

**The Proceedings of the 2007
North American Correctional &
Criminal Justice Psychology
Conference**

2008-02

Editors:

Guy Bourgon, R. Karl Hanson, Joanna D. Pozzulo,
Kelly E. Morton Bourgon, and Carrie L. Tanasichuk

© Her Majesty the Queen in Right of Canada, 2008

Cat. No.: PS3-1/2008-2E-PDF

ISBN No.: 978-0-662-48400-4

**THE PROCEEDINGS OF THE
2007 NORTH AMERICAN CORRECTIONAL &
CRIMINAL JUSTICE PSYCHOLOGY CONFERENCE**

NACCJPC

2008

Ottawa, Ontario



Editors:

Guy Bourgon, R. Karl Hanson, Joanna D. Pozzulo,
Kelly E. Morton Bourgon, & Carrie L. Tanasichuk.

Public Safety Canada

www.publicsafety.gc.ca



Criminal Justice Section of the
Canadian Psychological
Association (CPA)



Criminal Justice Section of
Division 18 of the American
Psychological Association (APA)

Introductory Comments to the Proceedings of the North American Correctional & Criminal Justice Psychology Conference

Correctional and Criminal Justice is a specialty area in psychology. We bring our skills to a specific and unique group of clients – often an underserved client group. It is not simply about taking what works for the general population and putting it to work within a correctional setting. It is about adapting psychology to the client, adapting to the organizations that we work for, and adapting to the systems that we serve with the goal to protect society and to improve the quality of life of individuals who are far too often marginalized.

We need to constantly and consistently frame what we do as a service to public safety. Society's demand for accountability too often translates into punishment. This makes the assessment portion of what we do palatable and popular as it too often focuses on the punitive while at the same time making the treatment portion of what we do optional and suspect. The message we need to repeat is a simple one – when our clients get better – society becomes safer. Treatment and intervention are not options: they are essential to public safety and a cornerstone to Correctional and Criminal Justice psychology.

As a specialty area in Psychology, we need ongoing training and education that meet our specific needs. Cooperative efforts between our respective Criminal Justice Sections such as the NACCJPC, will serve to meet the needs of practitioners who daily face the challenge of a practice behind the wall or fence of a correctional institution. My hope is that this conference may serve as a seed for future cooperative efforts that may take many forms in the years ahead.

The conference was a tremendous success with approximately 350 attendees and featuring well over 200 presentations. A number of those presenters have provided a synopsis of their presentations in these Proceedings, covering a wide area of topics of interest to the correctional and criminal justice psychologist.

I am very grateful to Dr. Guy Bourgon and his editorial volunteers R. Karl Hanson, Joanna D. Pozzulo, Kelly E. Morton Bourgon, & Carrie L. Tanasichuk for their hard work in editing this volume. They stayed long after the party was over to provide us all with a permanent record of some of the excellent presentations made at the conference.

Thank-you very much.

Dr. Jeremy Mills, Ph.D., C.Psych.

Chair - NACCJPC

Table of Contents

Section A: Challenges of Correctional & Criminal Psychology	1
Systemic Issues and Correctional Outcomes: Expanding the Scope of Correctional Psychology	2
<i>C.B. Clements, R. Althouse, R.J. Ax, P.R. Magaletta, T.J. Fagan, & J. S. Wormith</i>	
Extensions of the Risk-Need-Responsivity (RNR) Model of Assessment and Correctional Treatment	7
<i>D.A. Andrews</i>	
Structured Guidelines for Evaluating Study Quality	12
<i>L. Helmus</i>	
Sexual Conflict and Coercion	16
<i>V. Quinsey</i>	
Section B: Risk Assessment	19
Current status of Violence Risk Assessment: Is there a role for Clinical Judgment?	20
<i>M. Rice</i>	
Gender Comparisons on the Self-Appraisal Questionnaire (SAQ): A Tool for Assessing the Risk of Violent and Non-Violent Recidivism	24
<i>T. Hematti</i>	
Assessing the Predictive Validity of the Youth Level of Service/Case Management Inventory	27
<i>K. Bechtel, C.T. Lowenkamp, & E. Latessa</i>	
Predicting Criminal Recidivism in Adult Male Offenders: A Four Wave Prospective Study	33
<i>S.L. Brown & E. Zamble</i>	
Adapting the Sexual Offender Need Assessment Rating into a Self-Report Measure (SONAR-SR)	38
<i>M. Chajewski & K.A. Markus</i>	
Section C: Interventions	44
Risk, Need, and Responsivity: A Heuristic for Evaluating the “Quality” of Offender Interventions	45
<i>G. Bourgon, R.K. Hanson, & J. Bonta</i>	
Does Treatment Make Psychopaths Worse? A Meta-Analytic Review	50
<i>C.L. Tanasichuk & J.S. Wormith</i>	
A Psychologically Informed Meta-Analysis of Sex Offender Treatment Outcome Studies	55
<i>R.K. Hanson & G. Bourgon</i>	
The Vermont Cognitive Self-Change Program: The Case for Risk-Adjusted Classification	58
<i>C.M. Sadler & T.A. Powell</i>	
Prediction of Program Dropout in a High-Intensity Violent Offender Program in Canada	63
<i>R. Gobeil & R. Serin</i>	
Evaluating Sex Offender Treatment Efficacy: A Review of Research Findings from a High Intensity Sex Offender Program	68
<i>M.E. Olver & S.C.P. Wong</i>	
Care, Control, and Mental Disorder: Comparing Practices and Outcomes in Prototypic Specialty vs. Traditional Probation	73
<i>S.M. Manchak, J.L. Skeem, & S. Vidal</i>	

Section D: Police & Court Psychology	76
Clinical Versus Actuarial Geographic Profiling Approaches: A Meta-Analysis.....	77
<i>K.B. Emeno, C. Bennell, P.J. Taylor & B. Snook</i>	
Exploring the <i>CSI Effect</i> : Is it Real? If so, What is it?.....	81
<i>S.M. Smith, M.W. Patry, & V. Stinson</i>	
The <i>CSI Effect</i> : An Examination of the Source of Biases toward Forensic Science.....	85
<i>L.J. Hanby & C. Bennell</i>	
Blurring the Line between Fact and Fiction: Expert Opinions about Forensic Investigation Tools Represented on <i>CSI</i>	94
<i>M. W. Patry, S.M. Smith, V. Stinson, & T. McCulloch</i>	
Linking Serial Rapes: A Test of the Behavioural Frequency Hypothesis.....	98
<i>T. Melnyk, C. Bennell, & N.J. Jones</i>	
The Influence of Witness Age, Relation to Crime, and Eyewitness Identification Decision on Jurors' Perceptions and Verdicts	102
<i>J.L. Dempsey & J.D. Pozzulo</i>	
Preschoolers' Person Description and Identification Accuracy: A Comparison of the Simultaneous and Elimination Lineup Procedures	106
<i>C. Charmagne & J.D. Pozzulo</i>	
Section E: Professional & Ethical Issues	109
Ethical Dilemmas in Practicing Correctional Psychology	110
<i>J. McGuire</i>	
Ethical Dilemmas, Forensic Psychology and Therapeutic Jurisprudence.....	118
<i>I. Dickie</i>	
Training Needs for Graduate Students in Forensic Psychology	121
<i>B. Reynolds, A. Reinhardt, & I. Dickie</i>	
A Distance Learning Course on Mental Health Issues in Jails and Prisons.....	125
<i>M. Perlin</i>	
Prison, Stress and the Aging Process	129
<i>P.J. Donovick, J.S. Portocarerro, R. Holtzer, & R.S. Yee</i>	
The Importance of Context: Probation Officers Working With Adolescent Sex Offenders.....	134
<i>T.D. Griep</i>	
Juvenile Sex Offending Within an Attachment Theory Framework.	137
<i>A.L. Szielasko, D.K. Symons, & J. Boutilier</i>	
Section F: Special Topics	144
Criminal Trajectories from Adolescence to Adulthood in an Ontario Sample of Offenders.....	143
<i>D. Day, I. Beve, T. Duchesne, J.S. Rosenthal, Y. Sun, & F. Theodor</i>	
Criminal Attitudes in Young Female Offenders: A Psychometric Evaluation	149
<i>L. Greiner, S. Brown, & V. Jennings</i>	

The Jesness Inventory-Revised (JI-R) As a Measure of Psychopathology in a Sample of Adjudicated Juvenile Offenders	153
<i>A. Eisenbuch, M.R. Schneider, D. Martin, J. Manapace, & P. Reube</i>	
What’s Not Working within ‘What Works’: Executive Cognitive Functioning Capacity of First Time Offenders, Return Offenders, and Controls	156
<i>E.H. Ross, J. Neil, & P.N.S. Hoaken</i>	
Autonomous Classification of Sexual Offenders from Eye Pattern Behavior	158
<i>S. Chartier, P. Renaud, & S. Caron</i>	
Is Sexual Assault Committed by Hebephiles Different from that Committed by Pedephiles and Rapists?	162
<i>S. Desjardins & L. Granger</i>	
Personal and Criminal Characteristics Distinguishing Pedophiles, Hebephiles, and Rapists.....	166
<i>S. Desjardins & L. Granger</i>	
An Empirical Taxonomy of Incarcerated Male Sexual Offenders Using Finite Mixture Modeling: Adult Victims	170
<i>J. D. Fargo</i>	
Restorative Justice: What Role Can Psychologists Play?	175
<i>T. Ruge</i>	

Section A

Challenges of Correctional & Criminal Justice Psychology

Systemic Issues and Correctional Outcomes: Expanding the Scope of Correctional Psychology

Carl B. Clements
University of Alabama
E-mail: cclement@as.ua.edu

Robert K. Ax
Midlothian, VA

Thomas J. Fagan
Nova Southeastern University

Richard Althouse
Wisconsin Department of Corrections

Phillip R. Magaletta
Federal Bureau of Prisons

J. Stephen Wormith
University of Saskatchewan

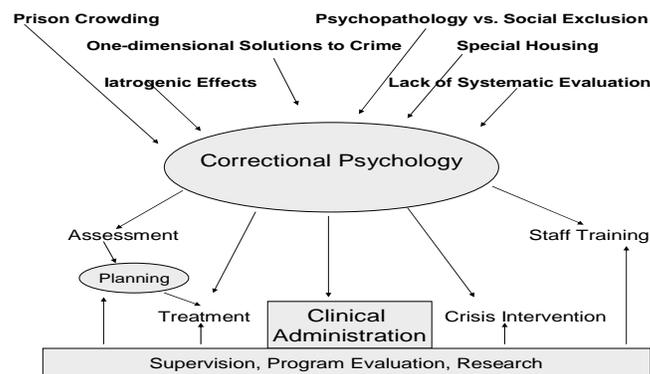
In this presentation we address issues that impact the delivery of correctional psychology services rather than the services themselves. While recognizing that the field is charged with such roles and functions as assessment, treatment, crisis intervention, staff training, planning, and program administration, we examine several systemic forces and conditions that may undermine the quality of these services. They include: prison crowding, unintended consequences, special housing, failure to address non-pathological factors in crime, and lack of systematic evaluation of prison practices. These relations are portrayed in the figure below.

A Systems Perspective

Represented by John F. Kennedy's quote: "Today's problems are the result of yesterday's solutions," a systems perspective acknowledges and examines the dynamic influence of interrelated components. Clearly, the justice

system is composed of many interacting forces. Some basic features of "systems" include the following: most identified "problems" are multiply caused but often addressed with singular solutions; policies based on single fix notions, of crime for example, often have unintended effects; such consequences are often delayed and negative. For example, the attempt to "reduce crime" via massive incarceration has given rise to the "Prison-Industrial Complex." Consequences of longer sentences include overcrowding, leading to fewer services and fewer offenders prepared for reintegration. "3-Strikes" laws and the "War on Drugs" have contributed to this cascade of effects.

Sentencing disparity may also give us a distorted picture of the offender population whom we are charged to manage. For example, minorities are overrepresented among those incarcerated for drug offenses despite the fact that they are less frequently arrested for these offenses. Poverty and limited opportunity also predict actual



incarceration. In addition to well-documented criminogenic risks and needs, it is likely that social exclusion has contributed to prison populations, particularly its disproportionate

minority representation. It is clear than single-cause, single-fix models (yesterday's solutions) create many negative consequences (today's problems).

Iatrogenic Effects

Systemic forces and their far-reaching consequences are mirrored by specific prison practices with potentially adverse effects. Unintended negative consequences are often identified in the medical context; the term “iatrogenic” translates “doctor-caused.” Prison-related examples, both medical and non-medical, have recently become more prominent. Although it has been argued that prisons inherently do harm, we raise the question of whether reducing *avoidable* harm should be a goal. Radical solutions have been proposed including the abolition of prisons. More realistic is the effort to identify and prioritize the goals of incarceration. As noted by our colleague Craig Haney,

“Contexts are shaped in part by the purposes they serve, and prison environments mandated to provide prisoners with opportunities that will facilitate their reintegration into the free world function differently from ones whose only purpose is to punish (p. 314).”

One way to address iatrogenic outcomes is to view them as dependent variables whose occurrence and causal forces are targets for research. We offer a three-part heuristic for classifying adverse outcomes: the creation or exacerbation of mental and/or physical illness or injury; the subversion of the rehabilitation mission of corrections; and the fostering of adverse societal effects. We recognize the need for more fine-grained analyses of the connections between specific practices and their effects (e.g., the role of individual differences and other moderating variables). In fact, we believe that such a program of research is needed.

With respect to *mental and physical consequences*, we need look no further than the high rates of victimization in prisons. Sexual assault occurs at alarming rates and is doubtless underreported. Victimization also has been associated with increased suicide rates which are comparatively high in US prisons. Its occurrence is often seen as imminently preventable despite a number of documented attitudinal barriers.

Subversion of the prison’s rehabilitation mission is another presumably unintended consequence. Failure to provide safe environments promotes a

survival mentality complete with recruitment to gangs whose anti-social values compete with the goals of reintegration. Likewise, failure to address the needs of offenders with mental disorders is also a source of negative outcomes. Given that 12-16% of US inmates are mentally ill and have a comparatively higher recidivism rate, the enrollment of such offenders in quality treatment programs is discouragingly low (e.g., 40-60%).

Secondary societal effects must also be included in any discussion of iatrogenesis. Examples include negative effects on inmates’ children, including school-related problems and peer relationships, and the comparatively higher suicide rate of correctional officers. Prisons also are potential sources of infectious disease, and inadequate treatment can threaten the community. In one example, “unacceptably low rates for the therapy completion” for inmate tuberculosis was seen as disturbing and putting communities at risk.

In the prison context, *iatrogenic* implies undue, avoidable harm—outcomes beyond the expected impact of correctional confinement. Prison life can have cumulative effects including depression, infectious disease, or more firmly entrenched antisocial attitudes. By virtue of training in research methods, assessment, and intervention, correctional psychologists are in a unique position to examine and perhaps mitigate or prevent iatrogenic outcomes.

The Experience of Segregation

One example of a widespread practice requiring further study is the use of “segregation” within correctional institutions. To date, the empirical evidence documenting the effects of segregation on offender functioning is inconclusive. Early studies lacked external validity. They often focused on tortured political and wartime prisoners, were based on sensory deprivation studies, or used volunteers and specified time limits. Although negative psychological states are often seen, particularly in the context of prison litigation, pre-existing measures are often absent. We recommend a three part conceptualization of segregation in order to examine which combinations of factors raise concerns about preventable harm and which adhere to the intent

of segregation without creating such harm. This model addresses person, context, and person-by-context.

Person or offender factors, which vary considerably, include mental health history and status, prior violence and impulsivity, personality factors, tolerance for reduced stimulation, and problem-solving skills. All have been demonstrated to impact an offender's response to segregation. Context variables reflect a similar diversity and include privacy, access to daylight, length of cell confinement per day, noise and overcrowding levels, staff functioning, physical layout, access to personal effects, programs and services, and recreational equipment and hobby items. Carried to the negative extreme, such factors may promote psychopathology and further behavioral dysfunction. Or they can be more benignly configured.

Combining person and context factors—and this is not beyond the scope of our research expertise—we suggest the question is properly, “What particular characteristics of offenders interact with which particular features of the segregation environment to produce what specific positive, neutral or negative outcomes?” From a systemic perspective, we need to re-examine segregation (or any similar policy) with these inquiries: What are its purposes? Does it achieve those purposes? Are there levels of harm that are counterproductive? Are there alternatives? The sometimes polarizing views of many prison practices, including segregation, can be addressed through competent research.

The Assessment of Correctional Settings

Attempts to reduce negative outcomes and to evaluate specific practices are but two components of broader efforts to assess the overall effectiveness of correctional settings and programs. Historically, we note five types of assessment, each progressively more systematic and structured. In most of these approaches, the relation between assessed quality and outcome generally focuses on various measures of recidivism, prison behavior, and a recently emerging favorite—the cost-benefit ratio.

