the specialized Youth Sex Offender Treatment Program (YSOTP).

A number of challenges arose while attempting to achieve representative and comparable samples from each geographical region. First, each YFPS outpatient clinic is responsible for servicing local youth and the youth of surrounding regions. As a result, samples from each clinic included youth who lived in the immediate locale of the clinic as well as in surrounding areas, such as suburbs, nearby towns, or rural communities. Information was collected when available about the location of the youths' residence at the time of assessment. Cases missing this information were excluded from analysis. The final sample of offenders was expected to adequately reflect sexually offending youth in the Greater Vancouver Area (GVA), Central Okanagan (CO), and Thompson Nicola (TN) regions. Second, while random selection was possible for the sample from the Burnaby clinic, it was not for Kamloops, Kelowna, and Penticton samples. The GVA sample was randomly selected from all youth who had committed a sexual offence and been admitted to the YSOTP located in Burnaby. However, the CO and TN samples could not be selected randomly for two reasons. First, the YSOTP has not been operating for as long in the Kamloops, Kelowna, and Penticton clinics as it has in the Burnaby clinic. Thus, there were fewer clients at these clinics who had been treated between 1997 and 2006 resulting in a smaller population to sample from. Second, although adolescents are responsible for a large proportion of sexual offences, the ratio of committed offences to arrests can be as low as 1:150 (Abel, Becker, Mittelman, & Cunningham-Rathner, 1987). Therefore, few offenders may actually receive intervention and services through YFPS. For smaller

communities, this means fewer convicted offenders resulting in a smaller population to sample from than from the larger region of GVA. Therefore, to ensure sufficient sample size, every youth who had committed a sexual offence and been admitted to the YSOTP in Kamloops, Kelowna, and Penticton was sampled for this study.

Considering that every youth was selected and assessed, it can be concluded with confidence that our sample represents the sexual offenders being treated in the Kamloops, Kelowna, and Penticton clinics. Refer to Table 3-1 for the average age at the time of assessment and the ethnic composition of the offenders treated in each region.

Table 3-1 Average Age and Ethnicity of Offenders in Each Region

Region	<i>Age</i>		<i>Ethnicity</i>									
	<i>M</i>	<i>SD</i>	<u><i>Caucasian</i></u>		<u><i>Aboriginal</i></u>		<u><i>Asian</i></u>		<u><i>Indo-Canadian</i></u>		<u><i>Other</i></u>	
			<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
GMA	16.7	1.4	9	50.0	4	22.2	3	16.7	2	11.1	0.0	0.0
CO	15.8	1.8	18	85.7	2	9.5	0	0.0	0	0.0	1	4.8
TN	15.7	2.0	21	72.4	6	20.7	0	0.0	0	0.0	2	6.9
Overall	16.1	1.8	48	70.6	12	17.6	3	4.4	2	2.9	3	4.4

3.2.2 *Classifying Regions – Independent Variable*

Consistent with previous ecological research, census metropolitan and agglomeration data were collected from Statistics Canada (2007a, b, & c) to categorize the regions for this study. Based on this information three regions were established for comparison: GVA, CO, and TN. The location of residence at the time of assessment was collected for each offender. Because it was rare for an offender to live in the immediate locale of one of these clinics, offenders were classified according to the census metropolitan, agglomeration, or city in which they were located. According to Statistics Canada, to be classified as a census metropolitan or agglomeration, there must be an urban core with a population of at least 50,000. In addition,

adjacent urban and rural areas must have a high degree of social and economic integration with the urban core, as evidenced by commuting flows to and from work.

Classification of regions was conducted in three steps. First, considering that these regions should be highly integrated, urban and rural areas within the census metropolitan or agglomeration were categorized into one of three regions: Vancouver, Kelowna, and Kamloops. Second, geographically adjacent census agglomerations were combined with nearby regions. For instance, the agglomerations of Penticton and Vernon geographically surround the Kelowna region, and therefore, were classified as Kelowna area. Third, communities outside of an agglomeration were combined with geographically adjacent agglomerations. For example, because of its geographical location, the small community of Lillooet was combined with the agglomeration of Kamloops. Therefore, the regions of GVA, CO, and TN include cities, suburbs, and in some cases rural areas. Although these categorizations may include a variety of community profiles, these communities should be relatively integrated, either because they are located within the census metropolitan or agglomeration areas or they are geographically near each other. Regardless, these classifications adequately reflect the youth from a variety of communities that are being treated in each respective YFPS clinic. Table 3-2 displays the offenders' location of residence at the time of assessment and the categorization of these communities into the three regions of GVA, CO, and TN.

Table 3-2 Location of Residence at Assessment

<u>Greater Vancouver Area</u>			<u>Central Okanagan</u>			<u>Thompson Nicola</u>		
Region	<i>n</i>	%	Region	<i>n</i>	%	Region	<i>n</i>	%
Abbotsford	2	8.3	Coldstream	1	4.3	108 Mile Ranch	1	3.3
Burnaby	3	12.5	Kelowna	13	56.5	Castlegar	1	3.3
Chilliwack	3	12.5	Naramata	1	4.3	Chase	1	3.3
Coquitlam	1	4.2	Penticton	4	17.4	Golden	1	3.3
Courtenay	1	4.2	Trail	1	4.3	Kamloops	21	70.0
Delta	1	4.2	Vernon	3	13.0	Lillooet	1	3.3
Langley	3	12.5				Salmon Arm	2	6.7
Mission	1	4.2				Sorrento	1	3.3
N. Vancouver	1	4.2				Williams Lake	1	3.3
Port Moody	1	4.2						
Surrey	4	16.7						
Vancouver	3	12.5						
Total	24	100		23	100		30	100

3.2.3 Categorizing Risk Factors – Dependent Variable

The Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II; Prentky & Righthand, 2003) is a 28-item evaluation divided into four scales: scale 1 - *sexual drive/preoccupation*, scale 2 - *impulsive/antisocial behaviour*, scale 3 - *intervention*, and scale 4 - *community stability/adjustment*. Scales 1 and 2 assess static risk factors and scales 3 and 4 assess dynamic factors. Items are scored from 0 to 2 indicating apparent absence to clear presence of the risk factor and final scoring involves summing the scores on each scale, adding the scales together and dividing by the total possible score for the final risk rating. The final risk rating is then expressed as a proportion (i.e., there is a 67% chance of reoffending sexually).

The individual items and scales of the J-SOAP-II have demonstrated relatively good internal consistency. For instance, Schoenfeld, Brown, Woodworth, & Gretton (n.d.) conducted a structure analysis of the J-SOAP-II on this same sample of offenders and found that while individual inter-item total correlations ranged from poor to excellent, each of the scales had good to excellent scale alpha's (.69 to .90). One of the goals of the J-SOAP-II is to provide clinicians

with information about offenders' risk factors that are stable or subject to change through treatment. Consequently, the scales of the J-SOAP-II are categorized to reflect static and dynamic risk. The purpose of this study, conversely, is to describe the influence that community composition has on individual risk factors associated with certain areas of psychosocial functioning. Therefore, the individual items of the J-SOAP-II were reorganized according to their conceived fit with other similar items.

For example, the item *sexual victimization* was originally included in the *sexual drive/preoccupation scale*. Although this classification is appropriate for the purpose of the J-SOAP-II, it is not an adequate classification for this study. All other items on this scale concern the sexual offending behaviour of the offender, whereas, sexual victimization history refers to the offender's own experience of abuse. Therefore, the items on this scale seem to be assessing two very different constructs, sexual abuse perpetrated by the offender and sexual abuse perpetrated against the offender. Although these items are obviously related, they reflect different types of experiences, and therefore, the item of *sexual victimization history* was included in the newly created *family/environment scale*. Four new scales were created by reorganizing the J-SOAP-II items. The components of the newly created scales, including newly calculated scale Cronbach alphas and inter-item total correlations (ITCs) are summarized below.