One early type of assessment, naïve in retrospect, used standardized psychological instruments (e.g.,

CPI, MMPI) administered to samples of staff or offenders. More directly relevant to the actual prison setting, the *Correctional Institution Environment Scale* (CIES) enjoyed some popularity, particularly in North American prisons. Judicial or special reviews were prevalent in the 1970s and 1980s, many in the form of class action suits which had a major impact on prison environments. Such reviews often included expert panels and/or the use of standardized survey methods. A third type of assessment, the audit, comes from the business tradition and is based on the increasingly popular position that good business practice makes for good correctional practice. Such audits typically monitor the extent to which correctional agency policies and procedures are followed. A fourth avenue has been the emergence of accreditation through independent associations such as the American Correctional Association (ACA), the National Commission on Correctional Health Care (NCCHC), and the Association of State Correctional Administrators. In a variant of accreditation, particularly popular in the Correctional Service of Canada and Her Majesty's Prison Service in the United Kingdom, the correctional agency typically convenes a panel of independent experts who evaluate a specific program (e.g., substance abuse treatment) based on criteria developed from the empirical literature, often from meta-analyses.

The final approach pairs empirical knowledge regarding effective correctional treatment, (e.g., “what works”) with a standardized psychometric methodology. The rationale is to use a corrections-specific instrument whose items are intentionally derived from the concepts of “best practices.” The only instrument to date that follows this two-part guideline is the *Correctional Program Assessment Inventory* (CPAI), which covers eight domains: Program Implementation and Maintenance, Organizational Culture, Management/Staff Characteristics, Client Risk/Need Assessment Practices, Program Characteristics, Core Correctional Practice, Inter-Agency Communication, and Program Evaluation. These domains are consistent with the principles of risk, need, and responsivity as fundamental characteristics of effective correctional intervention. Scores on the CPAI, computed on

the basis of independent expert assessment, have been shown to correlate with a range of impact measures (e.g., recidivism) across a number of settings and groups. By encouraging correctional evaluators and their client agencies to link the underlying principles of effective intervention with daily practice, the CPAI also promotes knowledge transfer to front line corrections staff about “what works.”

Conclusions

This review has addressed several systemic issues that impact professional work and influence correctional outcomes. It is clear that solutions to the context questions raised here are elusive. Nevertheless, we argue that state-of-the-art practice in corrections should consider systemic issues as co-equal to the traditional focus of professional psychology. As we move to the next era of correctional psychology, we must examine how these unique environments and macro-systems impinge on the delivery of best practices. Whether in recognizing and counteracting iatrogenic outcomes, addressing systemic contributors to offending, researching the impact of selected practices like segregation, or systematically evaluating the performance of institutions and programs, we have an obligation to provide high quality service and leadership to these critically important functions of the justice system.

Selected References

- Abramsky, S., & Ross, J. (2003). *Ill-equipped: U.S. prisons and offenders with mental illness*. New York: Human Rights Watch.
- American Association for Correctional Psychology. (2000). Standards for Psychology Services in jails, prisons, correctional facilities, and agencies: Second Edition. *Criminal Justice and Behavior*, 27, 433-494.
- Anderson, V., & Johnson, L. (1997). *Systems thinking basics: From concepts to causal loops* Waltham, MA.: Pegasus Communications.
- Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Twenge, J. M. (2005). Social exclusion impairs self-regulation. *Journal of Personality and Social Psychology*, 88, 589-604.
- Boothby, J. L., & Clements, C. B. (2000). A national survey of correctional psychologists. *Criminal Justice and Behavior*, 27, 716-732.
- Brodsky, S.L., & Scogin, F.R. (1988). Inmates in protective custody: First data on emotional effects. *Forensic Reports*, 1(4), 267-280.
- Clear, T. (1994). *Harm in American penology: Offenders, victims, and their communities*. Albany, NY: State University of New York Press.
- Clements, C. B., & McLearn, A. M. (2003). Research-based practice in corrections: A selective review. In T. Fagan and R. Ax (Eds.) *Correctional mental health handbook*. (pp. 273-302). Thousand Oaks, CA: Sage Publications.
- Cloyes, K., Lovell, D., Allen, D.G., & Rhodes, L.A. (2006). Assessment of psychosocial impairment in a supermaximum security unit sample. *Criminal Justice and Behavior*, 33, 760-781.
- Di Placido, C., Simon, T.L., Witte, T.D., Gu, D., & Wong, S.C.P. (2006). Treatment of gang members can reduce recidivism and institutional misconduct. *Law and Human Behavior*, 30, 93-114.
- Dumont, R. (2000). Inmate sexual assault: The plague that persists. *Prison Journal*, 80(4), 407-414.
- Dunaway, R. G., Cullen, F. T., Burton, Jr., V. S., & Evans, T.D. (2000). The myth of social class and crime revisited: An examination of class and adult criminality. *Criminology*, 38, 589-632.
- Dyer, J. (2000). *The perpetual prison machine*. Boulder, CO.: Westview Press.
- Elliott, R. L. (1997). Evaluating the quality of correctional mental health services. *Behavioral Science and the Law*, 15, 427-438.
- French, S. A., & Gendreau, P. (2006). Reducing prison misconducts: What works. *Criminal Justice and Behavior*, 33, 185-218.
- Gendreau, P. & Andrews, D. A. (2001). *Correctional Program Assessment Inventory – 2000 (CPAI-2000)*. Saint John, NB: University of New Brunswick.
- Griffin, M., & Hepburn, J. (2006). The effect of gang affiliation on violent misconduct among inmates during the early years of confinement. *Criminal Justice and Behavior*, 33, 419-466,
- Haney, C. (2003). Mental health issues in long-term solitary and “supermax” confinement. *Crime & Delinquency*, 49(1), 124-156.
- Haney, C. (2006). *Reforming punishment*. Washington, DC: American Psychological Association.
- Haney, C., & Zimbardo, P. G. (1998). The past and future of US prison policy: Twenty-five years after the Stanford Prison Experiment. *American Psychologist*, 53, 709-729.
- Hayes, L. (1999). Suicide in adult correctional facilities: Key ingredients to prevention and overcoming the obstacles. *Journal of Law, Medicine & Ethics*, 27(3), 260-268.

- Hochstetler, A., Murphy, D.S., & Simons, R.L. (2004). Damaged goods: Exploring predictors of distress in prison inmates. *Crime & Delinquency*, 50(3), 436-457.
- Kupers, T. (1999). *Prison madness: The mental health crisis behind bars and what we must do about it*. New York: Wiley/Jossey-Bass.
- Losel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works: Reducing reoffending*. Chichester: John Wiley.
- Lovell, D., & Jemelka, R. (1996). When inmates misbehave: The costs of discipline. *The Prison Journal*, 76(2), 165-179.
- Lowencamp, C. T., Smith, P., & Latessa, E. J. (in press). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology and Public Policy*.
- MacNeil, J.R., Lobato, M.N., & Moore, M. (2005). An unanswered health disparity: Tuberculosis among correctional inmates, 1993 through 2003. *American Journal of Public Health*, 95(10), 1800-1805.
- Magaletta, P.R., Ax, R.K., Patry, M.W., & Dietz, E.F. (2005). Clinical practice in segregation: The crucial role of psychologists. *Corrections Today*, 67(1), 34-37.
- Parke, R.D., & Clarke-Stewart, K.A. (December, 2001). *Effects of parental incarceration on young children*. Washington, DC: The Urban Institute.
- Restum, Z. (2005). Public health implications of substandard correctional health care. *American Journal of Public Health*, 95(10), 1689-1691.
- Schlosser, E. (Dec., 1998). The prison-industrial complex. *The Atlantic Monthly*, 282 (6), 51-77.
- Sturr, C. (2006). The political economy of the prison crisis: An introduction. *Dollars and Sense*, January/February, 8-9.
- Toch, H., Adams, K., Grant, J.D., & Lord, E. (2002). *Acting out: Maladaptive behavior in confinement*. Washington, D.C.: American Psychological Association.
- Tonry, M. (2004). Has the prison a future? In M. Tonry (Ed.), *The future of imprisonment* (pp. 3-24). New York: Oxford University Press.
- Travis J., & Waul M. (Eds.) (2003). *Prisoners once removed: The impact of incarceration and reentry on children, families, and communities*. Washington, D.C.: The Urban Institute Press.
- Waters, J. E., & Megathlin, W. L. (2002). Evaluating change in social climate in a close security state correctional facility. *Journal of Offender Rehabilitation*, 34, 71-84.
- Webb, L. R. (2003). Building constructive prison environments: The functional utility of applying behavior analysis in prisons. *The Behavior Analyst Today*, 4(1), 71-77.
- Welch, M. (2005). *The ironies of imprisonment*. Thousand Oaks, CA: Sage.
- Wright, K. (1991). The violent and victimized in the male prison. *Journal of Offender Rehabilitation*, 16, 1-25.
- Wright, K. N. (2005). Designing a national performance measurement system. *The Prison Journal*, 85, 368-393.
- Zinger, I., Wichmann, C., & Andrews, D.A. (2001). The psychological effects of 60 days in administrative segregation. *Canadian Journal of Criminology*, 43(1), 47-83.

Extensions of the Risk-Need-Responsivity (RNR) Model of Assessment and Correctional Treatment

D. A. Andrews
 Carleton University
 E-mail: daandrews@rogers.com

The principles of RNR are so well known within correctional psychology that many now consider them to be obvious: Draw upon a general personality and cognitive social learning perspective on human behaviour (GPCSL); Introduce human service; Work with the moderate and higher risk cases (risk principle); Focus on individualized dynamic risk factors as the predominant intermediate targets of change (need principle); Target a number of individualized dynamic risk factors (breadth); Use cognitive social learning influence strategies (general responsivity); Modify strategies in accordance with the strengths, motivation, personality, mental status, learning ability, learning style, circumstances and demographics of individual cases (specific responsivity); Community settings are preferred over institutions; Involve service delivery staff and volunteers who possess the key skills, key cognitions, and key association networks (staffing principle); Manage in such a way that RNR adherence is maximized (management principle); Whatever the setting, ethicality, decency, justice, legality, cost-effectiveness and other norms may be paramount.

In regard to staffing, the key service delivery skills are interpersonal relationship skills and the structuring skills that will facilitate change. The key cognitions are self-efficacy beliefs in regard to a) possession of the relationship and structuring skills and b) that adherence will be rewarded by positive outcomes (for the offender and the service provider). The key association networks include colleagues and friends who support RNR and membership in professional associations or groups that are supportive of RNR adherence.

The key management functions include the selection, training (pre-service and in-service) and clinical supervision of service delivery staff. The management functions extend to other ways of enhancing the quality of implementation, including the following: Adequate dosage; printed

/ recorded manuals and materials; monitoring for RNR adherence and intermediate gains and losses; maintain small units; involving researchers in the design, delivery and evaluation of services.

All of the above is obvious, perhaps, but adherence with RNR is a serious challenge in practice in everyday corrections (Andrews, 2006). Challenges also exist within the justice system itself wherein appeals for proportionality, due process and now the pursuit of personal well-being are weakening the crime prevention objective (Andrews & Dowden, in press). We argue that crime prevention through human service be extended through the wider justice system including forensic mental health. Additional extensions include making the reduction of antisocial behaviour an acceptable and legitimate primary objective within general clinical, counselling and community psychology. Indeed, I would like to see RNR-based programming become a legitimate and primary pursuit within human and social services generally, including child, youth, family and school services. With human service so fundamental to reduced recidivism, it is best not to rely exclusively on the expansion of human services within a justice context. Rather, the clinical skills that exist outside of justice and corrections should be drawn into crime prevention on a more systematic and normative basis (Andrews & Bonta, 2006, Chapter 13).

It will not be accomplished easily. In many health, human and social service agencies the reduction of antisocial behaviour is viewed as illegitimate unless it occurs incidentally to the enhanced well-being of the offender. The tension between enhanced well-being and social control is well-known in forensic mental health but a major source of that tension is rhetorical. For example, health services are described as enabling crime reduction while corrections is said to restrict rather than enable. The rhetoric of avoidance (“bad”) versus approach (“good”) goals is

similarly not helpful. Human service in corrections, as in any other area, is subject to ethical and professional codes of conduct. Our approach in this paper is to review the psychology of criminal conduct underlying RNR, review the research evidence in support of RNR, and review the challenges. The challenges come from critical criminology, critical feminism, humanistic feminism, and forensic mental health, including the Good Lives Model (GLM).

A General Personality and Cognitive Social Learning Perspective (GPCSL) on Human Behaviour, Criminal and Noncriminal

The personal, interpersonal and community-reinforcement perspective on criminal and noncriminal behaviour (PIC-R) is our version of GPCSL (Andrews & Bonta, 2006, Chapter 4). It explicitly recognizes antisocial attitudes, antisocial associates, a history of antisocial behaviour and antisocial personality pattern as indicators of the four sets of major causal variables. The big four determine whether the personal and interpersonal influences are favourable or unfavourable to crime. Additional predictive information is carried by consideration of patterns of rewards and costs in the major settings of home, school / work and leisure / recreation. Substance abuse completes the “central eight” risk / need factors.

For intervention purposes, the intermediate targets of change are enhancement of personal, interpersonal and community contingencies favourable to non-criminal alternatives to pro-criminal ways of thinking, feeling and acting. Increased rewards and satisfactions for non-criminal pursuits in the major behavioural settings of family, school / work and leisure / recreation is a major way of maintaining changes in attitudes and associates.

Empirical Support for GPCSL: Prediction

Summarizing the results of eight meta-analyses on risk / need factors, the grand mean validity estimate (r) for the big four (.26) was significantly greater than the grand mean estimate for the moderate four (.17) and the latter estimate was significantly greater than the grand mean for minor risk/need factors (.03; Andrews & Bonta,

2006, p. 66). LS/CMI General Risk/Need is a composite of the central eight and its mean predictive validity in relation to general criminal recidivism is .47, $k = 10$, $AUC = .776$. The mean validity estimate for LSI-R equalled or exceeded that of all competing risk/need scales (Gendreau, Little & Goggin, 1996).

The wide-applicability of the central eight has been evident across different types of crime (violent and general recidivism), the categories of age, gender, ethnicity, socioeconomic advantage and mental status, and different types of offenders (general, violent offenders, sex offenders, spouse abusers). Because of the strength and unanimity of arguments within and across critical criminology, critical feminism and humanistic feminism, the predictability of female crime has been a major focus of interest (Reisig et al., 2006). Three meta-analyses support the validity of GPCSL variables with female offenders. The mean predictive criterion validity estimates for LSI-R with female offenders range from .35 ($k = 25$; Lowenkamp, 2007) to .41 ($k = 14$; Goggin & Gendreau, 2004) and .42 ($k = 11$; Rettinger & Andrews, 2007). When computed within the same sample, the mean validity estimates for both men and women are substantial and statistically indistinguishable (Lowenkamp, 2007).

In the prediction of violent recidivism, the mean validity of mental health assessment approaches such as the HCR-20, PCL-R and VRAG are statistically indistinguishable from the mean validity of the LSI (Campbell, French & Gendreau, 2007). In the prediction of violent misconduct in institutions, the mean validity of the LSI equals or exceeds that of the forensic mental health instruments. Overall, for purposes of prediction, assessments based on GPCSL are doing very well.

Empirical Support for GPCSL: Effects of Intervention

I concentrate on Craig Dowden’s (1998) expansion of our original meta-analytic databanks. The key measure of RNR Adherence is a four level variable with the lowest score being “0” (not a human service program, or human service delivered to low risk cases, targeting noncriminogenic needs predominately, and not

employing behavioural or cognitive social learning strategies). The remaining scores (“1,” “2” and “3”) are reserved for human service programs that are in adherence with one, two or all three of risk, need and general responsiveness. Inspection of Table 1 reveals that the correlation of RNR Adherence with effect size is strong and robust across a variety of correctional contexts and a variety of types of offenders and types of programs. Indeed, the effect of RNR Adherence was apparent under all conditions except when non-criminogenic needs were targeted predominately: row 14, Table 1). Notably, non-compliance with the need principle was associated with non-adherence with the risk and general responsiveness principles (only two studies fell in level 2 of RNR Adherence in Row 14). Notable as well, when programs were in adherence with general responsiveness, they were also very likely to be in adherence with risk and need (see Row 11 in Table 1).

Challenges for RNR: Achieving RNR Adherence in Routine Corrections

Examples of strong positive effects on effect size come from well-controlled studies with small samples and with an evaluator who was involved in the design and/or delivery of the program (see Row 15 in Table 1 for demonstration projects and Row 16 for results from routine corrections). RNR Adherence was associated with enhanced effect sizes in routine corrections but the average effect was substantially lower than the average effect of adherence within demonstration programs. In routine or “real life” corrections, evidence of concurrent attention to staffing issues and to the management of program integrity is isolated and sparse (Andrews, 2007). Thus, it is not surprising that some very disappointing results have emerged over the last few years (see Andrews, 2006 for the steps required to enhance RNR adherence in everyday corrections).

Challenges for RNR: Intellectual and Professional Challenges

Some proponents of critical criminology, critical feminism, humanistic feminism, and forensic mental health rebuff RNR for displaying insufficient interest in their preferred issues and practices. Their position is that crime prevention

(reduced recidivism) is less important or less worthy than pursuit of their favoured primary objectives of a) removing structural sources of socio-economic disadvantage and victimization, b) providing relief from emotional distress and psychiatric symptoms, and c) overcoming barriers to living as fulfilling a life as is possible through maximization of personal well-being across the multiple domains of love, sex, friendship, work, intellectual challenge, spirituality, happiness, serenity, self-determination, personal autonomy, and intrinsically motivated pursuits.

Their model of crime (although rarely acknowledged) typically is some variation on the Freudian perspective of frustration-aggression or the 1930’s version of sociological strain / anomie theory. Their model of crime prevention contends that pursuit of their preferred objectives will reduce crime (and if it does not, so what, because their objectives and practices are more positive, noble and beneficial to the offender than the crime prevention focus of RNR).

Tony Ward and colleagues (2006) have summarized the three *crucial* clinical contributions of his positive approach: 1) Treating sex offenders with respect; 2) A sex offender may learn to transfer his pursuit of the goods of loving and helping vulnerable children to volunteer work with recovering and vulnerable adult drug addicts; 3) Substituting avoidance goals with approach goals creates a positive therapeutic environment. Frankly, Ward’s first crucial contribution was preceded by the golden rule and has been a fixture of human and social service for years. Second, the transfer of goods from the deviant to the prosocial is naive at best and dangerous at worst when the motivation for offending is misconstrued. In regard to number 3, the widely distributed RNR-based lists of promising intermediate targets of change are never cited by Ward as he continues to promote GLM while misrepresenting RNR as the Bad Lives Model.

The reduced criminal victimization associated with RNR adherence is compromised by the critics’ pursuit of the status quo, their theoretical naivety and by the lack of empirical support for their approaches. Tony Ward knows that risk reduction is a necessary condition for reduced recidivism through human service (Ward &

Table 1. Mean Effect Size (r) by Level of RNR Adherence In Various Justice and Treatment Contexts including Family, Academic, Vocational, and Substance Abuse Programs (k = number of tests of treatment. Portions of this table were presented in Andrews & Dowden, in press)

	<u>Level of RNR Adherence</u>				r with ES
	0 (k) None	1 (k) Low	2 (k) Moderate	3 (k) High	
1. Full Sample	-.02 (124)	.02 (106)	.18 (84)	.26 (60)	.56
2. Residential Corrections	-.10 (29)	.01 (32)	.20(34)	.17 (30)	.58
3. Restorative Justice	.02 (14)	.14 (4)	.16 (3)	.35 (1)	.59
4. Young Offenders	-.02 (62)	.01 (63)	.20 (39)	.28 (42)	.60
5. Female Offenders	.02 (14)	.03 (10)	.17 (9)	.36 (12)	.57
6. Non-white Offenders	-.08 (40)	.03 (33)	.20 (35)	.25 (15)	.65
7. Family Therapy	-.02 (6)	.06 (18)	.22 (17)	.40 (17)	.63
8. Academic Programs	.03 (6)	.07 (20)	.20 (31)	.32 (15)	.47
9. Vocational Programs	-.05 (5)	.05 (13)	.20 (16)	.38 (10)	.68
10. Substance Abuse	-.06 (5)	.07 (10)	.14 (17)	.30 (4)	.61
11. Cognitive-Behavioural	-.07 (1)	-.10 (1)	.14 (15)	.26 (60)	.36
12. A Forensic Mental Health Focus: Reduced Emotional Distress	.02 (20)	.01 (56)	.21 (21)	.14 (4)	.49
13. Some Noncriminogenic Needs Targeted	-.03 (62)	.01 (81)	.23 (26)	.24 (11)	.51
14. Noncriminogenic Needs Targeted Predominately	-.03 (55)	-.00 (59)	.16 (2)	---- (0)	.16 ns
15. Demonstration Project	.01 (1)	.07 (7)	.35 (16)	.38 (23)	.44
16. Real World Program	-.02 (93)	.04 (71)	.09 (34)	.15 (11)	.41

ns : statistically nonsignificant ($p > .05$)

Eccelston, 2004) and he knows that pursuit of intrinsically rewarding activities can be harmful to

others (Ward, Devon, & Beech, 2006). Yet he persists in promoting a perspective that seriously

threatens the effective pursuit of crime prevention (and has not even been shown to enhance personal well-being).

I remain confident that skilled clinicians in psychology and social work increasingly will be drawn into the ethical, legal, humane and decent pursuit of reduced antisocial behaviour. A re-examination of Table 1 will reveal that the noble and positive pursuits of restorative justice, enhanced emotional well-being, enhanced cognitive functioning, and enhanced functioning in family, school and work settings reduced criminal victimization only when those pursuits were in adherence with RNR.

References¹

- Andrews, D. A. (2007, June). Extensions of the RNR Model of Correctional Assessment and Treatment. Paper presented at the North American Conference of Correctional and Criminal Justice Psychology, Ottawa ON, Canada (PP file available: daandrews@rogers.com).
- Andrews, D. A. (2006). Enhancing adherence to risk-need-responsivity: Making quality a matter of policy. *Criminology and Public Policy*, 5 (3), 595-602.
- Andrews, D. A. & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark NJ: LexisNexis/Matthew Bender.
- Andrews, D. A. & Dowden, C. (in press). The Risk-Need-Responsivity Model of Assessment and Human Service in Prevention and Corrections: Crime Prevention Jurisprudence. *Canadian Journal of Criminology and Criminal Justice*.
- Campbell, M. A., French, S., & Gendreau, P. (2007), *Assessing the utility of risk assessment tools & personality measures in the prediction of violent recidivism for adult offenders*. Department of Public Safety & Emergency Preparedness, Ottawa, Canada
- Lowenkamp, C. (2007, September). Personal correspondence.
- Morgan, R. (2007, June). *Treatment of Mentally Disordered Offenders: A Research Synthesis*. Paper presented at the North American Conference of Correctional and Criminal Justice Psychology, Ottawa ON Canada
- Rettinger, J. & Andrews, D. A. (Submitted). General risk/need, emotional distress, poverty, abuse history and the criminal recidivism of adult female offenders.

¹ All references are listed in Andrews & Bonta (2006) except for those listed here. Available upon request to daandrews@rogers.com.

Structured Guidelines for Evaluating Study Quality

Leslie Helmus

Corrections Research, Public Safety Canada

E-mail: Leslie.Helmus@ps-sp.gc.ca

One of the most important questions in the field of correctional psychology is: “does this treatment/intervention work?” Given that knowledge is cumulative, this question has typically been answered through narrative review and more recently, through meta-analysis. Meta-analysis provides quantifiable answers and has been viewed as more objective than narrative reviews. The problem of subjectivity in study selection, however, is still a concern in meta-analysis. This problem can be illustrated with the example of sex offender treatment outcome research. Between 1995-2005, there have been six meta-analyses/reviews, with anywhere between 9-79 studies included (Gallagher, Wilson, Hirschfield, Coggeshall, & MacKenzie, 1999; Hall, 1995; Hanson et al., 2002; Kenworthy, Adams, Brooks-Gordon, & Fenton, 2004; Lösel & Schmucker, 2005; Rice & Harris, 2003). However, different studies were used in each review; studies rated as credible by one group were considered inherently biased by the other groups. Only one study was included in all reviews: California’s Sex Offender Treatment and Evaluation Project (SOTEP; Marques, Wiederanders, Day, Nelson, & van Ommeren, 2005). This variability in study selection demonstrates the difficulty of answering a research question when different researchers are considering different evidence. What is therefore needed is a consensus on study quality.

Although formal assessments of study quality are relatively new, a large number of scales and checklists have been developed within the medical field to assess the quality of randomized and clinical trials (see Juni, Witschi, Bloch, & Egger, 1999). Moher et al. (1995) identified 25 scales and nine checklists, noting considerable heterogeneity among them. Juni et al. (1999) found that the choice of quality assessment scale affected the results of a meta-analysis. They scored 17 studies using 25 different quality assessment scales. These scales did not consistently identify the same

studies as “high quality”, and the best studies identified by these different scales yielded different results. They concluded that the quality rating scales are heterogeneous and that many of the items concern reporting quality, ethical issues, and data interpretation rather than bias or internal validity.

In criminology, one of the most influential rating scales is the Maryland scale (Sherman et al., 1997), originally developed to help identify promising crime prevention programs. Raters using the Maryland scale consider seven elements of “methodological rigour” prior to forming an overall rating, ranging from 1 to 5, with. The Maryland scale has the advantage of being widely applicable to a broad selection of criminal justice intervention studies. However, it lacks a coherent definition of quality, combining concerns about statistical power and bias. In meta-analysis, concerns about statistical power or measurement error fade in comparison to concerns about bias. Despite the considerable progress in developing study quality rating schemes, none of the existing scales are well-suited to measuring the quality of treatment outcome studies in the correctional field.

Collaborative Outcome Data Committee Study Quality Guidelines

The Collaborative Outcome Data Committee was formed in 1997 with the goals of defining standards for research on treatment outcome for sexual offenders, organizing the existing sexual offender outcome studies, and promoting high quality evaluations. The committee’s first report – a meta-analysis of 43 studies (Hanson et al., 2002) concluded that there was a small positive effect for current treatments, but that firm conclusions awaited more and better research. A problem with this meta-analysis was that study inclusion was determined by vote, without a consensus on what constitutes study quality. The Committee has since developed guidelines for assessing study

quality¹. Although these guidelines were developed in the context of sexual offender treatment outcome studies, they are applicable to the study of any intervention in the corrections field.

Assumptions Guiding the Rating Scheme

The Collaborative Outcome Data Committee's Guidelines for the Evaluation of Sexual Offender Treatment Outcome Studies (CODC Guidelines) were based on the following assumptions:

A) *It is possible and desirable to rate study quality* - Studies vary in the extent to which they can inform research questions, and better studies should be given more weight than lesser studies.

B) *Knowledge is cumulative* - In offender treatment, there is unlikely to be a single definitive study involving a multi-site, randomized clinical trial involving thousands of patients. The complexity of the interventions and the long delays needed before knowing the ultimate outcome (i.e., recidivism) present significant technical obstacles. Therefore, researchers have to rely on the outcomes of numerous smaller studies.

C) *Multiple methods are needed* - There is no single method for determining the truth. Random assignment has been recognized as the gold standard for minimizing pre-existing differences between the treatment and comparison groups, but it is not the only source of information. Quasi-experimental designs can make important contributions to knowledge if special care is taken in their design, implementation, and interpretation.

D) *Program evaluation can and should contribute to cumulative knowledge* - Most studies of sexual offender treatment are program evaluations, not scientific experiments. In scientific experiments, the research is designed to address specific questions of scientific interest. In contrast,

program evaluations are concerned with the workings of a specific program. Well-designed program evaluations can contribute to cumulative knowledge, informing questions concerning the efficacy of both "this specific program" and programs "like this one."

Consequently, the CODC Guidelines focus considerable attention on how to maximize the contribution of program evaluations to cumulative knowledge.

Overview of the CODC Guidelines

The Guidelines were based on a review of existing study quality scales as well as specific concerns that have been raised about sexual offender research (e.g., Rice & Harris, 2003). Much of the content and structure of the CODC Guidelines were derived from an analysis of how CODC members described the strengths and weaknesses of individual research studies. Although the Committee originally envisioned separate criteria for specific designs (e.g., random assignment, cohort), the concerns raised for the different designs were remarkably similar. Consequently, the CODC Guidelines present general criteria for evaluating offender treatment outcome studies, and only occasionally provide distinct questions for specific research designs.

In order to rate study quality, it is necessary to have a definition of what is being rated. Ideally, the effect size calculated from a study would be wholly attributable to differences in treatment (plus random error). Bias is therefore the major criterion for judging study quality, but, it is also worth considering the *confidence* that can be placed in the finding. A random assignment study, for example, would not be expected to produce systematic differences between groups; nevertheless, increased confidence can be placed in the results when the researchers explore various potential threats to validity and are able to demonstrate that the study was implemented as intended. Consequently, a high quality study is one in which the judgement of *minimal bias* can be made with *high confidence*.

The CODC Guidelines contain 20 items organized into seven categories: a) administrative control of the independent variable; b) experimenter expectancies; c) sample size; d) attrition; e)

¹Contributors to the CODC guidelines are, in alphabetical order: Anthony Beech, Darren Bisshop, Guy Bourgon, Dawn Fisher, R. Karl Hanson, Andrew Harris, Calvin Langton, Roxanne Lieb, Janice Marques, Michael Miner, William Murphy, Michael Seto, Vernon Quinsey, David Thornton, and Pamela Yates

equivalence of groups; f) outcome variables; and g) correct comparison conducted. As well, there is one item to be rated only for cross-institutional designs (Sample Size of Institutions), and three checklists to help with the rating of Item 13 (A Priori Equivalence of Groups) for specific types of designs (random assignment, risk band/norm, and cohort studies). A flow chart, adapted from Zaza et al. (2000), is provided to help reviewers categorize studies. The 20 (21 for cross-institutional designs) items concern the extent to which the study's features introduce bias in the estimation of the treatment effect, or influence the confidence that can be placed in the study's findings. If the information is limited, raters are encouraged to seek out additional information and re-rate the item.

The overall judgement of study quality is a form of structured judgement. After rating the individual items, evaluators are asked to form global judgements as to the extent of "bias" inherent in the research design, and the "confidence" that can be placed in the bias rating. The bias ratings are as follows: a) no bias or minimal bias expected; b) some bias expected; and c) considerable bias expected. The overall confidence ratings similarly use a three-point scale: a) little or no confidence in the results; b) some confidence; and c) high confidence in the results as reported. Based on the ratings of bias and confidence, studies are placed in one of four categories:

- a) **STRONG.** High confidence that the study has minimal bias in estimating the effectiveness of sexual offender treatment.
- b) **GOOD.** High confidence that the studies have no more than a small amount of bias; OR, some confidence that the study has minimal bias (intermediate rating).
- c) **WEAK.** Some confidence that the studies have no more than a small amount of bias (intermediate rating).
- d) **REJECTED.** Low confidence in the results, *or* considerable bias.

Reliability

Two undergraduate students were given 5 days training on the use of CODC guidelines². This training primarily involved rating and reviewing eight practice studies with a trainer³. The two raters then independently coded 10 studies.

On the Global rating, the coders agreed on nine of the 10 studies ($ICC = .947$). There was 100% agreement on Global Confidence ($Kappa = 1.0$; $ICC = 1.0$), 90% agreement on Global Bias ($ICC = .690$; $Kappa$ could not be computed), and 70% agreement on Global Direction of Bias ($Kappa$ could not be computed). The level of agreement for the individual items was also high. For most of the categories, the median level of agreement was 1.0.

A second reliability study was conducted using 12 experts in the field of sexual offender research⁴. To examine the reliability of the expert's ratings, 10 hypothetical studies were used. Real studies were not used because the experts would be expected to have already formed opinions about the existing studies, either as authors or reviewers. The experts were not provided with any specific training in the use of the CODC guidelines, although half of them would have known about the guidelines through their membership on the CODC committee. Each expert rated between 1-6 studies, so that each study had three separate ratings.

Agreement among the experts was poor. The experts had moderate levels of agreement for the individual items, but disagreed on the overall ratings. Of the 10 studies, zero had agreement by all three raters, and only three had two common ratings (all rejected). Some of the disagreements were due to errors that could have been corrected given training and additional care and attention (failure to notice study features, misinterpretation of the coding rules). Subsequent discussions among the experts, however, revealed principled disagreements concerning the minimum features that required studies to be "good enough". There

² Leslie Helmus, Shannon Hodgson.

³ Guy Bourgon.

⁴ Guy Bourgon, Andrew Harris, Grant Harris, Niklas Långström, Roxanne Lieb, Ruth Mann, Robert McGrath, William Murphy, Vernon Quinsey, Marnie Rice, David Thornton, Pamela Yates.

was substantial agreement that the individual features identified in the CODC guidelines were important indices of study quality, but the experts had divergent views as to the relative importance of these features in influencing overall study quality.

The main conclusions of the reliability studies are that it is possible to train naïve raters to reliably use the CODC guidelines; simple exposure to the guidelines, however, was insufficient to change strongly held beliefs about the appropriate methodology to use in sexual offender outcome research (for similar findings in medical research, see Schroter et al., 2004).

Uses of the CODC Guidelines

Although designed for sexual offender treatment outcome studies, the Guidelines can be applied to any human service intervention with a criminal justice population. The only alteration necessary pertains to the item “length of follow-up,” where the criteria were based on the optimal follow-up to detect sexual recidivism, which is a low base rate event. There are three tasks for which the CODC Guidelines should be helpful: a) reviewing existing studies; b) evaluating existing programs; and c) designing new studies of treatment effectiveness. Reviewers can use the Guidelines as a part of their selection criteria for narrative or quantitative syntheses of the evidence of the effectiveness of treatment for offenders. The Guidelines should also be helpful to editors of professional journals (and reviewers) as a means of rating study quality and providing direction for improvements. For program developers, the guidelines can suggest features that facilitate future evaluation (e.g., routinely collecting information on the individuals not admitted to the program). Program evaluators are often given less than ideal conditions under which to determine the effectiveness of treatment. Nevertheless, evaluators can use the Guidelines to make design decisions that maximize information at a minimal cost.

References

Gallagher, C. A., Wilson, D. B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual

reoffending. *Corrections Management Quarterly*, 3, 19-29.

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology*, 63, 802-809.

Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J. K., Murphy, W., Quinsey, V. L., & Seto, M. C. (2002). First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment of sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14, 169-194.