Sexual Offending History Scale (Cronbach's $\alpha = .70$). Items on this scale pertain to the sexual offending behaviour of the offenders. It includes the number of *prior legally charged sex offences* (.15), *number of sexual abuse victims* (.71), whether there has ever been a *male child victim* (.32), the length of time the offending has been occurring (*duration of sex offence history* (.53)), the *degree of planning in sexual offence(s)* (.51), the use of *sexualized aggression* in the offence(s) (.20), and the *sexual drive and preoccupation* (.47) of the offender. Although the ITCs

were low on many of these items, the Cronbach's alpha suggests that the items of the scale fit well together overall.

Problem Behaviour Scale (Cronbach's $\alpha = .91$). This scale assesses a wide range of problem behaviours that may have manifested throughout the offender's development. Specifically, the items on this scale assess *school behaviour problems* (.74) from kindergarten through grade 8, *history of conduct disorder* (.82), *juvenile antisocial behaviour* (.75), whether the offender was *ever charged/arrested before age 16* (.59), the *multiple types of offences* (.65) the offender has been charged with, the *quality of peer relationships* he has (.56), the offender's *stability in school* (.66), and the *degree of pervasive anger* (.78) he expresses.

Intervention Scale (Cronbach's $\alpha = .86$). Items on this scale are essentially characteristics internal to the offender that may assist or impeded treatment success. Items include whether the offender is accepting *responsibility for offence(s)* (.61), has an *internal motivation for change* (.68), *understands risk factors* (.65), expresses *empathy* (.63), expresses *remorse and guilt* (.77), holds *cognitive distortions* (.55) about offending behaviours, has shown *management of sexual urges and desires* (.47), and *management of anger* (.51). All items on this scale are dynamic factors, and thus are subject to change. Changes in these items are reflected in scoring. For example, if an offender initially accepts no responsibility for his offence or is in complete denial, but over the course of treatment he is able to admit to his offence and accept responsibility, his score on this factor should reflect that change. Therefore, this offender would most likely receive a score of 1 suggesting partial acceptance of responsibility, rather than a score of 2 indicating no acceptance or a score of 0 indicating full acceptance of responsibility.

Family/Environment Scale (Cronbach's $\alpha = .74$). This scale mainly assesses the dynamics of the offender's family and social networks. The items of this scale evaluate the offender's

sexual victimization history (.40), history of physical assault and/or exposure to family violence (.57), the degree of caregiver consistency (.52) measured by the number of different caregivers up to age 10, the stability of the current living situation (.41), and evidence of positive support systems (.65).

3.2.4 Background Information

In addition to scoring the J-SOAP-II, background information was collected to assess demographic and treatment information.

3.2.5 Procedure

Trained researchers scored the J-SOAP-II and coded background variables using information in closed files on youth who had committed a sexual offence and been admitted into the YSOTP in the regions of GVA ($n = 24$), CO ($n = 23$), and TN ($n = 30$). Seven offenders were missing information about their location of residence at the time of assessment and were excluded from analysis. In addition, the J-SOAP-II is not intended to be scored on youth who commit noncontact sexual offences, such as exhibitionism. Two youth who committed noncontact offences were excluded, resulting in a final sample size of 77 offenders.

To ensure consistency between researchers, interrater agreement on the J-SOAP-II was calculated for a random sample of cases (20.0%). Considering that not all of the cases were scored by the same two raters, a two-way random effects model was selected. In addition, the more conservative test of absolute agreement was used because this test measures the potential differences between each rater on each case rather than examining patterns of scores as calculated by the consistency test (McGraw & Wong, 1996). Because there were two raters for only a selection of the cases, all further analyses used the score of rater 1 rather than an average

of rater 1 and rater 2. Thus, the more conservative test of single measure intraclass correlation coefficients (ICCs) are reported for the interrater agreement test.

The J-SOAP-II *proportion of risk score* demonstrated excellent interrater agreement (ICC = .94; for interpretation guidelines on the strength of agreement for ICCs, see Cicchetti et al., 2006). Individual items were found to have fair to excellent agreement between raters. Specifically, 24 of 28 items had moderate to perfect agreement between raters (ICCs = .41 to 1.0); whereas, the items *multiple types of offences*, *empathy*, *remorse and guilt*, and *management of sexual urges and desires* were only fair in their strength of agreement (ICCs = .25 - .35). Overall, however, the J-SOAP was shown to be relatively consistent between raters.

3.3 Results

3.3.1 Region and Risk

A one-way analysis of variance (ANOVA) was conducted to evaluate the relationship between region and the proportion of risk to reoffend as indicated by the J-SOAP-II. The ANOVA was significant $F(2, 72) = 6.17, p = .003, \eta^2 = .15, MSE = 342.32$. Follow-up tests using the Bonferroni adjustment were calculated to investigate differences among the means for each region. The results indicated that TN youth had a greater number of risk factors ($M = 47.05$) than CO youth ($M = 29.22$) resulting in a greater overall level of risk to reoffend. There was no difference found between youth residing in TN and GVA ($M = 35.44$), as well as CO and GVA.

To further assess the regional differences in overall risk levels, a Multivariate Analysis of Variance (MANOVA) was performed on the four newly developed J-SOAP-II scales (*sexual offending history*, *intervention*, *general problem behaviour*, and *family/environment*). Prior to analysis, all assumptions of MANOVA were assessed and were met.

The Wilks' Lambda criterion showed there to be a significant effect of region on the combined DVs, $F(8, 138) = 2.20, p = .031$. The strength of the relationship between region and J-SOAP-II scales, as assessed by eta squared, was moderate with region accounting for 11% of the variance of the combined DVs. Tests of between subjects effects indicated there to be no significant effect of region on *sexual offending history*, $F(2, 72) = 1.09, p = .343, \eta^2 = .03, MSE = 8.95$, but there was a regional effect on the *intervention scale*, $F(2, 72) = 5.45, p = .006, \eta^2 = .13, MSE = 64.48$, the *general problem behaviour scale*, $F(2, 72) = 7.79, p = .001, \eta^2 = .18, MSE = 156.59$, and the *family/environment scale*, $F(2, 72) = 3.36, p = .040, \eta^2 = .09, MSE = 18.64$.

To investigate the impact of each main effect on the individual DVs, the Bonferroni procedure was utilized. Table 3 displays the mean J-SOAP-II score and standard deviation for each scale and region. Post hoc analyses revealed that compared to CO and GVA, TN youth scored highest on the *intervention scale* and the *general problem behaviour scale* indicating that TN youth are higher risk in these areas. CO and GVA youth scored similarly on these scales. In addition, TN youth scored higher than those in CO on the *family/environment scale*, but there was no difference between youth in TN and GVA, and CO and GVA. These results suggest that TN youth generally are more at risk than youth in CO or GVA because they possess a greater number of risk factors in the areas of intervention, general problem behaviour, and family/environment functioning.

Table 3-3 Mean and Standard Deviation of J-SOAP-II Scale Scores

JSOAP-II Scale	Region					
	GVA		CO		TN	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Sexual Offending History	4.8	2.4	3.7	3.2	4.8	3.0
Intervention	5.6	4.1	5.0	2.9	8.0	3.3
Problem Behaviour	5.6	4.9	5.0	2.9	9.4	4.5
Family/Environment	3.1	2.7	2.4	2.2	4.1	2.1

To determine exactly which individual risk factors on each scale were affected by region, a series Kruskal-Wallis tests were conducted on the items of each of the significant scales (*intervention, problem behaviour, family/environment functioning*). The Bonferroni method was used to control the familywise error (FW) rate to .05 for each scale and to obtain an appropriate level for each of the comparisons.

Intervention scale. The Bonferroni method was used to adjust the alpha for eight comparisons ($\alpha = .006$). Only the item of *cognitive distortions* met the new significance level, $\chi^2(2, N = 73) = 13.65, p = .001$. Other individual risk factors to approach significance were *understands risk factors*, $\chi^2(2, N = 73) = 7.94, p = .019$, and *internal motivation for change*, $\chi^2(2, N = 73) = 6.87, p = .032$.