Juni, P., Witschi, A., Bloch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association*, 282, 1054-1060.

Kenworthy, T., Adams, C.E., Brooks-Gordon, B., & Fenton, M. (2004). Psychological interventions for those who have sexually offended or are at risk of offending (Cochrane Review). *Cochrane Library, Issue 3*. Chichester, UK: John Wiley & Sons.

Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, 1, 117-146.

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

Moher, D., Jadad, A.R., Nichol, G., Penman, M., Tugwell, P., & Walsh, S. (1995). Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists. *Controlled Clinical Trials*, 16, 62-73.

Rice, M. E., & Harris, G. T. (2003). The size and sign of treatment effects in sex offender therapy. *Annals of the New York Academy of Sciences*, 989, 428-440.

Schroter, S., Black, N., Evans, S., Carpenter, J, Godlee, F., & Smith, R. (2004). Effects of training on quality of peer review: Randomized controlled trial. *British Medical Journal*, 328, 673-5.

Sherman, L. W., Gottfredson, D., Mackenzie, D., Eck, J., Reuter, P., & Bushway, S. (1997).

Preventing crime: What works, what doesn't what's promising. A report to the United States Congress. College Park, Maryland: University of Maryland, Department of Criminology and Criminal Justice.

Zaza, S., Write-De Agüero, L. K., Briss, P. A., Truman, B. I., Hopkins, D. P., Hennessy, H., et al. (2000). Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. *American Journal of Preventive Medicine*, 18, 44-74.

Sexual Conflict and Coercion

Vernon L. Quinsey
Queen's University, Kingston, ON
E-mail: vern.quinsey@queensu.ca

I will describe an evolutionary or Darwinian approach to understanding sexual conflict and coercion. Evolutionary psychology aspires to develop a consilient theory that derives from, and is applied, to comparative, historical, anthropological, survey, archival, and laboratory data. *Consilience*, in this context, means that the theory is consistent with those of the more developed life sciences.

Sexual aggression forms a part of sexual conflict. Sexual conflict occurs when the reproductive interests of opposite-sexed individuals are opposed, in the sense that one individual's success occurs at the cost of another of the opposite sex. Sexual conflict is widespread among animals, including species that exhibit social monogamy and bi-parental care. Sexual conflict ultimately arises from sexually dimorphic reproductive strategies and, as we shall see, can drive genetic change.

An evolutionary account of sexual conflict distinguishes between *proximal causes*, the mechanisms that cause a particular characteristic in the present environment, and *ultimate causes*, the features of the ancestral environment that caused an adaptation to evolve. Proximal causation deals with mechanisms responsible for the development of characteristics in the current environment and ultimate causation addresses the question of why these characteristics developed. Characteristics that have developed over generations because of their relationship to reproductive success are genetic in nature and are termed *adaptations*. For example, life history strategy involves a suite of adaptations and refers to a genetically organized life course dictating how individuals allocate energy to aspects of reproductive fitness, such as body growth, mating effort, and parental effort. Adaptations are easy to postulate but have been hard to prove until recently.

Common errors in thinking about these matters include the following: a) to suppose that proximal and ultimate causes are at the same level of explanation (i.e., are opposed to each other); b) to assume that genetic causes of behavior must show heritability coefficients greater than zero (they only do if the trait shows variance in the population); c) to believe that adaptations are necessarily related to reproductive success in modern human environments; d) to think that adaptations are good for us or necessarily morally acceptable; and e) to believe that adaptations cannot be revealed by experimentation.

I will review theory and data on gender differences in sexual behaviour and preference and discuss the influence of these sex differences upon manifestations of sexual conflict, including cuckoldry and rape. I will conclude by describing individual differences among women and men that are related to sexual conflict and, in the case of male offenders, predict the likelihood of future instances of sexual coercion.

Sexual conflict between the sexes involves genes that confer a benefit to one sex and a cost to the other (e.g., a gene producing hirsute facial adornments). Because of the way genes are organized on chromosomes, genes involved in sexual conflict are close to the gender determining genes (and therefore travel with them to the appropriate sex most of the time). In order for sexually antagonistic genes to spread in the population they must provide a net reproductive benefit and thus there are mechanisms for the limiting the amount a particular gene is expressed in the "wrong" sex. These mechanisms include sex limited expression where the gene is only expressed in the presence of a hormone or amount of hormone that is present in only one sex.

Bill Rice, at the University of California, Santa Barbara, has documented sexually antagonistic genes at work in fruit flies through ingenious and elegant experiments. In one study, he prevented

females but not males from evolving. After 40 generations experimental males fathered more offspring, prevented their competition from siring offspring, and caused females to die young. Further work has shown that female fitness losses that are occasioned by exposure to multiple males are not compensated for by the reproductive performance of the females' grandsons. This is a true arms race.

There are sexually dimorphic reproductive strategies in humans. When men are unconstrained by circumstance, they prefer more sexual partners than women and when women are unconstrained by circumstance, they prefer fewer partners than men but more resources to be invested in the relationship and in parental assistance (e.g., Landoldt, Lalumière & Quinsey, 1995). Greater male mating effort, risk acceptance, and dominance striving is explained by greater male than female variance in reproductive success. Spectacular historical demonstrations of male reproductive success are provided by genetic studies of patrilineages begun by Niall of the Nine Hostages in Ireland and Genghis Khan in central Asia. These observations explain the differences in crime rates as a function of age and sex that are known as the fundamental data of criminology (for an extensive review, see Quinsey, Skilling, Lalumière & Craig, 2004).

Because sexual behavior and interests have been shaped by reproductive success in ancestral environments, rape is expected to be directed at reproductively relevant targets and involve reproductively relevant behaviours (Quinsey, 2003). Anything that causes men to disregard the preferred mating strategies of women is expected to increase the likelihood of rape. Anthropological, historical, and psychological evidence suggests that warfare, alcohol intoxication, psychopathic personality characteristics (Harris, Rice, Hilton, Lalumière & Quinsey, 2007), misogynist attitudes, and hyperdominant or sadistic sexual interests (Lalumière, Harris, Quinsey & Rice, 2003) contribute to rape. Oddly, men who perceive themselves as highly successful with women are more likely than other men to engage in date rape, presumably because, if their current dating partner breaks off their relationship because of sexual coercion, other partners are readily available

(Lalumière, Chalmers, Quinsey & Seto., 1996). For an extensive review of the rape literature, see Lalumière, Harris, Quinsey and Rice (2005).

Although women show greater preference than men for traits in sexual partners associated with long term mating strategies, such as resources and status, there is variance among women in their interest in casual relationships (e.g., Landoldt et al., 1995). Provost, Kosakoski, Kormos, and Quinsey. (2006) and Provost, Troje, and Quinsey (in press) found that women using short-term mating strategies appear to prefer genetic over parental contributions of mating partners more than women using long-term mating strategies--short-term partners were unlikely to make parental contributions in ancestral environments.

Although men interfere with women's reproductive strategies through sexual coercion, women interfere with men's reproductive strategy of paternal investment through cuckoldry. On average, ancestral men who invested in children who were unrelated to them were less reproductively successful than men who invested only in their children. There is evidence for genetic contributions to female infidelity and cuckoldry is common enough to lead us to expect that men may well have developed psychological adaptations to the threat of it. Volk and Quinsey (2002) showed for example that men but not women are more willing to adopt babies that they believe resemble them. In an offender sample, Camilleri and Quinsey (submitted) found, in support of the idea that men use sexual coercion to counter sperm competition in committed relationships, that partner rapists experienced cuckoldry risk events prior to committing their offense and had more such experiences than non-sexual partner assaulters. In a community sample, direct cues of infidelity predicted self-reported propensity for sexual coercion.

In summary, the conceptualization of sexual behavior as reproductive behavior and of sexual coercion and cuckoldry as manifestations of sexual conflict yields many interesting hypotheses for evaluation. This conceptualization also encourages the development of consilient theories and an interdisciplinary approach. The genome project, advances in neuro-imaging, and

conceptual advances in developmental and evolutionary biology mean that future research can solve rather than merely grapple with the fundamental problems of psychology. Future successful theories will situate proximal causal mechanisms in the context of ultimate causation.

References

- Camilleri, J. A. & Quinsey, V. L. (submitted). Testing the cuckoldry risk hypothesis of partner sexual coercion in forensic and community samples.
- Harris, G. T., Rice, M. E., Hilton, N. Z., Lalumière, M. L., & Quinsey, V. L. (2007). Coercive and precocious sexuality as a fundamental aspect of psychopathy. *Journal of Personality Disorders, 21*, 1-27.
- Lalumière, M. L., Chalmers, L. J., Quinsey, V. L., & Seto, M. C. (1996). A test of the mate deprivation hypothesis of sexual coercion. *Ethology and Sociobiology, 17*, 299-318.
- Lalumière, M. L., Harris, G. T., Quinsey, V. L., & Rice, M. E. (2005). *The causes of rape: Understanding individual differences in the male propensity for sexual aggression*. Washington, DC: American Psychological Association.
- Lalumière, M. L., Quinsey, V. L., Harris, G. T., Rice, M. E., & Trautrimas, C. (2003). Are rapists differentially aroused by coercive sex in phallometric assessments? *Annals of the New York Academy of Sciences, 989*, 211-224.
- Landolt, M. A., Lalumière, M. L., & Quinsey, V. L. (1995). Sex differences and intra-sex variations in human mating tactics: An evolutionary approach. *Ethology and Sociobiology, 16*, 3-23.
- Provost, M. P., Kosakoski, G., Kormos, C., & Quinsey, V. L. (2006). Female sociosexuality and preference for male masculinization. *Archives of Sexual Behavior, 35*, 305-312.
- Quinsey, V. L. (2003). Etiology of anomalous sexual preferences in men. *Annals of the New York Academy of Sciences, 989*, 105-117.
- Quinsey, V. L., Skilling, T. A., Lalumière, M. L., & Craig, W. (2004). *Juvenile delinquency: Understanding individual differences*. Washington, DC: American Psychological Association.
- Volk, A., & Quinsey, V. L. (2002). The influence of infant facial cues on adoption preferences. *Human Nature, 13*, 437-456.

Section B

Risk Assessment

Current Status of Violence Risk Assessment: Is there a Role for Clinical Judgment?

Marnie E. Rice

Mental Health Centre, Penetanguishene, ON

E-mail: mrice@mhcp.on.ca

One of the areas of greatest progress in the field of correctional psychology over the past two decades is the development of tools for the assessment of risk of violent recidivism. Among the first and well known of those instruments is the Violence Risk Appraisal Guide, an actuarial tool that contains 12 static predictors (Harris, Rice, & Quinsey, 1993; Quinsey, Harris, Rice & Cormier, 2006). The instrument yielded an effect size for predictive accuracy for the construction sample (ROC area = .76) that is considered large by conventional standards (Rice & Harris, 2005). There have now been nearly 50 replications of one or both of the VRAG and its sister instrument for sex offenders, The Sex Offender Risk Appraisal Guide (SORAG), yielding large median effect sizes (ROC areas) of .72 for the VRAG and .73 for the SORAG (<http://www.mhcp-research.com/ragreps.htm>).

As illustrated by the studies described on our website cited above, the VRAG has also been shown to have considerable generalizability. It has been shown to predict violence over as short a time period as 12 weeks (ROC area = .71), and over as long a period as 10 years (ROC area = .74). Although developed on Canadian offenders and mentally disordered offenders, it has been shown to predict violence for similar populations in the United States and Europe (20 studies, mean ROC area = .77). It has also been shown to predict self-reported violence (2 studies, mean ROC area = .73), institutional violence or misconduct (6 studies, mean ROC area = .68), offenses known from the criminal record to be sexual violence (7 studies, mean ROC area = .63; see Hanson & Morton-Bourgon, 2007), and any reoffense whether violent or not (two studies, mean ROC area = .76). It has also been shown to predict violent recidivism for offenders under the age of 18 (or even under the age of 16) at the time of the offense (Quinsey et al. 2006), developmentally delayed offenders, sex offenders, wife assaulters,

and emergency psychiatric patients, all with large effect sizes.

In summary, then, the VRAG has been shown to be a good predictor of future violence across many different jurisdictions, using many different definitions of outcome, across many different populations, and over a wide range of followup times.

Several authors have classified risk assessment tools as belonging to one of three or four “generations” (e.g., Campbell, French & Gendreau, 2007). First generation tools are unstructured, non-systematic, subjective clinical judgments, and research has convincingly shown these to be inferior to tools of subsequent generations (e.g., Grove et al., 2000). The VRAG and SORAG are among what the instruments called “second generation” inasmuch as they are purely actuarial – the item selection was based only on finding the maximally predictive combination, regardless of theoretical value or usefulness for management or treatment. All of the VRAG/SORAG items are “static” inasmuch as they are scored strictly based on historical items, usually measured shortly after an offender’s admission to hospital or prison. Critics have argued that the goal of risk assessment should be to inform effects to reduce risk rather than simply to assess it (e.g., Wong & Gordon, 2006), and that second generation tools fail to measure theoretically relevant factors or the change in risk over time (Andrews, Bonta, & Wormith, 2006).

In response to these criticisms, so-called “third generation” risk measures have been developed that include empirically-supported risk factors that are also theoretically-motivated and/or “dynamic” – that is, that are thought to change with time or treatment and improve upon the prediction of risk based on static predictors. An example of such a tool is the LSI –R (Andrews & Bonta, 1995). Most recently, “fourth generation” risk

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 20-23. Public Safety Canada.

Author note: Thanks are due to Grant Harris and Karl Hanson for comments on earlier versions of this article.

assessments have been developed that are similar to those of the third, but designed to be administered on multiple occasions through an offender's contact with the criminal justice system. These instruments are meant to be used to select treatment targets, evaluate treatment progress, and facilitate risk management (Campbell, French, & Gendreau, 2007). Examples of fourth generation tools are the LS/CMI (Andrews, Bonta, & Wormith, 2004) and the VRS (Wong & Gordon, 2006).

In addition to differing "generations" of instruments, there are also tools that, although not constructed by empirically testing the variables for inclusion, contain many variables shown in other studies to be related to violent recidivism. They also contain other clinical variables that have less empirical support, but are thought to be relevant to treatment and management. In addition, the creators of these tools encourage users to use their clinical judgment to adjust the score once calculated. An example of these "structured professional judgment" tools is the HCR-20 (Webster, Douglas, Eaves, & Hart, 1997).

Are the so-called "third" or "fourth" generation or "structured professional judgment" tools really an advance over purely actuarial static tools such as the VRAG? Hanson and Morton-Bourgon (2007) examined the accuracy of prediction of three types of recidivism (sexual, violent, or any) for sex offenders. They compared the accuracy of actuarial tools (all were "second generation") with structured professional judgment tools. In every case, the actuarial tools significantly outperformed the structured professional judgment tools. Campbell, French and Gendreau (2007) examined whether the predictive accuracy of risk prediction tools varied across generation and according to whether or not they included "dynamic" variables. They examined both institutional violence and violence upon release. Overall, their results gave no reason to conclude that third or fourth generation tools outperformed those of the second generation.

At the present time, then, there is scant evidence that third or fourth generation tools, or tools that contain "dynamic" variables outperform purely actuarial tools. Similarly, there is scant evidence

that clinical judgment can improve upon the prediction possible using purely static actuarial instruments. Other than overwhelming evidence that actuarial or formulaic tools outperform first generation (purely unstructured clinical judgment) tools (e.g., Grove, Zald, Lebow, Snitz, & Nelson, 2000; Aegisdottir et al., 2006; Hanson & Morton-Bourgon, 2007), there is little evidence that we have made progress in violence risk assessment since the development of the second generation purely static actuarial tools.

Why might this be? One possible reason is that truly "dynamic" variables have not yet been identified. In my view (see also Kraemer et al., 1997), before a variable can be said to be truly dynamic, several criteria must be met. First, the variable must be scoreable before, and related to, recidivism. Second, an individual's score on the variable must be changeable as a result of time or treatment. Third, the changed score must predict significantly better than the unchanged score. Finally, the changed score must add to the predictive accuracy possible based on the best available purely actuarial tool for the outcome and population in question. As I outline below, it is my view that no one has yet demonstrated that such variables have been identified, at least among adult offenders.

One example from our own research illustrates the challenge in finding truly dynamic variables (Rice, Quinsey, & Harris, 1991). In that study, we endeavored to reduce the risk of violent recidivism among child molesters by changing their phallometric age preferences. Because phallometric age preferences were measured before release and because they have been shown to be related to subsequent violent and sexual recidivism (Hanson & Bussière, 1998), phallometric preferences meet the first criterion for a dynamic variable. In that study, we found that we could alter phallometric preferences during a laboratory-based treatment, thus fulfilling the second requirement for a dynamic variable. However, when we examined the relationship of the pre-treatment and post-treatment scores to subsequent violent (including sexual) recidivism, we found that the pre-treatment scores predicted better than the post-treatment scores. We were forced to conclude that the treatment was not effective and that the study provided no evidence

to support the idea that phallometric age preferences were truly dynamic.