A series of Mann-Whitney U tests were conducted to evaluate which particular region scored highest on the significant risk factor of *cognitive distortions*. The Bonferroni correction was used to adjust the alpha of .05 for three comparisons of the independent variable ($\alpha = .02$). There was a significant difference between TN and CO, $z = -2.33, p = .020$, as well as between TN and GVA, $z = -3.405, p = .001$. There was, however, no significant difference between CO and GVA, $z = -1.571, p = .116$. Evaluation of the mean rank scores for each comparison revealed that TN youth generally scored higher on the item of *cognitive distortions* than did CO and GVA youth.

Problem behaviour scale. The Bonferroni adjustment for eight comparisons ($\alpha = .006$) resulted in a significant association between region and three of the risk factors on this scale: *quality of peer relationships*, $\chi^2(2, N = 74) = 14.67, p = .001$, *juvenile antisocial behaviour*, $\chi^2(2, N = 75) = 12.35, p = .002$, and *school behaviour problems*, $\chi^2(2, N = 74) = 12.47, p = .002$. *Ever charged or arrested before age 16*, $\chi^2(2, N = 75) = 9.93, p = .007$ approached, but did not reach

significance. There were no associations between region and *pervasive anger*, $\chi^2(2, N = 75) = 7.25, p = .027$, *history of conduct disorder*, $\chi^2(2, N = 72) = 7.05, p = .029$, *stability in school*, $\chi^2(2, N = 73) = 6.73, p = .035$, and *multiple types of offences*, $\chi^2(2, N = 75) = 4.57, p = .102$.

The three significant risk factors were further evaluated to better assess the differences between regions. The Bonferroni method was used to adjust the alpha of .05 for three comparisons of the independent variable ($\alpha = .02$). The results of the Mann-Whitney U tests and evaluation of the mean rank scores indicated that TN youth scored significantly higher on the item *quality of peer relationships* than youth in CO, $z = -3.53, p < .001$, and in GVA, $z = -2.51, p = .012$. Similarly, TN youth scored higher than youth in CO, $z = -3.33, p = .001$, and GVA, $z = -2.48, p = .013$, on the item *juvenile antisocial behaviour*. Moreover, youth in TN had significantly higher scores on the item *school behaviour problems* than did youth in CO, $z = -3.31, p = .001$, and GVA, $z = -2.58, p = .010$. There were no significant regional differences in scores between CO and GVA on the individual items of this scale.

Family/environment scale. The Bonferroni method for five comparisons resulted in an adjusted alpha of .01. There was a significant association between region and *caregiver consistency*, $\chi^2(2, N = 75) = 11.60, p = .003$. No other items of this scale reached or approached significance (all p 's $\geq .05$) indicating no regional differences on the items *physical abuse/exposure to family violence*, *social support*, *stability of living situation*, and *sexual victimization history*.

Mann-Whitney U tests of the regional comparisons using the adjusted alpha of .02 revealed that TN youth had significantly less *caregiver consistency* than did youth in GVA, $z = -3.08, p = .002$, and CO, $z = -2.71, p = .007$.

3.3.2 *YSOTP Initiation, Duration, and Completion*

ANOVAs were calculated and a chi-square contingency table was created to assess the time lapse from the index offence to YFPS assessment, age at treatment initiation, duration of treatment, and completion rates for youth in each region. The Bonferroni adjustment was used to control the FW error rate for four comparisons ($\alpha = .01$). Youth in TN, CO, and GVA did not differ in the time lapse from the commission of the index offence to YFPS assessment, $F(2, 74) = .95, p = .393, \eta^2 = .03, MSE = 309.69$, the age at beginning treatment, $F(2, 74) = 2.49, p = .09, \eta^2 = .06, MSE = 8.18$, the number of months in treatment, $F(2, 74) = 2.72, p = .072, \eta^2 = .07, MSE = 155.77$, and the degree of treatment success, $\chi^2(4, N = 76) = 2.10, p = .718, V = .166$.

3.4 Discussion

Ecological theories of crime have traditionally argued that the organization of a community can influence crime rates and offending patterns. Conversely, proponents of the individual-level perspective identify and assess individual risk factors for engaging in criminal behaviour. This study was the first to provide a multilevel analysis of sexually offending youth by combining individual-level and ecological perspectives on crime. Specifically, risk assessments were conducted on sexually offending youth from the diverse regions of Thompson Nicola (TN), Central Okanagan (CO), and Greater Vancouver Area (GVA) to evaluate the criminal behaviour and individual characteristics of offenders in each region. As a result, three key findings emerged. First, compared to CO and GVA, sexually offending youth in the least populated, least dense, and least heterogeneous region of TN were highest risk to reoffend sexually. Second, TN's elevated overall risk level was primarily due to the youth having more risk factors pertaining to behavioural problems, intervention, and family/environment functioning. Third, despite regional differences in overall risk levels, youth in GVA, CO, and TN

did not differ in their sexual offending histories or in treatment duration and completion rates. Although this study did not find support for the argument that urbanized communities are most at-risk for crime, it does highlight the need to assess individual offenders in conjunction with their community and social contexts.

Recent Canadian research has indicated that violent crime may be a characteristic of small, rather than large urban areas (Francisco & Chenier, 2007). Accordingly, our findings suggested that youth in less urbanized areas have more risk factors for engaging in criminal behaviour than do youth residing in more urbanized centers. One possible explanation for this finding is that smaller communities do not have access to the services and resources that may be available to individuals in larger and more inhabited areas. For instance, Gumpert and Saltman (1998) examined rural treatment practices and indicated unique treatment issues for rural clinicians to be a lack of resources and skilled workers. Similarly, a recent study evaluating treatment services for sexually offending youth in two communities found that a treatment facility in the non urban area of Prince George was unable to recruit an adequate number of specialists despite national advertisements for available positions (Smiley, 2007). Consequently, specialists must travel from urban to rural areas to conduct assessments or treatment resulting in scheduling and travelling conflicts for offenders and their families, as well as longer waits for treatment. Further, Jennissen (1992) argued that the problem of few specialists in rural areas is exacerbated by a lack of information and support in a number of preventative areas, such as sex education and family violence.

TN has a well-established, high-quality treatment facility for sexually offending youth. However, young offenders are only referred to this clinic as a result of formal intervention through a charge and/or conviction. Considering that less urbanized areas struggle with a lack of

resources, it is possible that many of these youth do not receive adequate prevention and informal intervention before they commit their first sexual offence. For instance, prevention programs, such as sex education, are often offered at the community level through the medium of schools. If there are few teachers and money is lacking for education, it is likely that these non-academic prevention programs are first to be omitted from the curriculum. In addition, TN youth possessed a greater number of risk factors pertaining to general delinquency and school behaviour problems. If resources are lacking in these small communities, it is possible that teachers, law officers, and other professionals do not have the time or the resources to adequately intervene when problem behaviours are evident. Without adequate intervention, it is likely that problem behaviours will persist and potentially escalate with age. Therefore, scarce resources for prevention and early intervention in smaller communities such as TN may result in the offending youth there having more behavioural difficulties, as well as being criminally entrenched and high risk to reoffend by the time formal intervention is finally provided.

In addition to lacking resources for prevention and informal intervention, there is some evidence to suggest that smaller communities may have slower judicial processes resulting in slower intervention once a charge is formally laid. For instance, Smiley (2007) compared the characteristics of sexually offending youth treated in two YFPS clinics located in Prince George and Burnaby, British Columbia. Although the young sexual offenders in Northern British Columbia were before the courts on their first sexual offence charge more often than their Burnaby counterparts, the time lapsed from the date of the offence to the date of being judicially processed was greater for Prince George than for Burnaby youth. Specifically, the average length of time from the index offence to referral for assessment or treatment in Prince George was 16.92 months, compared to only 8.64 months for youth in Burnaby. The results suggest that

youth in the metropolitan area of the province are before the courts and referred for assessment or treatment in almost half the time as youth in the smaller more rural region of Northern British Columbia.