There have been a few other attempts to identify truly dynamic risk prediction. Hudson, Wales, Bakker, & Ward (2002) identified 26 variables they expected to find were related to sexual recidivism among 246 sex offenders and which they endeavored to change through treatment. Only 6 of the so-called “dynamic” predictors turned out to meet the first criterion – i.e., only 6 turned out to be actually related to sexual recidivism – whether they were measured before or after treatment. After treatment, scores on the 6 were changed (the second criterion) but for only 3 of those 6 were post-treatment scores more strongly related to recidivism than pre-treatment scores (the third criterion). In addition, for 2 of the 3 that did predict better on post-test, offenders who “improved” on the measure as a result of treatment were more likely to reoffend than those who did not. Clearly, clinicians’ assumptions about dynamic risk factors were incorrect.

Most recently, Olver, Wong, Nicholaichuk, & Gordon (2007) reported that post-treatment scores on the VRS:SO (a sex-offender specific version of the “fourth generation” VRS) predicted violent recidivism better than pre-treatment scores and that the addition of dynamic items to the purely static items improved predictive accuracy. Although the direction of the effect supported the post-treatment dynamic over pre-treatment dynamic assessments, a careful reading of their results shows that the improvement in predictive accuracy was not significant. As well, the post-treatment scale that included both static and dynamic items did not significantly outperform the purely static actuarial scale (in fact, it did nonsignificantly worse). At the present time, then, there is no evidence that any truly dynamic treatment targets have been identified, or that judgments about how much offenders have changed on putative dynamic measures add to actuarially-determined risk.

Is it possible that we have been so eager to believe that our treatments for offenders do reduce their risk of violence that we have failed to examine the issue in a rigorous and objective manner? Is it

possible that we have not yet developed theories of criminal and violent reoffending that will yield truly dynamic predictors? Although it is not popular to say so, that is how it seems to me. Although I had hoped that after over 30 years in the field I would be able to come to a more positive conclusion, I regret that, as I see it, a positive conclusion is not yet warranted. When my career started, I, along with most of my colleagues, assumed that social learning explanations of crime were going to lead us to better prediction and treatment. It now appears that play they a much more minor role than we thought. Furthermore, using less than rigorous methods to persuade ourselves that we have indeed developed valid theories and valid dynamic predictors prevents us from continuing to search for more accurate theories and effective interventions.

On the other hand, psychologists in the criminal justice system have much to be proud of. In the past twenty years, the development and application of actuarial risk assessment tools has made possible substantial improvements to public safety without sacrificing the rights of offenders (Rice, 1997). The application of rigorous scientific methods was key to the development of these tools. I believe that continued application of rigorous methods will lead to better theories and the discovery of truly dynamic variables. Clinical and professional judgment is required to formulate the theories we test, and develop the intervention, management, and prevention strategies that follow from empirically-validated theories. Although clinical and professional judgment have an important role to play in risk assessment, that judgment should, for now, be confined to selecting the best actuarial tool for the outcome in question, and in scoring that tool as accurately as possible.

References

- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A. Cooke, R. S., et al. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*, 341-382.
- Andrews, D. A., & Bonta, J. (1995). *The Level of Supervision Inventory- Revised*. Toronto, Canada: Multi-Health Systems.

- Andrews, D. A., & Bonta, J., & Wormith, S. J. (2004). *The Level of Service/ Case Management Inventory (LS/CMI)*. Toronto, Canada: Multi-Health Systems.
- Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime & Delinquency*, 52, 7-27.
- Campbell, M. A., French, S., & Gendreau, P. (2007). Assessing the utility of risk assessment tools and personality measures in the prediction of violent recidivism for adult offenders. (Department of Public Safety and Emergency Preparedness Canada Report No. 2007-04). Ottawa, ON: Department of Public Safety and Emergency Preparedness Canada.
- Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19-30.
- Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362.
- Hanson, R. K., & Morton-Bourgon, K. E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis*. (Public Safety and Emergency Preparedness Canada Report No. 2007-01). Ottawa, ON: Public Safety and Emergency Preparedness Canada.
- Harris, G.T., Rice, M.E., & Quinsey, V.L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior*, 20, 315-335.
- Hudson, S. M., Wales, D. S., Bakker, L., & Ward, T. (2002). Dynamic risk factors: The Kia Marama evaluation. *Sexual Abuse: A Journal of Research and Treatment*, 14, 103-120.
- Kraemer, H.C.; Kazdin, A.E.; Offord, D.R.; Kessler, R.C.; Jensen, P.S.; Kupfer, D.J. (1997). Coming to terms with the terms of risk. *Archives of General Psychiatry*, 54, 337-343.
- Olver, M. E., Wong, S. C. P., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the violence risk scale – sexual offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment*, 19, 318-329.
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (2006). *Violent offenders: Appraising and managing risk* (Second Edition). Washington, DC: American Psychological Association.
- Rice, M.E. (1997). Violent offender research and implications for the criminal justice system. *American Psychologist*, 52, 414-423.
- Rice, M.E. & Harris, G.T. (2005). Comparing effect sizes in follow-up studies: ROC, Cohen's *d* and *r*. *Law and Human Behavior*, 29, 615-620.
- Rice, M.E., Harris, G.T., & Quinsey, V.L. (1991). Evaluation of an institution-based treatment program for child molesters. *Canadian Journal of Program Evaluation*, 6, 111-129.
- Webster, C. D., Douglas, K., Eaves, D., & Hart, D. (1997). *HCR-20: Assessing risk for violence (Version 2)*. Vancouver, BC: Mental Health, Law, & Policy Institute, Simon Fraser University.
- Wong, S. C. P., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment friendly violence risk assessment tool. *Psychology, Public Policy, and the Law*, 12, 279-309.

Gender Comparisons on the Self-Appraisal Questionnaire (SAQ): A Tool for Assessing the Risk of Violent and Non-Violent Recidivism

Toni Hemmati
Carleton University
E-mail: toni@kingston.net

In federally and provincially sentenced populations, females are under-represented in comparison to males. Many studies have provided data on differences between male and female inmate samples on a variety of rater-based risk assessment measures, yet there is a paucity of research examining differences between genders on self-reported risk assessment. The Self-Appraisal Questionnaire (SAQ; Loza, 2005) is a well established self-report designed to assess general and violent recidivism and institutional adjustment in inmate samples. Multiple studies have supported the reliability and validity of the SAQ in male offender samples (e.g., Loza & Loza-Fanous, 2001), but there is a limited amount of research on its application in female inmate samples (e.g., Loza, Neo, Shahinfar, & Loza-Fanous, 2005). However, the female samples have varied in different respects from the male samples (e.g., jurisdiction) making direct comparisons between the two samples difficult. The purpose of the current study is to compare the responses of male and female inmates within the same jurisdiction on the SAQ.

Participants

The participants were 176 (88 male, 88 female) volunteers serving sentences in one of several Ontario Provincial correctional

institutions, and who had no further matters before the court. Table 1 shows the age and sentence length of the participants. Table 2 shows the percentage of offenders who had been convicted for various offence types.

Materials

The Self-Appraisal Questionnaire (SAQ) is an empirically based, theoretically and rationally derived assessment instrument designed to assess the risk of violent and general recidivism. The SAQ consists of 72 items comprising seven subscales: Criminal Tendencies, Antisocial Personality Problems, Conduct Problems, Criminal History, Alcohol/Drug Abuse, Antisocial Associates, and Anger.

The Level of Service/Case Management Inventory (LS/CMI) is a rater-scored comprehensive risk, needs, and responsivity assessment designed to predict the likelihood of recidivism and assist in case treatment, planning, and overall management. The 43-items making up the actuarial risk assessment portion are divided into eight subscales: Criminal History, Education/Employment, Family/Marital, Leisure/Recreation, Companions, Alcohol/Drug Problems, Procriminal Attitude/Orientation, and Antisocial Pattern.

Table 1. Gender Comparison for Ages and Sentence Length

Gender	n	Age	Sentence Length
Males	88	M=31.9 (10.5) range 19-64	M=190 days (144) range 14-729
Females	88	M=34.9 (10.5) range 20-66	M=187 days (187) range 15-729

Table 2. Gender Comparison for Index Offence Type

Offence type	Males %	Females %
Property	18.5	18.5
Robbery	3.7	4.9
Drugs	4.9	28.4
Assault, Threats	25.9	14.8
Weapons	1.2	2.5
Sex offences	7.4	-
Criminal negligence	19.8	4.9
Fraud, Forgery	2.5	14.8
Fail, Breach, Escape	11.1	6.2
Forcible confinement	2.5	1.2
Hate crimes	2.4	-
Misc. minor offences	-	3.7

Method

As part of a larger study, participants were asked to complete a battery of self-report instruments including the SAQ. The LS/CMI (aka Level of Service Inventory-Ontario Revision) scores, as rated by Correctional Classification personnel, were taken from the offender files.

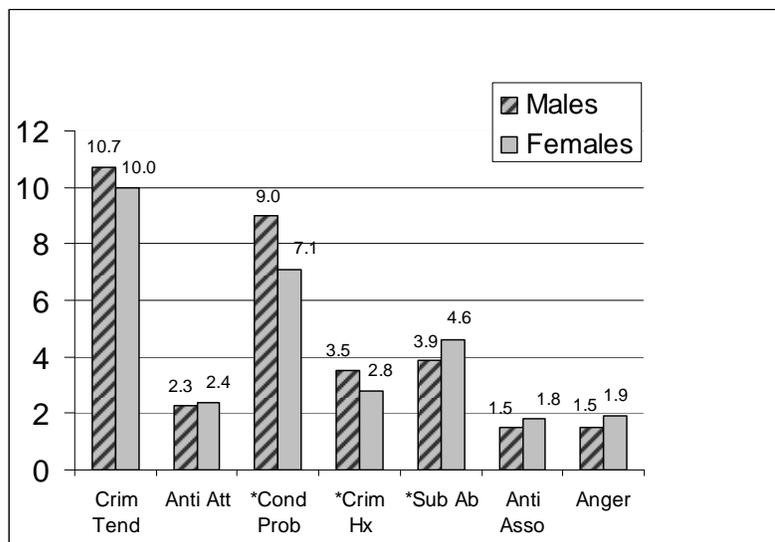
Results

Mean scores for the Total SAQ and LS/CMI are shown in Table 3. There were no significant differences between genders in the Total Scores for either instrument.

Table 3. Male and Female Total Mean Score Comparison

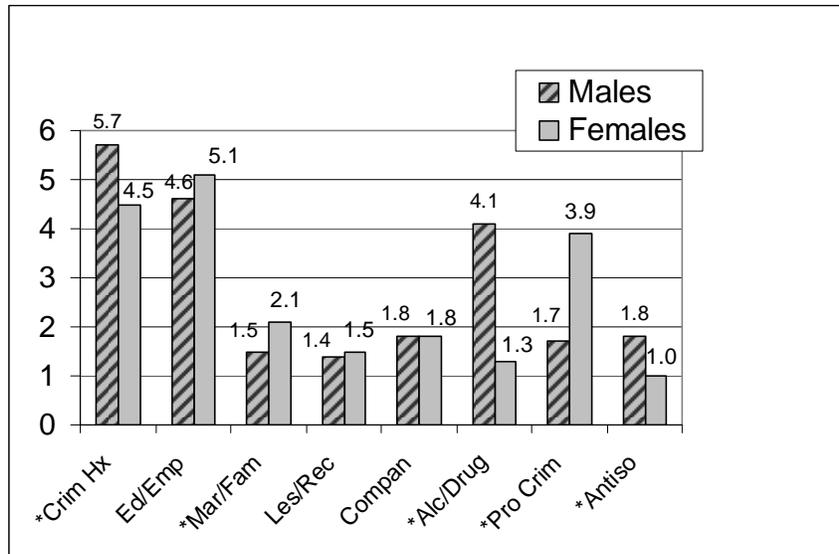
	Males	Females	<i>p</i>
SAQ Total	30.98 (11.67) range 8-53	28.70 (14.28) range 2-61	<i>ns</i>
LS/CMI Total	22.63 (7.74) range 9-37	21.29 (7.83) range 5-36	<i>ns</i>

Mean scores for the SAQ subscales are shown in Figure 1. Three subscales exhibited significant differences between the groups: Conduct Problems (*p*=.000); Criminal History (*p*=.006); and Substance Abuse (*p*=.035). Mean scores for the LS/CMI subscales are shown in Figure 2. Scores were significantly different on five subscales: Criminal History (*p*=.004); Family/Marital (*p*=.021); Alcohol/Drugs (*p*=.000); Procriminal Orientation (*p*=.000); and Antisocial Pattern (*p*=.000).



Note. * = significant difference

Figure 1. Mean SAQ Subscale Scores



Note. * = significant difference

Figure 2. Mean LS/CMI Subscale Scores

Discussion

Similar Total scores between gender groups for both instruments is consistent with the literature, however this is the first time two gender samples from the same jurisdiction were compared on the SAQ. Also consistent with the literature, females had significantly lower scores on SAQ Conduct Problems and Criminal History, but it was surprising that they had significantly higher scores on Substance Abuse. It could be that this difference is spurious, in that a number of females were arrested for drug-related offences (importing narcotics) but were not necessarily drug users: one of the items in this subscale deals with substance-related offences, not necessarily use. In the corresponding LS/CMI Alcohol/Drug subscale, the mean score of the males was more than twice that of the females, however, these LS/CMI items are concerned with alcohol and drug use.

Also surprising was the high rating for females on LS/CMI Procriminal Orientation. The comparable

SAQ Criminal Tendencies and Antisocial Attitudes subscales did not reflect a gender difference. This finding suggests that the item content might differ substantially between the scales; or that there might have been a gender bias when rating the LS/CMI subscale; or the results were sample-specific. Further research to examine self-report vs. rater-scored items on these subscales is suggested.

References

- Loza, W. (2005). *Self-Appraisal Questionnaire Technical Manual*. Toronto: Multi-Health Systems.
- Loza, W. & Loza-Fanous, A. (2001). The effectiveness of the Self-Appraisal Questionnaire (SAQ) in predicting offenders' post-release outcome: A comparative study. *Criminal Justice and Behaviour*, 28, 105-121.
- Loza, W., Neo, L. H., Shahinfar, A., & Loza-Fanous, A. (2005). Cross validation of the Self-Appraisal Questionnaire: A tool for assessing violent and non-violent recidivism with female offenders. *International Journal of Offender Therapy and Comparative Criminology*, 49, 1-14.

Assessing the Predictive Validity of the Youth Level of Service/Case Management Inventory

Kristin Bechtel
University of Cincinnati
E-mail: kristin.bechtel@uc.edu

Christopher T. Lowenkamp
University of Cincinnati
E-mail: christopher.lowenkamp@uc.edu

Edward Latessa
University of Cincinnati
E-mail: edward.latessa@uc.edu

Originally created from an earlier version of the Youth Level of Service Inventory and the adult version of the Level of Service Inventory-Revised, the Youth Level of Service/ Case Management Inventory (YLS/CMI) has become a popular risk/needs assessment tool for juveniles that is being utilized in various correctional, residential and community settings. Updated features to the instrument include addressing responsibility factors as well as providing a format to develop case planning strategies (Hoge, 2001, 2002; Hoge & Andrews, 2002; Hoge, Andrews, & Leschied, 2002).

Following a semi-structured interview with the juvenile and a review of available official documentation, the YLS/CMI produces a summary score based on 42 items which are categorized into eight domains that examine a juvenile offender's risk for delinquent behavior (Hoge & Andrews, 2002). These domains focus on the following areas of risk: (1) prior and current offenses/ adjudications, (2) family circumstances and parenting issues, (3) education and employment, (4) peer relations, (5) substance abuse, (6) leisure and recreation, (7) personality and behavior, and (8) attitudes and orientations. Each of these items was created from multiple studies and meta-analyses on criminality which indicated that these risk factors are significant correlates of youthful offending behavior (Cottle, Lee & Heilburn, 2001; Farrington, 1997; Gendreau, Little & Goggin, 1996; Hoge, 2001; Lipsey & Derzon, 1998; Loeber & Dishion, 1983; Simourd & Andrews, 1994).

While Hoge (2001) suggested that the YLS/CMI is relevant for case planning in both institutional and community settings, given the recent application of this tool, it is advisable to examine the predictive validity of the instrument. In

particular, Gottfredson and Moriarty (2006), caution practitioners to examine the predictive validity of a risk/needs assessment on their targeted population. They acknowledge that samples used in the construction of a risk/needs tool should be representative of the population that the instrument will be used upon; however, small variations between samples may impact the validity of the instrument to predict recidivism. Hence, the process of norming the tool on the targeted population is appropriate when determining the instrument's reliability and validity with their offender base. The current study examines the predictive validity of the YLS/CMI on a sample of youthful offenders that are given either community or institutionally based sentences and attempts to identify if the instrument is more reliable in predicting outcome in either setting.