This study found no significant regional differences in time from the commission of the index offence to YFPS assessment. However, this is not to say that youth in each region were processed at a similar speed. The data in this study only describe the time lapse between the actual commission of the index offence and the time of assessment. There were, unfortunately, no data indicating the date the actual charge was laid for the index offence. It is possible that youth in the more metropolitan regions were not being charged for the index offence as quickly, but received swift processing once formal charges were laid. In addition, although there were no regional differences in the time lapse from offence commission to YFPS assessment for sexual offences, this is not to say that other problem behaviours receive quick or appropriate intervention especially considering the potential for limited services in smaller communities. Given that early intervention can be a powerful deterrent for engaging in further criminal behaviour and that treatment can reduce the risk for reoffending (Hall, 1995; Loeber & Farrington, 2000; Welsh & Farrington, 2007), communities that do not provide swift interventions may be unknowingly perpetuating offending behaviour and creating high risk youth, such as in TN.

This study suggests that there are community characteristics that increase the propensity to engage in offending behaviour. There may also be community factors, however, that protect youth from initiating and maintaining criminal behaviour. Wikstrom and Loeber (2000) argue that highly disadvantaged structural characteristics tend to foster social processes that create a high risk social context, whereas, communities with highly advantaged structural factors tend to

produce social processes that provide a protective social context. Income, social cohesiveness among members, and religious affiliation are potential protective factors that can provide a low risk social context for youth resulting in lower risk offenders. While the community factors of the TN region may be producing high risk offenders, it is just as likely that the regions of CO and GVA have community factors that are protecting the youth there from engaging in or maintaining criminal behaviour. To adequately reduce the risk of young offenders it would be beneficial to evaluate risk factors in conjunction with protective factors to identify which community characteristics are increasing or impeding criminal offending.

TN youth had the most risk factors pertaining to general offending and behavioural problems. The youth, however, did not differ in their sexual offending histories. Specifically, youth in each region were found to have committed a similar number of sexual offences, had similar types of victims, and displayed a similar number of sexually inappropriate behaviours. There are two potential explanations for this finding. First, sexual offending is generally not a social crime. Most often sexual offences are committed in private, in someone's home, and in secrecy. General offending and other behavioural problems, on the other hand, are often influenced by peer groups and can be reinforced through approval and status achievement. Sexual offending is considered by most to be an intolerable crime. Youth who commit these offences are often ostracized by their peers and other community members. If youth are not motivated socially to engage in sexual crimes, by peers or to acquire social status, it is unlikely that some of the traditional sociological factors such as poverty, heterogeneity, and density will influence the specific characteristics of sexual offences.

Second, ecological theories propose that criminal offending occurs when there is a lack of cohesiveness among community members. Communities that are well-integrated through

common values and goals should deter crime because if community members are known to each other and care about each other they theoretically should not want to commit crimes and harm against each other. This is not necessarily the case for sexual offences. For instance, in general population studies, it has been found that 80% of female and 60% of male child sexual assaults were committed by someone known to the victim or by a family member (Finkelhor, Hotaling, Lewis, & Smith, 1990). Moreover, approximately half of the offenders in clinical samples of child victims are parental figures or relatives (Elliott & Briere, 1994). The nature of the offender/victim relationship for sexual offences seems to conflict with the theoretical assumptions of ecological research, which could explain the lack of regional differences found for this type of crime. Overall, this study suggests that sexual offending is a unique offending type and characteristics of the offence may not be as readily determined by social or structural factors.

While this study found that community characteristics were associated with individual risk, it was unable to support traditional ecological theories on crime for three plausible reasons. First, traditional ecological research has largely evaluated American neighbourhoods and cities. Therefore, results from these studies have driven ecological theorists to assume that crime rates are generally an urban phenomenon. However, there are structural characteristics that are unique to American cities, which limit the generalizability of these results to Canadian communities. For instance, access to firearms, as well as segregation of the poor and the prevalence of ghettos in the United States produce social contexts that are unique to that country (Ouimet, 2000). Research evaluating ecological theories on Canadian communities thus far has been inconclusive at best and it has been argued that more ecological research on Canadian communities is much needed (Jacob, 2006). This study suggests that while community factors may be related to

offending behaviour, the structural factors influencing crime in Canadian communities might be quite different than those in communities south of the border.

Second, the differences found between this study and previous ecological studies could be attributed to particular structural variables that were not directly tested. This study was unlike traditional ecological research in that it did not directly test the association between structural factors and overall crime rates. Rather, the community factors of population size, density, and cultural heterogeneity were used to categorize the regions accordingly. Some structural variables, however, such as income and residential mobility were not used for these descriptive purposes. Assessing these particular variables in future studies could add insight into the factors that influence Canadian crime, particularly in small urban communities.

Lastly, while ecological research has found support for the association between community composition and general offending patterns, it has not specifically evaluated the influence of communities on the progression of serious offending or criminal careers. This study was unique in that it examined the communities' influence on individual characteristics that have been identified as predictors of reoffending. Therefore, it could be that traditional ecological theories are adequate explanations for current crime rates, but not for the potential to commit future crimes and especially future sexual offences.

While there are benefits to not using traditional ecological research methods, the non-traditional approach of comprehensively evaluating each offender's individual risk factors resulted in the limitation of a small sample size. Given resource and time constraints, it was not possible to do a large scale study including more youth in various regions, as is often the case with ecological research. The effect of a small sample size was that potential and less salient differences among the regions may have been missed because of a loss of statistical power. For

instance, an anecdotal comparison of the regional means seemed to indicate that compared to TN and GVA, CO youth were considerably lower risk on all scales of the J-SOAP-II. Statistical analyses, however, did not indicate these to be significant differences. In addition, when the risk factors were assessed individually, rather than as scales, few regional differences were found. This could possibly be because less salient community effects were missed with the small sample of youth from each region. Therefore, a larger sample size could possibly have demonstrated an interesting trend between small, medium, and large urban centers, as well as identified more individual risk factors as being associated with particular regions.

Another limitation of this study was that it was difficult to decipher the exact contribution that each community had on its individual offenders. For instance, regions were categorized according to census agglomeration and metropolitan data, as well as by geographical location. For this reason, it was assumed that there was some degree of integration among the communities of each region. A larger sample size, however, could have resulted in a more precise classification system that better illustrated which particular communities were most strongly associated with risk to reoffend. Nevertheless, this classification of the three regions was relatively consistent with previous ecological research and provides invaluable information to YFPS about the unique characteristics of youth being assessed and treated in each region.

YFPS offers an empirically guided, highly structured, and standardized treatment program for sexually offending youth. The Youth Sex Offender Treatment Program (YSOTP) is a combination of individual and/or group treatment sessions aimed at educating and rehabilitating young sexual offenders. Though a standardized treatment program is beneficial for ensuring empirically sound and consistent intervention strategies, this study suggests that it may be advantageous to consider the characteristics of the community in order to deliver successful

treatment. For instance, the standardized YSOTP is offered in the regions of TN, CO, and GVA, yet this study found that the youth in TN may be in need of unique and potentially more intensive treatment targeting both sexual and nonsexual problem behaviours. Youth in TN were higher risk because they had more risk factors pertaining to general problem behaviours, such as school difficulties and prior nonsexual offences. These results suggest that sexually offending youth in TN may benefit from a treatment approach that is devised for the particular needs of youth in that region. If the fundamental goal of treatment is to reduce the risk of reoffending, TN youth may benefit from treatment that not only targets sexually offending behaviours, but also emphasises improving general psychosocial functioning and preventing or reducing other nonsexual criminal behaviour. In addition, the risk assessments for this study were conducted only on youth who had already completed or been discharged from treatment at YFPS.

Therefore, even after treatment had been provided, youth in TN youth continued to have the greatest number of risk factors. Considering that service deliverers at YFPS are often limited in the amount of treatment they can provide by probation orders and sentencing, there were no significant regional differences in the amount of time spent in treatment and the degree of treatment success. This study suggests that while youth in the TN region receive the same type of treatment for the same amount or time and have the same rate of treatment success as youth in CO and GVA, they remain at highest risk. Treatment that addresses the unique needs of TN youth until risk is effectively lowered may result in fewer future offences, justifying the need for regionally tailored intervention strategies.

This study is believed to be the first to simultaneously evaluate structural and individual factors of sexually offending youth. It was demonstrated that there may be an association between community factors and the levels of risk for reoffending. Specifically, this study

indicated that sexually offending youth in the region of TN had a greater number of risk factors than did youth from CO and GVA suggesting that community characteristics may be associated with problem behaviours and other difficulties. These findings highlight the need to consider youth as inseparable from their community and social context. Intervention strategies that target the individual characteristics of offenders are essential. However, if factors associated with risk at the community level are ignored, then individual offenders are not likely to reach their potential for success in treatment. Further, if the goal is to treat offenders within their communities or to safely return the offender to his community, it is necessary to treat the individual, as well as to prepare him to return to the community that might be associated with his risk for reoffending. Therefore, intervention modalities that treat individual offenders as well as their immediate social ecologies should prove to be a successful strategy for treating sexually offending youth.

References

- Abel, G. G., Becker, J. V., Mittelman, M., & Cunningham-Rathner, J. (1987). Self-reported sex crimes of nonincarcerated paraphiliacs. *Journal of Interpersonal Violence*, 2(1), 3-25.
- Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behaviour*, 17(1), 19-52.
- Bourke, M. L., & Donohue, B. (1996). Assessment and treatment of juvenile sex offenders: An empirical review. *Journal of Child Sexual Abuse*, 5(1), 47-70.
- Cicchetti, D., Bronen, R., Spencer, S., Haut, S., Berg, A., Oliver, P., & Tyrer, P. (2006). Rating scales, scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. *Journal of Nervous and Mental Disease*, 194(8), 557-564.
- Christie, N., Andenaes, J., & Skirbekk. S. (1965). A study of self-reported crime. *Scandinavian Studies in Criminology*, 1, 86-116.
- Efta-Breitbach, J., & Freeman, K. A. (2004). Recidivism and resilience in juvenile sexual offenders: An analysis of the literature. *Journal of Child Sexual Abuse*, 13(3), 257-279.
- Elliott, D. M., & Briere, J. (1994). Forensic sexual abuse evaluations of older children: Disclosures and symptomatology. *Behavioural Sciences & The Law*, 12(3), 261-277.
- Farrington, D. P., Sampson, R. J., & Wikstrom, P. H. (1993). *Integrating individual and ecological aspects of crime*. Stockholm, Sweden: Swedish National Council for Crime Prevention.
- Finkelhor, D., Hotaling, G., Lewis, I. A., & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: Prevalence, characteristics, and risk factors. *Child Abuse & Neglect*, 14(1), 19-28.

- Fitzgerald, R., Wisener, M., & Savoie, J. (2004). Neighbourhood characteristics and distribution of crime in Canada. Statistics Canada Catalogue no. 85-561-MWE2004004. Retrieved February 26, 2008 from <http://www.statcan.ca/english/freepub/85-002-XIE/85-002-XIE2007003.pdf>
- Francisco, J., & Chenier, C. (2007). *A comparison of large urban, small urban and rural crime rates, 2005*. Statistics Canada Catalogue no. 92-591-XWE, vol. 27, no. 3. Retrieved February 26, 2008 from <http://www.statcan.ca/english/freepub/85-002-XIE/85-002-XIE2007003.pdf>
- Gottfredson, D. C., McNeil, R. J., III, & Gottfredson, G. D. (1991). Social area influences on delinquency: A multilevel analysis. *Journal of Research in Crime and Delinquency*, 28(2), 197-226.
- Gumpert, J., & Saltman, J. E. (1998). Social group work practice in rural areas: The practitioners speak. *Social Work With Groups*, 21(3), 19-34.
- Hall, G. C. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology*, 63(5), 802-809.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1998). Appraisal and management of risk in sexual aggressors: Implications for criminal justice policy. *Psychology, Public Policy, and Law*, 4(1), 73-115.
- Hartnagel, T. F., & Lee, W. (1990). Urban crime in Canada. *Canadian Journal of Criminology*, 32, 591-606.
- Jacob, J. C. (2006). Male and female youth crime in Canadian communities: Assessing the applicability of social disorganization theory. *Canadian Journal of Criminology and Criminal Justice/Revue Canadienne de Criminologie et de justice Penale*, 48(1), 31-60.

- Jennissen, T. (1992). *Health issues in rural Canada*. Retrieved February 26, 2008 from <http://dsp-psd.pwgsc.gc.ca/Collection-R/LoPBdP/BP/bp325-e.htm#AVAILABILITYtxt>
- Leonard, T. (1997). *Crime in major metropolitan areas, 1991-1995*. Statistics Canada Catalogue no. 85-002, vol. 17, no. 5. Retrieved February 26, 2008 from <http://www.statcan.ca/english/freepub/85-002-XIE/85-002-XIE2007003.pdf>
- Loeber, R., & Farrington, D. P. (2000). Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Development and Psychopathology*, 12(4), 737-762.
- McCarthy, B. R. (1991). Social structure, crime, and social control: An examination of factors influencing rates and probabilities of arrest. *Journal of Criminal Justice*, 19(1), 19-30.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30-46.
- Moore, T., Franey, K. C., & Geffner, R. (2004). Introduction: Assessment and treatment of youth who sexually offend: An overview. *Journal of Child Sexual Abuse*, 13(3), 1-13.
- Osgood, D. W., & Chambers, J. M. (2000). Social disorganization outside the metropolis: An analysis of rural violence. *Criminology*, 38(1), 81-115.
- Ouimet, M. (2000). Aggregation bias in ecological research: How social disorganization and criminal opportunities shape the spatial distribution of juvenile delinquency in Montreal. *Canadian Journal of Criminology* 42(2), 135-156.
- Ouimet, M. (1999). Crime in Canada and in the United States: A comparative analysis. *La Revue Canadienne de Sociologie et d'Anthropologie/The Canadian Review of Sociology and Anthropology*, 36(3), 389-408.

- Prentky, R., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol-II (J-SOAP-II) manual*. Washington, District of Columbia, US: US Department of Justice, Office of Justice Programs; Office of Juvenile Justice and Delinquency Prevention.
- Reiss, A. J., Jr. (1986). Why are communities important in understanding crime? *Crime and Justice*, 8, 1-33.
- Righthand, S., & Welch, C. (2004). Characteristics of youth who sexually offend. *Journal of Child Sexual Abuse*, 13(3), 15-32.
- Sampson, R. J. (1997). Collective regulation of adolescent misbehaviour: Validation results from eighty Chicago neighbourhoods. *Journal of Adolescent Research*, 12(2), 227-244.
- Sampson, R. J. (1986). The effects of urbanization and neighbourhood characteristics on criminal victimization. In R. M. Figlio, S. Hakim, & G. F. Rengert (Eds.), *Metropolitan Crime Patterns* (pp. 3-25). New York: Criminal Justice Press.
- Schoenfeld, T. M., Brown, J. E., Woodworth, M. & Gretton, H. (2008). *A comparison of the Juvenile Sex Offender Assessment Protocol-II (JSOAP-II) and the Estimate of Risk of Adolescent Sexual Offence Recidivism-II (ERASOR-II)*. Manuscript in preparation.
- Schulenberg, J. L. (2003). The social context of police discretion with young offenders: An ecological analysis. *Canadian Journal of Criminology and Criminal Justice/Revue Canadienne de Criminologie et de justice Penale*, 45(2), 127-157.
- Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. Chicago, IL, US: University of Chicago Press.
- Smiley, T. (2007). *Youth who sexually offend: Comparing treatment populations in British Columbia*. Unpublished master's thesis, University of Northern British Columbia, Prince George, British Columbia, Canada.