Framework

While there are relatively few studies that have examined the predictive validity of the YLS/CMI; overall, there appears to be empirical support concerning the predictive validity and reliability of the YLS/CMI. First, Catchpole and Gretton (2003) indicated that the YLS/CMI was significantly related to risk for both violent and general recidivism. Second, Marczyk et al. (2005), compared three types of juvenile assessment tools, the YLS/CMI, Psychopathy Checklist: Youth Version (PCL:YV), and the Massachusetts Youth Screening Instrument (MAYSI) to determine if either of the three instruments could predict if the adult courts would decertify a juvenile offender and return the youth to the juvenile justice system or keep the youth bound over in the adult criminal justice system. As demonstrated by their findings, the YLS/CMI total score was a significant predictor of

certification status and those with higher scores remained in the adult system (Marczyk et al., 2005). Third, previous research indicates that the YLS/CMI has predictive validity across sex, age and race (see Flores, Travis & Latessa, 2004). Fourth, psychometric support for the YLS/CMI has been noted in the manual for the assessment tool (Hoge & Andrews, 2002) in addition to Jung and Rawana's (1999) study. Specifically, Jung and Rawana (1999) demonstrated that the YLS/CMI was able to significantly distinguish between youthful recidivists and non-recidivists.

While these studies have demonstrated the predictive validity of the YLS/CMI, limitations are noted. This includes small sample sizes (Catchpole & Gretton, 2003; Jung & Rawana, 1999; Schmidt, Hoge, & Gomes, 2005), a relatively short follow-up period for the recidivism measure (Catchpole & Gretton, 2003; Jung & Rawana, 1999), scoring the instrument retrospectively (Catchpole & Gretton; Marczyk, Heilburn, Lander, & DeMatteo, 2005) or limiting its applicability to one type of setting (Schmidt et al., 2005).

This study attempts to add to the limited body of knowledge for juvenile offenders risk/needs assessments as well as to address some of the previously mentioned limitations concerning the predictive validity of the YLS/CMI.

Research Questions

Several research hypotheses are proposed to examine Hoge's claim that the YLS/CMI is a valid predictor of youthful offending in both community and institutional settings. The sample is comprised of institutionalized and community-sentenced youth from Ohio. To address the limitations in previous research, this study includes a longer follow-up period (3.4 years) and a larger sample (N=4,482). In addition, the assessments used in the current research are based on an interview with the youth and a review of collateral information at the time of intake to the community or institutional setting. Research questions examined in this study are as follows:

1. Does the YLS/CMI have predictive validity for juvenile offenders?

2. Does the YLS/CMI have predictive validity for community-sentenced juvenile offenders?
3. Does the YLS/CMI have predictive validity for institutionally-sentenced juvenile offenders?
4. Does the strength of the relationship between the total YLS/CMI score and outcome differ for the institutional-based or the community-based juvenile offenders?

Method

Sample

Subjects included in this study are juvenile offenders from all 88 counties in the State of Ohio. Of the 4,482 juveniles included in the sample, there were 3,376 experiencing an institutionally based sentence and 1,106 that were given a community based sentence. In addition, the sample included both males (N=3,884) and females (N=598) and had a racial composition of 52 percent white offenders and 47 percent non-white offenders.

Data Source

Data for this study were obtained from a larger scale project conducted by the University of Cincinnati (see Lowenkamp & Latessa, 2005).

Measures

For this analysis, the total YLS/CMI score was the primary independent variable. Control variables examined in the current research include sentencing type, age, sex, and race. For the current study, age remained a continuous variable, sex was coded as 0= males, 1= females and race was coded 0= whites and 1=non-whites. For interpretation of the findings, each of these dichotomous controls is identified as category 1 in the tables.

Recidivism is defined as a youth who received any type of conviction or commitment post-release from their community or institutionalized sentence, coded as 0= no new conviction or commitment following release from the earlier sentence type and 1= a post-release new conviction or commitment occurred for the juvenile.

Statistical Analysis

Both bivariate and multivariate analyses were conducted for this study. In addition, a z-test was calculated to assess whether or not the magnitude between the regression coefficients significantly differs for the institutionalized and community based youth (Clogg, Petkova & Haritou, 1995). Further, a ROC curve, or receiver operating characteristics analysis was performed as a final measure of the predictive validity of the YLS/CMI (Schmidt et al., 2005; Rice & Harris, 1995).

Results

Descriptive statistics

Table 1 presents the descriptive statistics on all of the variables included in the analyses. The mean score on the YLS/CMI for the entire sample was 21.48, while the institutionalized and community based samples averaged 23.08 and 16.60 respectively¹.

Males comprised the majority of both groups. The average age of the juveniles for the overall sample was slightly over 17 years and the institutionalized youth were an average age of approximately 18 years. The community based youth had a mean age of nearly 16 years. White youth comprised 52% of the sample. A majority of the entire sample was white; however, there were some differences by sentencing type. In particular, institutionalized youth had slightly more non-white juveniles (N=1,687) than white (N=1658). The majority of community sentenced offenders were white (N=673) rather than non-white (N=416).

Regarding outcome for the entire sample, nearly half had recidivated. By sentencing type, 53% of the institutionalized group recidivated in comparison to nearly 35% of the juveniles of the community group.

Correlations

Table 1. Descriptive Statistics

Variable	Total Sample (N=4482)		Institutionalized (N=3376)		Community (N=1106)	
	Mean	SD	Mean	SD	Mean	SD
Independent						
YLS/CMI Score (0-42)	21.48	7.23	23.08	6.29	16.60	7.70
Female	.13	.34	.10	.30	.25	.43
Non-white	.47	.50	.50	.50	.38	.49
Age	17.29	1.64	17.72	1.40	15.98	1.59
Sentence Type	.25	.43				
0 = institutionalized, 1 = community						
Dependent						
Recidivism	.49	.50	.53	.50	.36	.48
0 = no, 1 = yes						

Note: Small differences in the total N for the Institutionalized and Community samples exist due to missing data.

¹ To further describe the groups, means and standard deviations were also calculated for each of the eight domains in the YLS/CMI. In order to distinguish the differences in the means, an independent samples t-test was conducted which revealed in the Levene's test for equality of variances that the only dimension, education and employment, was not significant ($p=.627$) and the F value was .236. However, the t-test for equality of means suggests that there is a significant mean difference for each category and the overall risk score for each group. While not necessarily revealing, this suggests that there is a difference between the static and dynamic risk factors for the two groups.

Table 2. Correlations Between the Youth Level of Service/Case Management Inventory (YLS/CMI) and Outcome

Outcome	Total Sample ^a	95% Confidence Interval for <i>r</i>		Institutionalized ^b	95% Confidence Interval for <i>r</i>		Community ^c	95% Confidence Interval for <i>r</i>	
		Lower	Upper		Lower	Upper		Lower	Upper
Commitment or Conviction	.196	.168	.223	.122	.089	.155	.229	.173	.284
Males	.201	.171	.230	.146	.112	.180	.249	.185	.311
Females	.167	.089	.243	.149	.041	.253	.115	-.003	.230
Whites	.209	.170	.247	.151	.104	.197	.210	.137	.281
Non-Whites	.160	.119	.201	.096	.049	.143	.227	.134	.316
White Males	.227	.185	.268	.187	.138	.235	.242	.159	.322
White Females	.076	-.031	.181	-.051	-.201	.101	.059	-.094	.209
Non-White Males	.149	.105	.193	.106	.057	.155	.224	.117	.326
Non-White Females	.276	.160	.385	.333	.186	.466	.190	-.003	.369

^a With the exception of White Females, all correlations are significant at the $p \leq .01$.

^b With the exception of White Females, all correlations are significant at the $p \leq .01$.

^c With the exception of Females and White Females, all correlations are significant at the $p \leq .01$.

Table 2 illustrates the bivariate correlations between the YLS/CMI total score and recidivism. It should be noted that all correlations were significant except for white females for the entire sample and separately for the institutionalized and the community juveniles. Moreover, females were not significant with respect to the community group².

² This is not surprising as the number of females, especially white females within each group was rather small. Specifically, there were 333 white females in the total sample. Of that total, there were nearly the same in both the institutionalized group (N=167) and the community group (N=166). For the community group itself, there were a total of 273 females. This finding should not be misinterpreted as though the YLS/CMI does not predict for community-sentenced females, or white females as a whole, rather, the small N may suggest that a larger sample is needed to determine if the YLS/CMI is a valid predictor for this group. As noted previously, the YLS/CMI has been able to predict recidivism for females (Jung & Rawana, 1999).

Logistic Regression

First, each variable in the logistic regression models significantly predicted recidivism for the institutionalized group. Several observations can be made with respect to the values of B for the control variables, which suggest that recidivism is more likely to be predicted for youthful, male and non-white offenders. In comparison, regression models for the community group indicated that the total YLS/CMI score, race and sex significantly predict recidivism, but not age. For the community group, these findings suggested that recidivism was more likely for non-white male youth.

Disaggregation of Sentencing Types

Both sentencing type models were found to be significant, $p \leq .001$, with χ^2 values of 292.389 for the institutionalized group and 124.363 for the community group. A comparison of the summed -2 log likelihood values subtracted from the total

sample was calculated to justify the disaggregating of these groups from the same dataset. This value of 28.098 and 5 degrees of freedom justifies the separate models. Table 3 also presents the z-scores which were calculated to compare the parameter estimates for the two models. These z-tests were calculated to assess if the magnitude of the regression coefficients statistically differs between the two sentencing types (Clogg et al., 1995). These two groups do not differ significantly ($z = .71$) based on their total YLS/CMI scores; however, there is a significant difference between these two sentencing types ($z = 2.66$) with respect to the z value for the constant term. This finding offers additional support that the YLS/CMI is a valid predictor for both the institutionalized group as well as the community sample.

Table 3. Logistic Regression Models Predicting Recidivism for Each Sample

Variable	Institutionalized		Community		z
	β	SE	β	SE	
YLS/CMI	.05*	.01	.06*	.01	-.71
Female	-1.46*	.14	-1.19*	.18	-1.18
Non-White	.69*	.07	.52*	.14	1.09
Age	-.16*	.02	-.05	.04	-2.46
Constant	1.58	.48	-.70	.71	2.66**

Institutionalized Sample: * $p \leq .001$; -2 Log Likelihood = 4331.447; $\chi^2 = 292.389$; Cox and Snell R Square = .084; Nagelkerke R Square = .112

Community Sample: * $p \leq .001$; -2 Log Likelihood = 1292.590; $\chi^2 = 124.363$; Cox and Snell R Square = .108; Nagelkerke R Square = .148

** $p \leq .05$

ROC Analysis

The last statistical analysis in the current study explores the ROC curve and specifically the value for the area underneath the curve (AUC).³ For the overall sample, the AUC is .60, while the

institutionalized group revealed an AUC of .56. For the community sample, the AUC is .64, which suggests that the predictive validity of the community sample is relatively strong. Given these findings, the YLS/CMI appears to have predictive validity for both disaggregated groups.

Conclusion

Each analyses offered significant support concerning the predictive validity of the YLS/CMI. Further, the YLS/CMI may be a stronger predictor within community settings than institutional. Future research is recommended since the generalizability of these findings is limited and the study only examined one measure of recidivism; however, support for Hoge's claim was demonstrated.

References

- Catchpole, R. E. H., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders. *Criminal Justice and Behavior, 30*, 688-708.
- Clogg, C. C., Petkova, E., Haritou, A. (1995). Symposium on applied regression: Statistical methods for comparing regression coefficients between models. *American Journal of Sociology, 100*, 1261-1293.
- Cottle, C. C., Lee, R. J., & Heilburn, K. (2001). The prediction of criminal recidivism in juveniles. *Criminal Justice and Behavior, 28*, 367-394.
- Farrington, D. P. (1997). Early prediction of violent and non-violent youthful offending. *European Journal of Criminal Policy and Research, 5*, 51-66.
- Flores, A. W., Travis III, L. F., & Latessa, E. J. (2004). Case classification for juvenile corrections: An assessment of the youth level of service/case management inventory (YLS/CMI), final report. Unpublished manuscript.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*, 575-607.
- Hoge, R. D. (2001). A case management instrument for use in juvenile justice systems. *Juvenile and Family Court Journal, 52*, 25-32.
- Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior, 29*, 380-396.
- Hoge, R. D., & Andrews, D. A. (2002). *The Youth Level of Service/Case Management Inventory manual and scoring key*. North Tonawanda, NY: Multi-Health Systems.

³ Values higher than 50 percent for the AUC reveal a strong predictive validity for the YLS/CMI total score in predicting recidivism (Rice & Harris, 1995; Schmidt et al., 2005).

- Hoge, R. D., Andrews, D. A., & Leschied, A. (2002). *The Youth Level of Service/Case Management Inventory*. Toronto, Canada: Multi Health Systems.
- Jung, S., & Rawana, E. P. (1999). Risk and need assessment of juvenile offenders. *Criminal Justice and Behavior*, 26, 69-89.
- Lipsey, M.W., & Derzon, (1998). Predictors of violent or serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* pp. 86-105. Thousand Oaks, CA: Sage.
- Loeber, R., & Dishion, T. J. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin*, 94, 68-99.
- Lowenkamp, C. T., & Latessa, E.J. (2005). Evaluation of Ohio's Reclaim funded programs, community corrections facilities, and DYS facilities. Unpublished manuscript.
- Marczyk, G. R., Heilburn, K., Lander, T., & DeMatteo, D. (2005). Juvenile decertification: Developing a model for classification and prediction. *Criminal Justice and Behavior*, 32,278-301.
- Rice, M. E., & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology*, 63, 737-748.
- Schmidt, F., Hoge, R. D., & Gomes, L. (2005). Reliability and validity analyses of the youth level of service/case management inventory. *Criminal Justice and Behavior*, 32, 329-344.

Predicting Criminal Recidivism in Adult Male Offenders: A Four Wave Prospective Study

Shelley L. Brown
Carleton University

E-mail: shelleyb@connect.carleton.ca

Edward Zamble
Queen's University

The ability to accurately detect changes in an offender's risk level once he/she has been released into the community is an important function for any correctional organization. Research that explores the extent to which changes in risk factors can predict criminal recidivism has tremendous implications not only for practice but also for theory development as Andrews aptly observed twelve years ago, "Overall, exploration of the predictive validities of assessments of change remains a major issue and is perhaps, the major issue for the development of theory and practice in the psychology of crime" (Andrews, 1995; p.54). The question that remains however is what advancements have occurred since Andrew's seminal statement.

The number of genuine multi-wave prospective studies has grown considerably since 1995 (e.g., Bonta, 1996; Brown, 2002a; Hanson & Harris, 2000; Hanson, Harris, Scott & Helmus, 2007; Motiuk, 1999; Quinsey, Book, & Skilling, 2004; Quinsey, Coleman, Jones, & Altrows, 1997; Quinsey, Jones, Book, & Barr, 2006). However, mixed results coupled with theoretical, methodological or statistical shortcomings associated with some of these studies have lead certain scholars to conclude that there is no conclusive evidence to support the hypothesis that the systematic assessment and re-assessment of dynamic risk adds incrementally to single-wave assessment studies that measure static or dynamic factors on one occasion (e.g., Douglas & Skeem, 2005; Rice, 2007).

Accordingly, the objective of the study is to determine whether or not the re-assessment of prospectively-rated dynamic risk can improve predictive accuracy over and above static risk. This objective is accomplished by extending previous work (e.g., Brown, 2002a) that originally assessed dynamic factors over three waves during an average 10 month follow-up in a sample of 136 adult male offenders. In contrast, this study 1) incorporates an additional fourth wave of dynamic

data collected at the six month post-release phase, 2) extends the average follow-up to 68 months and 3) increases the sample size to 157. It was hypothesized that the strongest prediction model would include both static measures assessed pre-release as well as re-assessments of prospectively-rated dynamic risk.

Method

Participants

One-hundred and fifty seven male offenders about to be released from minimum-, medium-, or maximum-security federal institutions located in Ontario participated in the study. Offenders were selected to participate if they: 1) consented to take part (consent rate: 56.4%); were scheduled to be released on either discretionary (parole) or non-discretionary release (statutory release) within 45 days of the initial pre-release assessment; understood English, were neither actively psychotic nor eligible for deportation; and lastly, would not reach warrant expiry for at least six months from the date of release.

On average, the sample was 33.2 years old ($SD = 10.3$), serving a four-year sentence for a variety of crimes including murder, assault, sexual assault, robbery, drug offences, and property-related crimes. While 54.9% of the sample was released on parole, 45.2 was released on statutory release. Approximately, two-thirds of the sample was Caucasian and 70% was single at time of release.

Measures

Measures were selected to ensure that they were theoretically congruent with the coping-relapse model of criminal recidivism (Zamble & Quinsey, 1997). The coping-relapse model measures four constructs hypothesized to be associated with criminal recidivism: 1) static factors such as criminal history and personality, 2) acute situational factors that may trigger a criminal relapse such as sudden employment loss or heated

argument, 3) acute cognitive and emotional appraisal factors that emerge in response to acute situational factors such as negative affect and perceived stress, and 4) stable dynamic response mechanisms including criminal attitudes, criminal associates, coping efficacy, social support, and substance abuse.

Static Factors. The following six static factors were included: age at time of release, the Statistical Information on Recidivism Scale (SIR-R1; Nuffield, 1982), the Hare Revised Psychopathy Checklist (PCL-R; Hare, 2003), the Childhood Adolescent Taxon Scale: Self Report Version (CATS-SR; Harris, Rice & Quinsey, 1994) and the number of prison misconducts incurred during one year prior to release.

Acute situational factors. The Problem Survey Checklist (PSC; Brown & Zamble, 1998a) was used to assess seven acute situational factors: marital/family, employment, accommodations, finances, leisure, interpersonal conflict and physical/mental health.