- Statistics Canada. (2007a). *Central Okanagan, British Columbia* (table). *2006 Community Profiles*. 2006 Census. Statistics Canada Catalogue no. 92-591-XWE. Ottawa.
Released March 13, 2007. Retrieved December 18, 2007 from
<http://www12.statcan.ca/english/census06/data/profiles/community/Index.cfm?Lang=E>
- Statistics Canada. (2007b). *Greater Vancouver, British Columbia* (table). *2006 Community Profiles*. 2006 Census. Statistics Canada Catalogue no. 92-591-XWE. Ottawa.
Released March 13, 2007. Retrieved December 18, 2007 from
<http://www12.statcan.ca/english/census06/data/profiles/community/Index.cfm?Lang=E>
- Statistics Canada. (2007c). *Thompson-Nicola, British Columbia* (table). *2006 Community Profiles*. 2006 Census. Statistics Canada Catalogue no. 92-591-XWE. Ottawa.
Released March 13, 2007. Retrieved December 18, 2007 from
<http://www12.statcan.ca/english/census06/data/profiles/community/Index.cfm?Lang=E>
- Van Dijk, J. M. (1999). The experience of crime and justice. In G. Newman (Ed.), *Global report in crime and justice* (pp. 25-42). New York: Oxford University Press.
- Weisheit, R. A., Falcone, D. N., & Wells, L. E. (1999). *Crime and policing in rural and small-town America* (2nd ed.). Prospect Heights, IL: Waveland Press.
- Wells, L. E., & Weisheit, R. A. (2004). Patterns of rural and urban crime: A county-level comparison. *Criminal Justice Review*, 29(1), 1-22.
- Welsh, B. C., & Farrington, D. P. (2007). Scientific support for early prevention of delinquency and later offending. *Victims & Offenders*, 2(2), 125-140.
- Wikstrom, P. H. (1998). Communities and crime. In M. Tonry (Ed.), *The handbook of crime and punishment* (pp. 269-301). New York: Oxford University Press.
- Wikstrom, P. H. (1991). *Urban crime, criminals and victims*. New York: Springer-Verlag.

Wikstrom, P. H., & Loeber, R. (2000). Do disadvantaged neighbourhoods cause well-adjusted children to become adolescent delinquents? A study of male juvenile serious offending, individual risk and protective factors, and neighbourhood context. *Criminology*, 38(4), 1109-1142.

Wirth, L. (1938). Urbanization as a way of life. *The American Journal of Sociology*, 44, 1-24.

4 CONCLUSION: RECOMMENDATIONS FOR PRACTICE

Research on sexual offending has traditionally focused on identifying, evaluating, and reducing the risk posed by sexual offenders (e.g., Efta-Breitbach & Freeman, 2004; Righthand & Welch, 2004). The results of these research endeavours have led to a relatively comprehensive understanding of the factors that lead one to initiate and to maintain this criminal behaviour. The common agreement that there are individual factors associated with sexual offending has led to the development of standardized approaches for systematically evaluating offender risk levels. However, ecological researchers have long proposed that criminal behaviour is influenced by macro structures and that individuals may be products of their social and community contexts (Sampson & Groves, 1989; Schulenberg, 2003; Shaw & McKay, 1942). This study sought to add to the literature concerning sexually offending youth in two ways. The first goal was to evaluate two risk assessment instruments developed specifically for sexually offending youth. The second goal was to provide a multilevel analysis of sexually offending youth, to assess the potentially inextricable relationship between the community contexts of young offenders and individual risk factors that influence sexual offending behaviour. The results of this research can be used to help provide assessment and treatment techniques that are empirically supported and are effective in reducing the potential for future harm to the community.

Risk assessments have been shown to be valuable tools for providing suitable interventions for young sexual offenders. To treat these youth justly, and to protect the safety of the community, important decisions must be made about the level of supervision required and the amount of time needed for rehabilitation. In addition, considering the goal is to rehabilitate youth rather than simply incarcerate, clinicians have a responsibility to provide treatment that is suitable for individual offenders and that is effective at reducing recidivism. Risk assessment

instruments have been shown to play a vital role in assisting justice professionals and treatment planners with these difficult decisions. Unfortunately, few standardized assessment instruments have been developed for this population. Given the invaluable contribution these tools can have in managing sexually offending youth, this study further evaluated two risk assessment instruments that are widely used with this population and that have the potential to be valid measures of risk.

This study found that compared to the ERASOR-II, the J-SOAP-II was a more reliable and consistent measure. The J-SOAP-II demonstrated good to excellent interrater reliability, had internal consistency for some of the items, and some statistical justification for the scales. Conversely, fewer of the items on the ERASOR-II were scored alike among raters, there was less consistency among the items, and little statistical justification for the scales. Moreover, the clinical rating of the ERASOR-II was not as reliable as the J-SOAP-II *proportion of risk rating* or the ERASOR-II *total score*, probably because it was heavily reliant on clinical judgment. Despite this, however, the J-SOAP-II and the ERASOR-II concluded a similar level of risk for each offender ($r_s = .74$).

These results suggest that if deciding between the J-SOAP-II and the ERASOR-II, the J-SOAP-II is a better choice. However, the J-SOAP-II was not without its limitations. For instance, although there was excellent agreement between raters on the overall risk rating ($r = .96$), scoring among raters on individual items was not as consistent. In addition, although the scale alphas indicated good internal consistency of the items, individual item-total correlations (ITCs) and scale alphas were far from perfect. Interestingly, some of the items with low ITCs were items that had been recently added to this newest version of the instrument (e.g., *sexualized aggression* and *sexual victimization history*). Finally, although the confirmatory factor analysis

indicated that the items and the scales of the J-SOAP-II were a better fit than the ERASOR-II, this model failed to reach significance. This was mainly attributed to an insignificant relationship between the *sexual drive/preoccupation scale* and the three other scales of the instrument.

This evaluation is a valuable contribution to the study of risk assessments for sexually offending youth because it assesses two newly developed instruments in need of further validation. In addition, the results of this study can have a number of implications for clinicians who assess the risk of sexually offending youth to predict future violence or use risk assessments to guide treatment planning. Although it was beyond the scope of this study to test the predictive validity of these instruments, our results suggest that it is premature to use either of these instruments as a predictive measure. This is because neither of the instruments had excellent interrater reliability for all of its items. Further, there were many items on both of these instruments that did not contribute well to the overall risk rating. For instance, 3 of the 28 items of the J-SOAP-II and 10 of the 25 items on the ERASOR-II had low ITCs raising concern about their inclusion. Given that conclusions about risk can guide decisions about the level of sentencing and degree of supervision required, statements regarding risk to reoffend should be highly consistent and should be based on a number of factors that are relevant to risk. Basing risk ratings on factors that do not ultimately contribute to the final score is unnecessary and it could mislead the assessor into believing that all of the items are equally important in formulating the risk score.

While these risk assessments should not yet be used as a predictive measure of risk, these results suggest that a systematic evaluation of risk factors may result in relatively consistent ratings between assessors. For instance, although the *final risk rating* of the ERASOR-II was not as reliable as the J-SOAP-II *proportion of risk rating*, these scores were highly correlated with

each other. Reviewing empirically supported risk factors may, therefore, result in more consistent ratings of risk even if a number of the individual items are not reliable. In addition, basing risk ratings on individual risk factors may improve the reliability of risk ratings using clinical judgment. However, risk ratings based on clinical judgement were still not as reliable as the ratings of the J-SOAP-II that were derived from tallying scale scores. Therefore, while the J-SOAP-II and the ERASOR-II concluded similar levels of risk, for the most reliable risk ratings it is still advisable to use a more structured approach of tallying scores, as required by the J-SOAP-II.

Both the J-SOAP-II and the ERASOR-II were designed to measure risk for committing future sexual crimes. However, many of the items on these instruments concern nonsexual offending and other psychosocial functioning because these factors have been suggested in the literature to be associated with sexual recidivism (e.g., Efta-Breitbach & Freeman, 2004). Interestingly, this study indicated that the scale on the J-SOAP-II and on the ERASOR-II that included items pertaining to sexual offending specifically (e.g., *number of sexual abuse victims*, *duration of sex offence history*) were not correlated with all of the other scales. If all of the items on an instrument are supposed to be measuring the same underlying construct of risk, it would be expected that all of the scales would be correlated with each other. These results suggest that items pertaining to sexual and nonsexual offending may not be as interrelated as previously thought. Nevertheless, research has shown that many sexual offenders who reoffend do so with nonsexual crimes (Righthand & Welch, 2001). Accordingly, these instruments can still be used to identify areas to target in treatment, which could ultimately result in the reduction of risk. However, it is problematic to conclude a level of risk based on items that are uncorrelated with

each other because these items may potentially be measuring two very different underlying constructs.

Although these instruments are not valid measures of risk and some of the psychometric properties are limited, both the J-SOAP-II and the ERASOR-II could still help provide informative assessments for targeted treatment plans. For instance, the systematic evaluation of empirically derived items can provide information about an offender's history, behavioural problems, and other psychosocial difficulties that could be targeted and treated. For instance, the criminally entrenched offender who has a number of delinquent peers and a history of disruptive and problematic behaviours is in need of a very different treatment program than the offender who does not have a history of problematic behaviours and has strong positive social networks. Therefore, despite their limitations, the J-SOAP-II and the ERASOR-II could inform clinicians about each offender's unique service needs allowing for tailored and potentially more effective treatment, which could ultimately result in lowered risk for individual offenders and their communities.

To accurately use the J-SOAP-II and the ERASOR-II for treatment planning it is important that clinicians are first aware of the limitations of the instruments. One such limitation is the lack of scale consistency, particularly for the ERASOR-II. Although the J-SOAP-II had three items that did not contribute well to their respective scale, all scale alphas were good to excellent indicating adequate categorization of the items. The majority of the items on the ERASOR-II, conversely, contributed little to their respective scales. In addition, while the scale alphas ranged from poor to excellent, confirmatory factor analysis failed to yield support for the five scales of the instrument. These results suggest that while there is some statistical justification for the scales of the J-SOAP-II, there is little for the ERASOR-II. As such, it would

be highly inappropriate to describe an offender in terms of the ERASOR-II scales (e.g., offender was high risk in the area of family/environment functioning) because there is no rationale for the categorization of the items in this manner. Rather a more appropriate strategy may be to use the scales of the J-SOAP-II and the ERASOR-II simply as informal guides for structuring clinical assessment interviews. As mentioned above, both of these instruments contain valuable information for devising targeted treatment plans. However, it would be inappropriate to assess risk according to the predetermined categories of the ERASOR-II and to avoid misuse of this instrument the manual should make this perfectly clear.

Although both the J-SOAP-II and the ERASOR-II have some utility as treatment planning tools, they are both in need of further testing and potentially revisions before they can be used entirely as intended. This is particularly necessary given that the results of this study contradicted previously reported evaluations of both the J-SOAP-II and the ERASOR-II. Specifically, this study found that the interrater reliability and internal consistency of the items on the J-SOAP-II and the ERASOR-II were not as strong as reported by Righthand et al. (2005) and Worling (2004). Moreover, the J-SOAP-II has undergone revisions since its last published evaluation and this is believed to be the first study to evaluate this new version of the instrument. This study found that some of the modifications from version one to version two resulted in improvements in interrater reliability or internal consistency (e.g., *sexual victimization history*), whereas others were worsened (e.g., *prior legally charged sex offences*). These inconsistent results highlight the need for further testing of the psychometric properties of these instruments.

Standardized risk assessment instruments are lacking for sexually offending youth. Given that the assessment of risk is the first step in providing offenders with appropriate and effective services, it is especially necessary that youth who commit sexual crimes also receive this

advantage. Thus, risk assessment instruments that have demonstrated some promise with this population should be repeatedly evaluated and revised until consistent results are produced. Surprisingly, there is very little research evaluating the predictive validity of these instruments. This is largely attributed to the small number of offenders who actually reoffend sexually and the difficulties in tracking these offenders. Nevertheless, if these instruments are ever to be used as accurate predictors of recidivism, longitudinal testing of their predictive validity is essential. Moreover, while psychometric testing has been conducted on both the J-SOAP-II and the ERASOR-II, these studies have yet to generate consistent results and have been disadvantaged by small sample sizes. Future replication studies with large sample sizes could potentially improve confidence in these instruments. Despite the limitations of these instruments, however, they mark a crucial first step in providing sexually offending youth with appropriate and effective services.

This study also conducted a multilevel assessment of sexually offending youth by considering both individual-level and ecological perspectives on crime. It was found that sexually offending youth in the Thompson Nicola (TN) region were generally had a greater number of risk factors than did youth in the Central Okanagan (CO) and Greater Vancouver Area (GVA). Although there were no regional differences among sexually offending behaviours and histories, TN youth had more risk factors pertaining to the areas of general offending and other behavioural difficulties, intervention, and family/environment functioning. These findings have a number of important implications for devising and delivering effective treatment modalities for youth who commit sexual crimes.

The Youth Sex Offender Treatment Program (YSOTP) offered by Youth Forensic Psychiatric Services (YFPS) is an empirically supported and highly structured treatment program

for youth who commit sexual crimes. Like most empirically driven treatment modalities for this population (Borduin & Schaeffer, 2001), it offers sex education as well as cognitive behavioural therapy to reduce denial and increase accountability for the sex offence, increase remorse and empathy for the victim, and modify cognitive distortions and deviant arousal patterns. The primary focus of treatment is on sexual offending behaviours, and the goal is to adequately reduce the risk of these offenders before probation is completed. YFPS offers the YSOTP in the regions of TN, CO, and GVA.

The results of this study suggest that a standardized treatment program may not be equally suitable for all offenders in every region. For instance, it was found that while there were no regional differences in sexually offending behaviours *per se*, youth in TN were participating in more antisocial and delinquent behaviours. For example, TN youth had poor quality of peer relationships, participated in more antisocial behaviours, and had more school difficulties than did youth in CO and GVA. A treatment program in this region that focuses solely on sexual offending behaviours and attitudes may be neglecting to treat other behavioural difficulties that could heighten the risk for recidivism. This is especially problematic considering that sexually offending youth are more likely to reoffend by committing nonsexual rather than sexual crimes (Righthand & Welch, 2001). Therefore, to adequately reduce the risk of TN youth, it would be advisable to match the treatment program to the needs of the offenders there, and target nonsexual as well as sexual behaviours. Targeting a wider range of problem behaviours in treatment may result in better treatment outcomes and subsequently reduce the threat of future harm to the community.

The need to treat a variety of problem behaviours is further supported by the regional differences found in the area of intervention. Specifically, TN youth scored the highest of all

regions in the area of intervention, suggesting that these youth were high risk in this area at the beginning of treatment and were resistant to change. This scale assesses beliefs about all types of offending behaviour, such as cognitive distortions, degree of remorse and guilt, and motivation for change. If youth are entering treatment with lengthy histories of generally problematic behaviours, it is likely that they will also present with generally antisocial beliefs and attitudes. Negative attitudes and resistance to change have the potential to impede treatment success. In addition, if clinicians are unaware of the root of these antisocial attitudes because they have ignored a variety of problem behaviours associated with offending, then these attitudes are not likely to be understood or improved. Treatment programs that neglect to consider a variety of determinants for offending behaviour may be unknowingly ignoring variables that are essential to positive treatment outcomes.