Acute emotional and cognitive appraisals. The following measures assessed this domain: the Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983), the Perceived Problem Index (PPI; Zamble, 1998), and a slightly modified version of the Positive Affect Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988).

Stable Response Mechanisms. This domain included seven constructs: substance abuse, supervision compliance, coping ability, social support, criminal associates, and criminal attitudes. The Problem Survey Checklist (Brown & Zamble, 1998a) measured substance abuse and supervision compliance. Coping was assessed in the institution using the Coping Situations Questionnaire (CSQ; Zamble, 1989) while the Coping Interview (CI; Zamble & Porporino, 1988) was used for the community-based assessments. The Social Support Scheme (SSS; Brown & Zamble, 1998b) measured both social support and criminal associates. The construct of criminal attitudes was assessed using two measures: 1) the Criminal Self-Efficacy Scale-15 (CSES-15; Brown, Zamble, & Nugent, 1998) and 2) a slightly modified version of the Expected Value of Crime Inventory (EVC; Harris, 1975)

that assessed the anticipated positive and negative consequences of crime. Detailed information regarding the development, reliability and validity of each measure is available in Brown (2002b).

Procedure

A four-wave prospective panel design was utilized. Static and dynamic variables were first assessed within 45 days of release from prison. This assessment was labeled Wave 1. Waves 2, 3, and 4 occurred in the community after the offender had been released occurring at one, three, and six month post-release intervals, respectively. Trained graduate level research assistants collected the data.

Cox regression survival analysis with time dependent covariates was used to ascertain which static and dynamic factors predicted recidivism (Allison, 2000). X¹Beta scores generated as a result of the survival analyses were saved and analyzed using ROC analysis to ascertain the relative predictive accuracy of each model (see Brown, 2002b for further details).

Results

Data Screening

The following variables were dropped from further analysis due to significant missing data at Wave 4 (e.g., >40%) or poor reliability: interpersonal conflict, supervision compliance, perceived stress, negative affect, positive affect, social support and criminal self-efficacy.

Recidivism

The follow-up period ranged from 5.9 to 7.3 years ($M = 6.6$; $SD = 0.33$). During this time 68.2% of the sample was either revoked while under parole supervision or was convicted of a new criminal offence post warrant expiry. The exact timing of recidivism was as follows: 19.1% recidivated between Wave 1 (pre-release) and Wave 2 (1 month post-release); 10.1% recidivated between Wave 2 (1 month post-release) and Wave 3 (3 months post-release); 9.6% recidivated between Wave 3 and Wave 4 and 29.3% recidivated after Wave 4.

Predictive Validity

To test the study hypothesis the following analytic steps were followed. First, each of the 18 predictor variables was analyzed individually using Cox regression survival analysis. Variables that were significant at this stage ($p < .05$) were then entered into one of three Cox stepwise regression analyses. The first Cox stepwise regression included the significant static subset (only age at release was excluded due to univariate non-significance). The second Cox stepwise regression analysis included nine of the original twelve Time 1 dynamic factors (accommodations, finance, and leisure were deleted due to non-significance). The final Cox regression analysis included eight of the original twelve Time Dependent dynamic factors. Variables that remained significant during the stepwise analyses were retained. See Table 1 for an overview of the variables that contributed to each final model.

X¹Beta scores were then outputted for the following models: 1) static, 2) combined static and wave 1 dynamic, 3) time dependent, and 4) combined static and time dependent. To test the relative predictive merit of static versus dynamic predictors, X¹Beta scores were then used to generate AUC values corresponding to each model (see Table 2). Importantly, partial support for the hypothesis was obtained given that the absolute AUC value associated with the combined static and time dependent dynamic model was higher (AUC = .89) than the absolute value of the static model AUC (.84). One could argue however that the presence of overlapping confidence intervals (although marginal) renders the differences inconsequential.

Discussion

This four wave prospective panel study investigated the extent to which the systematic

Table 1. Cox Stepwise Regression Survival Results: Model Comparisons

Model	Standardized Beta	% change in the hazard rate (standardized) ^a	Chi square
<i>Best Static</i>			
SIR-R1 ^b	-.56	-42.65	19.10***
# of prison misconducts	.25	27.94	6.71**
PCL-R ^c	.22	24.89	4.21*
<i>Best Wave 1 Dynamic</i>			
Perceived problem level	.36	43.33	12.76***
Substance abuse	.21	1.53 ^d	4.24*
Criminal associates	.21	1.55 ^d	4.62*
Strong marital support	-.21	0.65 ^d	4.23*
Good physical/mental health	-.29	-25.17	9.60*
<i>Best Time Dependent Dynamic</i>			
Positive employment	-.41	-33.64	35.49***
Strong marital support	-.37	0.46 ^d	19.31***
Perceived negative crime outcomes	-.49	-31.20	4.27*
Good physical/mental health	-.23	-30.99	5.49*

Note. *** $p < .001$; ** $p < .01$; * $p < .05$; ^athis value represents the percentage change in the hazard rate for an increase of one standard deviation in the variable while holding all other variables in the model constant; ^bSIR-R1 = Statistical Information on Recidivism Scale; ^cPCL-R = Hare Revised Psychopathy Checklist; ^dThese variables are dichotomous thus these values represents relative hazard risk rather than % change in hazard rate (i.e., individuals with criminal associates are 1.5 times more likely to recidivate than individuals without criminal associates while holding all other variables in the model constant).

Table 2. ROC Results: Comparisons of Predictive Models

Model	Area Under the Curve (AUC)	Confidence Interval
Static	.84	.77 - .89
Static & Wave 1 Dynamic	.87	.81 - .92
Time Dependent Dynamic	.82	.74 - .88
Static & Time Dependent Dynamic	.89	.84 - .93

assessment and re-assessment of dynamic risk could add incrementally to static risk estimates in a sample of 157 male offenders released from federal prisons located in Ontario. The results indicated that the strongest static predictors were the Statistical Information on Recidivism Scale, the Hare Revised Psychopathy Checklist, and the number of prison misconducts incurred during one year prior to release. Notably, the strongest time dependent dynamic factors included employment, marital support, criminal attitudes, and surprisingly a composite measure of physical and mental health. Importantly, the ROC analyses revealed that the combined static and time dependent model outperformed the strongest static model. However, the confidence intervals did overlap to a small degree.

This study is not without limitations including small sample size, missing data and the use of relatively infrequent assessment waves—a factor that most likely minimized the predictive potential of rapidly changing acute factors. Nonetheless, this research does justify cautious optimism in favour of dynamic risk in the ongoing debate regarding the relative efficacy of static versus dynamic assessment.

References

- Allison, P.D. (2000). *Survival analysis using the SAS system: A practical guide* (4th ed.). Cary, North Carolina: SAS Institute Inc.
- Andrews, D.A. (1995). The psychology of criminal conduct and effective treatment. In J. McGuire (Ed.), *What works: Reducing reoffending: Guidelines from research and practice* (pp. 35-62). New York: Wiley.
- Bonta, J. (1996). Risk-needs assessment and treatment. In A.T. Hartland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks: Sage Publications.
- Brown, S.L. (2002a). The dynamic prediction of criminal recidivism: A three-wave prospective study. *Forum on Corrections Research, 14*(1), 24-27.
- Brown, S.L. (2002b). *The dynamic prediction of criminal recidivism: A three-wave prospective study*. Unpublished Doctoral Dissertation. Queen's University, Kingston, ON.
- Brown, S.L., Zamble, E., & Nugent, P. (1998). *Criminal Self-efficacy Scale-15 (CSES-15)*. Unpublished Test. Queen's University, Kingston, Ontario.
- Brown, S.L., & Zamble, E. (1998a). *Problem Survey Checklist*. Unpublished Test. Queen's University, Kingston, Ontario.
- Brown, S.L., & Zamble, E. (1998b). *Social Support Scheme*. Unpublished Test. Queen's University, Kingston, Ontario.
- Brown, S.L., Zamble, E., & Nugent, P. (1998). *Criminal Self-efficacy Scale-15 (CSES-15)*. Unpublished Test. Queen's University, Kingston, ON.
- Cohen, S., Kamarck, T., Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behaviour, 24*, 385-396.
- Douglas, K.S., & Skeem, J.L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law, 11*, 347-383.
- Hanson, R.K. (2005). Stability and change: Dynamic risk factors for sexual offenders. In W. L. Marshall, Y. M. Fernandez, L. E. Marshall, & G. A. Serran (Eds.), *Sexual offender treatment: Issues and controversies* (pp. 17-31). West Sussex, UK: John Wiley.
- Hanson, R.K., & Harris, A.J.R. (2000). Where do we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior, 27*, 6-35.
- Hanson, R.K., Harris, A.J.R., Scott, T.L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The dynamic supervision project*. Ottawa, Canada: Public Safety Canada.
- Hare, R.D. (2003). *Hare Psychopathy Checklist-Revised (PCL-R)* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.
- Harris, A.R. (1975). Imprisonment and the expected value of criminal choice. *American Sociological Review, 40*, 71-87.
- Harris, G.T., Rice, M.E., & Quinsey, V.L. (1994). Psychopathy as a taxon: Evidence that psychopaths are a discrete class. *Journal of Consulting and Clinical Psychology, 62*, 387-397.

- Motiuk, L.L. (1999). Assessment methods in corrections. In P.M. Harris (Ed.), *Research to results: Effective community corrections: Proceedings of the 1995 and 1996 conferences of the International Community Corrections Association (ICCA)* (pp. 171-190). Lanham, MD: American Correctional Association.
- Nuffield, J. (1982). *Parole decision-making in Canada: Research toward decision guidelines*. Ottawa, Ontario, Canada: Communications Division.
- Philipse, M.W.G., Koeter, M.W.J., van der Staak, C.P.F., & van den Brink, W. (2006). Static and dynamic patient characteristics as predictors of criminal recidivism: A prospective study in a Dutch forensic psychiatric sample, *Law and Human Behaviour*, 30, 309-327.
- Quinsey, V.L., Book, A., & Skilling, T.A. (2004). A follow-up of deinstitutionalized men with intellectual disabilities and histories of antisocial behaviour. *Journal of Applied Research in Intellectual Disabilities*, 17, 243-253.
- Quinsey, V.L., Coleman, G., Jones, B., & Altrows, I. (1997). Proximal antecedents of eloping and reoffending among supervised mentally disordered offenders. *Journal of Interpersonal Violence*, 12, 794-813.
- Quinsey, V.L., Jones, B.G., Book, A.S., & Barr, K.N. (2006). The dynamic prediction of antisocial behavior among forensic psychiatric patients: A prospective field study. *Journal of Interpersonal Violence*, 21, 1539-1565.
- Rice, M.R. (2007, June). *Current status of violence risk assessment: Is there a role for clinical judgment?* Keynote address presented at the 1st meeting of North American Correctional and Criminal Justice Psychology Conference, Ottawa, Ontario.
- Watson, D., Clark, L.A., & Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Zamble, E. (1989). *Coping Situation Questionnaire*. Unpublished Test. Queen's University, Kingston, ON.
- Zamble, E. (1998). *Perceived Problem Index*. Unpublished Test. Queen's University, Kingston, ON.
- Zamble, E., & Porporino, F.J. (1988). *Coping, behaviour, and adaptation in prison inmates*. New York: Springer-Verlag.
- Zamble, E., & Quinsey, V.L. (1997). *The process of criminal recidivism*. Cambridge: Cambridge University Press.

Adapting the Sexual Offender Need Assessment Rating into a Self-Report Measure (SONAR-SR)

Michael Chajewski & Keith A. Markus
John Jay College of Criminal Justice
E-mail: chajewski@fordham.edu

It is estimated that 1-2% of the male population in the United States will be convicted of a sexual crime over the course of their lifetime (Marshall, 1997; Hanson & Morton-Bourgon, 2005). In an effort to reduce the concentration of active sexual predators in our communities, risk classifications were developed in order to improve risk assessment methods by objectively predicting risk levels.

Predictors of sexual recidivism, such as sexual deviancy, antisocial orientations, sexual preoccupation, impulsivity, intimacy deficits and offense oriented attitudes, are considered among the staples of risk evaluations (Hanson, 2000; Hanson & Morton-Bourgon, 2005). Current assessment tools by Hanson and Harris (such as the Stable-2007 and Acute-2007) measure dynamic factors that the authors believe to be key in making informed judgment about sex offenders' changes in risk levels.

Guided by Social-Cognitive Theory (Bandura, 1986) Hanson and Harris (2000) developed a structured interview incorporating 9 areas of sex offender need. Of these 9 factors, 5 factors were designated as dynamic-stable factors (factors changing slowly over time) and 4 as dynamic-acute (those changing more rapidly). Intimacy deficits, social influence, attitudes, sexual self-regulation and general self-regulation made up the five stable-dynamic factors; substance abuse, negative mood, anger/hostility, and opportunity for victim access were the remaining four acute-dynamic factors.

The SONAR (Hanson & Harris, 2000) is typically administered by the probation officer or mental health professional. Each stable-dynamic scale on the SONAR receives a score of 0 to 2, while each acute-dynamic scale is scored from -1 to 1. The attempted exploratory adaptation was to create a self-report measure capable of soliciting structured information pertaining to potential recidivism in a time and cost effective manner.

The objective of this research was to analyze the Sexual Offender Need Assessment Rating (SONAR; Hanson & Harris, 2000, 2001) for subject-specific content and to create a self-report dynamic risk factor tool version (SONAR-SR).

Method

Participants

Participants were recruited from two distinct sources. The research, conducted at John Jay College of Criminal Justice, The City University of New York (CUNY), utilized 107 (52.4%) students in the school's research participation pool, consisting of students enrolled in a Psychology 101 course, as well as 97 (47.6%) community members, solicited through the internet. Ages of participants (all male) ranged from 18 to 65, with a $M = 27.0$ ($Mdn = 23$, $Mode = 19$). Of our 204 participants, 36.3% identified themselves as Caucasian, 17.2% as African American, 28.9% as Hispanic, 8.8% as Asian American and 8.8% as Other.

Materials and Apparatus

The SONAR-SR consists of 170 items, of which 116 items compose the SONAR-SR (SONAR resembling; 9 risk factor scales). All scale items were scored by participants on a scale from 1 to 9. The focus during tool construction was on maintaining the integrity of conceptual factors used in the SONAR, rather than constructing new, unexplored, sexual offender risk predictive factors.

The adaptation of items was done by reviewing the original article published by Hanson and Harris (2000) as well as the structural interview guide of the SONAR and the scoring sheet used to score the SONAR protocols. When creating items for the SONAR-SR, one of four methods was used to adopt the structural framework of the original measure. The methods used were the following: direct adaptation, interpolating items,

newly constructed items which conceptually followed the orientation of the SONAR scale, and exploratory new items.

Demographic and exploratory new items were used for two purposes: 1) to establish a better understanding of the demographic distribution and characteristics of the sample on which the SONAR-SR was evaluated, and 2) to create additional experimental scales.

In total, the SONAR-SR contains 56 directly adapted items, 38 interpolated items, 22 new scale items, 27 exploratory new items and 27 demographic items (see Table 1). Cumulatively 116 items constitute the SONAR-SR risk scales. The remaining 54 items, 27 exploratory new items and 27 demographic questions, compose the sub scales, and experimental validity indices.

Direct adaptation was utilized whenever the structured interview guide or the scoring sheet

provided vocabulary that was suitable for a direct transformation into a self report question. For example, Hanson and Harris state that the primary designation for scoring risk factor 1 is to determine whether the offender is in a relationship. The appropriate vocabulary was then subsequently used to form the questions “Are you in a relationship?”

Interpolated items were directly based on the narratives and descriptions in the interview guide and scoring sheet. The difference between directly adapted items and interpolated items is that no exact wording was provided. If the conceptual aim for a scale was clear, but with little or no usable vocabulary, items could be created by interpreting the narratives into a self report question, extracting those particulars delineated by Hanson and Harris.

The second scale, social influence, is a good

Table 1. Number of SONAR-SR items for the SONAR 9 scales

SONAR scales	Directly adapted items	Interpolated items	New items		Total
			Scale items	Experimental items	
Intimacy deficits	8	1	2	3	14
Social influences	0	1	0	7	8
Attitudes ^a					
Rape attitudes	5	3	5	1	14
Child molesting attitudes	4	3	2	0	9
Sexual self-regulation ^a					
Sexual entitlement	4	2	4	1	11
Sexual preoccupation	8	5	4	2	19
General self-regulation	6	12	3	9	30
Substance abuse	2	2	1	2	7
Negative mood	9	1	0	0	10
Anger / Hostility	6	3	1	0	10
Victim access	4	5	0	2	11
Demographic items	0	0	0	27	27
Total	56 32.9%	38 22.4%	22 12.9%	54 31.8%	170 100%

a. Both Attitudes and Sexual Self-Regulation consist of two sub-categories, each scored separately.

example of the interpolation method. Hanson and Harris describe that the probation officer should “name all people in the offender’s life who are not paid to be with him. For each one, is the influence positive, negative or neutral?” (Hanson, & Harris, 2000, p. 20). The interpolated question pertaining to the second scale asked participants to list, by relations only, all individuals from their immediate social circle, with whom they have had continues contact in the past 12 months (excluding paid professionals). For each individual, participants were also asked to indicate the type of influence that individual had on them (from 1-positive to 9-negative).