Adequate resources are needed to provide offenders with comprehensive treatment programs that target all appropriate behavioural difficulties. Unfortunately, there is generally a lack of resources, such as few qualified professionals and less money for services, in small communities (Gumpert & Saltman, 1998; Jennissen, 1992; Smiley, 2007). The results of this study demonstrate that it is problematical to deprive these small communities of resources for youth. A better distribution of resources at the community level could help prevent the initiation of sexual offending and reduce the potential for future offending. This study found that youth entering treatment in the TN region had a number of behavioural difficulties, in addition to sexual offending. Had preventative strategies been employed at the first sign of behavioural problems, many of these offenders may not have escalated to sexual offending. With more resources, community organizations such as schools and services for children and families, could potentially prevent offending behaviour or at least reduce the risk levels of offenders. Further,

this study suggests that the YFPS clinic in TN may need more resources for intervention services. If youth entering the treatment facility are higher risk than youth in CO and GVA, resources should be distributed accordingly. That is, the YFPS in TN could potentially benefit from more clinicians to design and implement targeted treatment plans and money to research and develop programs that are suitable for the offenders there. This study shows how the assessment of risk can be a valuable method for making decisions regarding resource distribution at the community level.

Overall, this study found evidence that sexually offending youth are influenced by their social and community contexts. Whether these communities act as high risk social contexts that perpetuate offending behaviour or as protective environments that deter criminal offending, it is evident that there are multiple determinants for offending both at the individual and community level. As such, to be effective in reducing risk, it is essential that treatment modalities consider that individual offenders are interrelated with their families, peers, schools, and neighbourhoods. Treatment strategies that simultaneously incorporate all relevant aspects of an offender's environment, including individual and social factors, ought to reduce the likelihood of future offending more effectively.

One such treatment modality that has demonstrated success with sexually offending youth is Multisystemic Therapy (MST; Borduin & Henggeler, 1990; Henggeler & Borduin, 1990). MST is an ecologically-based treatment model that simultaneously addresses multiple determinants of serious offending behaviour among adolescents. The basic premise of this therapy is that offenders and their families are imbedded within multiple systems, and it is the interplay of these systems that can influence or impede treatment success (Borduin & Schaeffer, 2001). Using well-validated treatment strategies, MST addresses intrapersonal (e.g., cognitive),

familial, and extrafamilial (e.g., peers, school, neighbourhood) factors by treating offenders as inseparable from their communities. It has been argued that MST is more effective than traditional individual orientated therapies because it addresses known correlates of sexual offending in a comprehensive fashion and it supports healthy adaptation and reduces risk for reoffending by improving the social ecologies of young offenders (Swenson, Henggeler, Schoenwald, Kaufman, & Randall, 1998). This study supports the use of a treatment modality like MST that simultaneously addresses individual and community factors for sexually offending youth.

This study addressed two areas of research that can inform about the service needs of sexually offending youth. It was demonstrated that risk assessments designed for youth who commit sexual offences are much in need of further evaluation, but currently used tools have shown some promising results. This study also found that there is an inseparable relationship between individual offenders and their community contexts, supporting the need for treatment modalities that combine individual and ecological factors. To treat sexual offenders justly and effectively, while protecting the community from future harm, it is necessary to further evaluate assessment and intervention strategies that have shown some potential to reduce the risk of these young offenders.

References

- Borduin, C. M., & Henggeler, S. W. (1990). A multisystemic approach to the treatment of serious delinquent behaviour. In R. J. McMahon & R. Dev. Peter (Eds.), *Behavior disorders of adolescence: Research, intervention, and policy in clinical and school settings* (pp. 62-80). New York: Plenum.
- Borduin, C. M., & Schaeffer, C. M. (2001). Multisystemic treatment of juvenile sexual offenders: A progress report. *Journal of Psychology & Human Sexuality*, 13(3), 25-42.
- Efta-Breitbach, J., & Freeman, K. A. (2004). Recidivism and resilience in juvenile sexual offenders: An analysis of the literature. *Journal of Child Sexual Abuse*, 13(3), 257-279.
- Gumpert, J., & Saltman, J. E. (1998). Social group work practice in rural areas: The practitioners speak. *Social Work With Groups*, 21(3), 19-34.
- Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond: A mutisystemic approach to treating the behaviour problems of children and adolescents*. Pacific Grove, CA: Brooks/Cole.
- Jennissen, T. (1992). *Health issues in rural Canada*. Retrieved February 26, 2008 from <http://dsp-psd.pwgsc.gc.ca/Collection-R/LoPBdP/BP/bp325-e.htm#AVAILABILITYtxt>
- Righthand, S., Prentky, R., Knight, R., Carpenter, E., Hecker, J. E., & Nangle, D. (2005). Factor structure and validation of the Juvenile Sex Offender Assessment Protocol (J-SOAP). *Sexual Abuse: Journal of Research and Treatment*, 17(1), 13-30.
- Righthand, S., & Welch, C. (2004). Characteristics of youth who sexually offend. *Journal of Child Sexual Abuse*, 13(3), 15-32.

- Righthand, S., & Welch, C. (2001). *Youths who have sexually offended: A review of the professional literature*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.
- Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, 94(4), 774-802.
- Schulenberg, J. L. (2003). The social context of police discretion with young offenders: An ecological analysis. *Canadian Journal of Criminology and Criminal Justice/Revue Canadienne de Criminologie et de justice Penale*, 45(2), 127-157.
- Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency and urban areas*. Chicago, IL, US: University of Chicago Press.
- Smiley, T. (2007). *Youth who sexually offend: Comparing treatment populations in British Columbia*. Unpublished master's thesis, University of Northern British Columbia, Prince George, British Columbia, Canada.
- Swenson, C. C., Henggeler, S. W., Schoenwald, S. K., Kaufman, K. L., & Randall, J. (1998). Changing the social ecologies of adolescent sexual offenders: Implications of the success of multisystemic therapy in treating serious antisocial behaviour in adolescents. *Child Maltreatment*, 3(4), 330-338.
- Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR): Preliminary psychometric data. *Sexual Abuse: Journal of Research and Treatment*, 16(3), 235-254.

APPENDIX



The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road,
Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - FULL BOARD

PRINCIPAL INVESTIGATOR: Michael W. Woodworth	INSTITUTION / DEPARTMENT: UBC/UBCO IKE Barber School of Arts & Sc/UBCO Admin Unit 4 Arts & Sci	UBC BREB NUMBER: H06-04063
INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:		
Institution	Site	
N/A	N/A	
Other locations where the research will be conducted: Youth Forensic Psychiatric Services (YFPS), Kelowna and Burnaby		
CO-INVESTIGATOR(S): Tara Schoenfeld		
SPONSORING AGENCIES: N/A		
PROJECT TITLE: Risky business: A regional comparison of juvenile sex offenders' risk and service needs.		
REB MEETING DATE: March 8, 2007	CERTIFICATE EXPIRY DATE: March 8, 2008	
DOCUMENTS INCLUDED IN THIS APPROVAL:		DATE APPROVED: March 12, 2007
Document Name	Version	Date
<u>Questionnaire, Questionnaire Cover Letter, Tests:</u>		
ERASOR-II Manual	N/A	February 1, 2001
Demographics	N/A	January 11, 2007
J-SOAP-II Manual	N/A	January 1, 2003
<u>Other Documents:</u>		
YFPS Letter of Support	N/A	January 4, 2007
The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.		
<p>Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:</p> <div style="border-top: 1px solid black; width: 60%; margin: 0 auto; text-align: center;"> Dr. Peter Suedfeld, Chair Dr. Jim Rupert, Associate Chair Dr. Arminee Kazanjian, Associate Chair Dr. M. Judith Lynam, Associate Chair </div>		