New scale items, which were in line with the proposed aims of each SONAR scale, were counted as a part of the SONAR-SR scorable protocol. Experimental new items, which have a common conceptual ground with the SONAR scale for which they were created, were added into the measure in order to provide additional scales and indices within the tool. These items, however, were not added into the SONAR-SR final score.

The SONAR-SR can be scored by utilizing the structured interview guide and scoring sheet associated with the original SONAR. Hanson and Harris’s guidelines for determining the participant’s risk factor score (risk classification) have been faithfully adapted into the main structure of the SONAR-SR. Scales do not overlap in items, i.e., items belonging to one scale will not add or subtract from the composite score achieved on another scale in the tool.

Design and Procedure

The procedure for administering the survey packet was the same for both the John Jay and community members groups. Upon arrival, participants were introduced to the study and presented with an informed consent. Following the completion of the survey packet (approximately 40 minutes) participants were debriefed.

Results

Given that the risk classification scale of the SONAR-SR has not been empirically adjusted to the self-report format, raw scores and their

derivatives were the primary data for the evaluation of the scales.

The first factor of the SONAR-SR (in the form of one complex item) assessed the individual’s relationship status. Over fifty percent (51.0%) of the men in the study reported being single, while 24.0% reported currently seeing someone (or casually dating) and 25.0% endorsed being in a committed relationship.

The second risk factor scale was constructed like a free response section. Over eighty-eight percent (88.7%) of the participants received a Hanson and Harris scaled factor score of 0, while only 2.5% of participants received a score of 2. As expected, individuals tended to report positive influence over negative ones accounting for the high percentage of zero scores (Table 3).

Items belonging to risk factor 3 were divided according to Hanson and Harris’s stipulations into two sub categories, a) rape attitudes and b) child molesting attitudes. Subscale raw scores ranged from 13 to 91 and 9 to 54 respectively. The means were $M = 51.07$ and $M = 22.21$. Risk factor 3a scale Cronbach α was .81, $SE = .02$. The deletion of any of the 13 items would not have improved the scale’s alpha. Risk factor scale 3b assessed child molesting attitudes. The scale consists of 9 items questioning participants’ level of agreement with statements pertaining to the initiation, appropriateness, and willingness of engagement in sexual relationships with children. Items which correlated with each other (no items correlated above $r = .66$, $p < .01$) assessed similar concepts, which was expected. The Cronbach α for risk factor scale 3b was .81, $SE = .02$ (Table 2).

The fourth risk factor scale measures sexual (4a) and general (4b) self-regulation. Raw risk factor sums for 4a and 4b ranged from 10 to 90 and 17 to 121 respectively. The means of the two sub-scales were $M = 45.4$ and 59.7 . The 10 items from risk factor scale 4a produced a Cronbach α of .81, $SE = .02$. Risk factor scale 4b, sexual preoccupation, contains 17 items. Only one pair of items correlated strongly, $r = .76$, $p < .01$, items 76, “Have you fantasized about raping someone,” and item 80, “how frequently do you think about rape.” Considering the nature of these two questions, high correlation between the two

Table 2. SONAR-SR Raw Risk Factor Sums Scale Reliabilities^a

Scale	<i>N</i>	<i>M</i>	<i>SD</i>	Cronbach α	Alpha <i>SE</i>	<i>SE</i> of Measurement
Rape attitudes	13	51.1	17.2	.812	.019	7.46
Child molesting attitudes	9	22.2	11.1	.814	.018	4.78
Sexual entitlement	10	45.3	16.6	.809	.020	7.26
Sexual preoccupation	17	59.7	20.4	.772	.023	9.72
General self-regulation	21	78.0	27.4	.827	.017	11.41
Substance abuse	5	14.4	8.19	.737	.029	4.20
Negative mood	10	39.5	19.9	.931	.007	5.24
Anger / Hostility	10	32.7	20.3	.945	.006	4.74
Victim access	9	39.3	14.6	.813	.019	6.33

a. SONAR-SR Intimacy Deficits and Negative Social Influences are single items so internal consistency cannot be calculated

questions is to be expected. Risk factor scale 4b produced a Cronbach $\alpha = .77$, $SE = .02$.

The last of the stable-dynamic risk factors, risk factor 5, general self-regulation, ranged from 29 to 154, $M = 78.0$, $SE = 1.92$, $SD = 27.4$, the original Cronbach α for risk factor scale 5 was .77. The deletion of nine items from its scale significantly increased the scale's alpha. Deleted items' corrected total item correlations in the reliability analysis were either very low or near zero (the highest being $r = -.29$ for item 91) and correlated negatively with the scale. The modified risk factor 5 scale contained 21 items, with a Cronbach α of .83, $SE = .02$ (Table 2).

Risk factor 6 scale assesses the extent to which the individual's substance abuse has deteriorated or improved since the last assessment, in the case of this assessment as compared to 30 days earlier. Risk factor 6 raw scores ranged from 0 to 33, $M = 14.4$, $SE = .57$, $SD = 8.12$. Of the individuals partaking in the study, 51% of individuals received a Hanson and Harris scaled factor sum of

-1, 39.2% a 0, and only 9.8% a score of 1 (Table 3). Originally, the scale contained 6 items. Following the pilot and focus group, however, the scale was reduced to 5. Item 115, "Number of visits to the doctor," was determined to be vague and to contain a double meaning. The final scale contained 5 items, SONAR-SR's shortest scale, and produced a Cronbach $\alpha = .74$, $SE = .03$.

The 10 questions of the acute risk factor scale 7 correlate relatively strongly with one another, which was expected given the nature of the questions. Most items assess ideations about loneliness, depression, and stress, concepts which are most certainly related. Cronbach α for risk factor scale 7 was .93, $SE = .01$.

Risk factor 8 scale measures the participant's change in anger and hostility as compared to the previous month. The raw risk factor sum for the scale ranged from 0 to 76, $M = 32.7$, $SE = 1.42$, and $SD = 20.3$. 46.6% of the respondents had a Hanson and Harris scaled factor score of -1,

Table 3. Frequency of SONAR-SR Hanson and Harris Scaled Factor Scores (SFS)

	SFS				
	Score	Percent (%)	<i>M</i>	<i>SE</i>	<i>SD</i>
Intimacy deficits	0	4.4	1.71	.04	.54
	1	20.1			
	2	75.5			
Social influences	0	88.7	.12	.03	.41
	1	8.8			
	2	2.5			
Attitudes ^a	0	1.5	1.78	.03	.44
	1	17.6			
	2	80.9			
Rape attitudes	0	2.5	1.75	.03	.49
	1	20.1			
	2	77.5			
Child molesting attitudes	0	28.9	1.05	.06	.80
	1	36.8			
	2	34.3			
Sexual Self-Regulation	0	.5	1.96	.02	.23
	1	3.4			
	2	96.1			
Sexual entitlement	0	2.5	1.86	.03	.41
	1	9.3			
	2	88.2			
Sexual preoccupation	0	1.0	1.88	.03	.36
	1	10.3			
	2	88.7			
General self-regulation	0	.0	1.75	.03	.43
	1	25.0			
	2	75.0			
Substance abuse	-1	51.0	-.41	.05	.66
	0	39.2			
	1	9.8			
Negative mood	-1	44.6	-.07	.06	.90
	0	18.1			
	1	37.3			
Anger / Hostility	-1	46.6	-.27	.05	.76
	0	34.3			
	1	19.1			
Victim access	-1	34.8	.06	.06	.86
	0	26.0			
	1	39.2			

34.4% a score of 0, and 19.1% a score of 1. Of the 10 items of the risk factor 8 scale, SONAR-SR items 130, “losing your temper,” and 131, “being angry,” correlated strongly, $r = .83$, $p < .01$. The scale had a Cronbach α of .94, $SE = .01$.

The last SONAR-SR risk factor scale, risk factor scale 9, assesses the opportunities for victim

access. Such opportunities can range from communication capabilities (such as cell phone possession and internet access) to actual victim grooming habits. The raw scores ranged from 0 to 77, with a mean of 34.4, $SE = 1.03$, $SD = 14.6$. The scales 10 items produced a Cronbach $\alpha = .81$, SE of $\alpha = .02$.

Discussion

It is the ultimate goal of this line of research is to validate the risk factor selection used in the SONAR-SR and to establish a reliable and valid self-report measure for use in research and in risk evaluations. The SONAR-SR, as a newly developed measure, is currently inappropriate for the classification of risk. The current study was primarily intended to shed light on the adapted internal structure of the self-report measure. Currently, the SONAR-SR is limited in that it has only been administered to a non-sexual offender population. Calculated scores and profiles are therefore not representative of the target group for which the SONAR and subsequently SONAR-SR were designed. Future research will have to be conducted in order to norm SONAR-SR items and risk factors for use in sex offender risk evaluation.

Moreover, future research needs to compare SONAR profiles with SONAR-SR profiles. Such an endeavor in test equating is essential, not only for tool validation, but also to draw conclusions about the new measure's applicability and interpretation. Perhaps most importantly, future research should examine recidivism rates of tested individuals in order to see if the SONAR-SR has predictive utility among this population. Preferably, a sample of community members and one of released sexual offenders should be evaluated using both tools.

References

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Hanson, R. K. (2000). Will they do it again? Predicting sex-offense recidivism. *Current Directions in Psychological Science*, 9, 106-109.
- Hanson, R. K., & Harris, A. (2000). *The Sex Offender Need Assessment Rating (SONAR): A method for measuring change in risk levels*. Ottawa, Ontario, Canada: Department of the Solicitor General Canada.
- Hanson, R. K., & Harris, A. (2001). A structured approach to evaluating change among sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 105-122.

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

Marshall, P. (1997). *The prevalence of conviction for sexual offending (Research Finding No. 55)*. London: Research and Statistic Directorate, Home Office.

Section C

Interventions

Risk, Need, and Responsivity: A Heuristic for Evaluating the “Quality” of Offender Interventions

Guy Bourgon, R. Karl Hanson & James Bonta
Public Safety Canada
E-mail: guy.bourgon@ps-sp.gc.ca

In 1990, at a time when the general view was that ‘Nothing Works’ to change the behaviour of criminals, a group of researchers (Andrews, Zinger, Hoge, Bonta, Gendreau & Cullen) asked an important question. Rather than the simple question, “Does treatment work?”, they asked “When does treatment work?”. This group proposed and tested a model for why some interventions worked and others did not. It was this seminal meta-analysis that brought the clinically relevant and psychologically informed principles of Risk, Need, and Responsivity (RNR) to the field. Since that time, there has been an abundance of research demonstrating that treatment programs and services that are designed, implemented and managed to maximize adherence to these principles have greater reductions in recidivism than programs that do not and these effects hold regardless of gender, age, or type of offender (Andrews & Bonta, 2006; Bonta & Andrews, 2007; Dowden & Andrews 1999; 2000; 2003; Hanson & Bourgon, 2008). These principles of effective correctional treatment continue to provide the corrections and criminal justice field with an empirically based paradigm to assist in developing evidence-based policies and services. In addition, they provide a framework through which research efforts and empirical knowledge continue to expand. However, translating this research work into high quality ‘real world’ programs and services beyond small demonstration projects has proven difficult (Wormith et al., 2007).

When we talk about ‘quality’ treatment, what does this mean? In the criminal justice context, high quality programs are those that reduce recidivism. It is important not to lose sight of the fact that the primary concern for correctional agencies is a reduction of criminal behaviour because fewer crimes and fewer serious crimes mean fewer victims and less serious victimization. In this context, adherence to the principles of RNR is a

credible indicator of program quality given their established relationship to reducing recidivism (Andrews, 2006; 2007; Bonta & Andrews, 2007). Presently, there are two general approaches to evaluating and measuring quality or adherence to the RNR principles. One approach has an organizational focus and the other has a research or knowledge building focus. The Correctional Program Assessment Inventory (CPAI-2000: Gendreau & Andrews, 2001) embodies the organizational approach. The CPAI-2000 is a tool that gathers detailed program, staff, and agency related factors through on-site evaluations. By measuring an organization’s (or agency, or program) quality within the context of adherence to the RNR principles and other implementation issues, the CPAI-2000 assists existing programs to monitor, revise, and strengthen their programs and services. These on-site evaluations can be costly (e.g., time and resources) and, so far, they have rarely been reported in the literature. It has been the work of the group at the University of Cincinnati (Lowenkamp, Latessa, & Smith, 2006) that has demonstrated the empirical relationship between the CPAI and recidivism reductions, illustrating the importance of treatment quality in the ‘real world’.

The other method, seen in a number of meta-analyses, serves knowledge accumulation. The Carleton university group has rated literally hundreds of programs on adherence to the RNR principles. Using a coding scheme to rate adherence, this approach typically evaluates programs by reviewing available documentation (e.g., articles and program manuals). By utilizing this approach to evaluating quality, adherence to RNR principles and other core correctional practices have been linked to treatment effectiveness (Andrews & Bonta, 2006). Despite these developments, the specifics of translating each of the principles to everyday correctional practices are far from clear. Examining how these

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 45-49. Public Safety Canada.

The views expressed are those of the authors and not necessarily those of Public Safety Canada.

principles are translated in the 'real world' illustrates the challenges our field faces but can provide insight into how we can improve.

The risk principle centres on providing more intensive services to those individuals who are higher risk; it is about matching levels of service to the risk level of the offender. Although translating this principle into practice appears to be simple, in the 'real world' it is much more difficult. First, adherence to the risk principle requires valid and reliable assessments of risk. There are still correctional agencies and even more programs and services that do not employ such instruments or worse yet, pay little heed to the results of such instruments. Too often, we see policies, procedures or opinions regarding service levels that favour weak predictors of risk, such as the type of index offence or severity of violence over validated measures of risk. Factors such as motivation and compliance often play substantial roles in program admission, often resulting in lower risk offenders gaining admission over higher risk offenders. Secondly, individuals designing programs and drafting policies around treatment are faced with questions regarding what is an appropriate and effective level of service (i.e., dosage and intensity). There is very little research that provides answers to the question of intensity. One recent study, however, examined the issue of dosage with adult offenders (Bourgon & Armstrong, 2005). This study found that higher risk offenders require much more than the previously believed 100 hour minimum as set out by Lipsey in 1995. It is these factors that erode a program's quality or adherence to the risk principle.

The need principle stipulates that more effective services are those that target and primarily focus on a specific set of needs related to offending (i.e., criminogenic), such as antisocial personality, antisocial attitudes and cognitions and impulsivity. Many criminogenic needs have been outlined in the literature. Given that there is a list of what these needs are, it seems that translating the need principle into practice would be fairly simple, but this too has its challenges. One problem is that to bring about change in criminogenic needs, programs must facilitate change on a series of intermediate targets. For example, increased awareness of situations in

which an individual has cravings or urges for drugs could be considered an intermediate target. Although awareness in itself is not the targetted criminogenic need, awareness could be an important early step in learning how to appropriately manage the potential to abuse drugs. It is crucial that changes to these specific intermediate cognitive-behavioural targets are empirically linked to changes in criminogenic needs. One example of the failure to do this can be found in drug treatment courts. In these programs, substance abuse is the primary target (criminogenic need). However, upon closer examination of the services and programs, one can see that a considerable amount of resources are devoted instead to health, welfare and general social services, which have not been demonstrated to reduce substance abuse. The assumption of course is that criminal behaviour is a result of the addiction, which is a result of a lack of social-welfare support and/or opportunities.

Closely related to the problem with intermediate targets, the second challenge of adhering to the need principle is the comprehensiveness or breadth of programs. Although many offender needs are inter-related, all too often treatment programs strive to be comprehensive and holistic to ensure that they meet all the needs of the offender. It is common to view these comprehensive programs and services as better, of 'higher quality' than very focused and targeted programs and services. Such comprehensive programs may sound good on paper, but realistically diminish the focus on key criminogenic needs and increase focus on non-criminogenic needs. Marlowe (2006) specifically mentions this problem when discussing the failures of Project Greenlight, a 'real world' reentry program developed from the RNR principles (Wilson & Davis, 2006). Treatment and services may believe that their focus is on criminogenic needs; however, too often comprehensive services really means providing a little of everything.

To assist the translation of the need principle into the 'real world', emphasis must be placed on measuring and reporting changes in criminogenic needs and intermediate treatment targets. Research must establish the relationship between these intermediate targets and criminogenic needs,

and there needs to be more comparisons between comprehensive programs and narrowly focussed programs. Not enough of our attention has been paid to what goes on behind closed doors. As a result, too many programs and services sell themselves as adhering to the need principle, when in fact they do not.

The responsivity principle addresses the manner and way the services are delivered. There are two components to the responsivity principle. There is general responsivity (services should be based on a cognitive-behavioural model of change) and there is specific responsivity (services delivered in a manner that matches the styles and modes in which offenders learn). With many programs today purporting to be cognitive-behavioural, it is the translation of the general responsivity principle into practice that poses the significant challenge. We must start asking how can a cognitive-behavioural model of change be translated into a high quality integrated cognitive-behavioural program? Employing a cognitive-behavioural model to develop a program is more than simply utilizing a multitude of cognitive-behavioural interventions and employing an eclectic approach in treatment. We would suggest that a high quality cognitive-behavioural program has the following four characteristics through which we can evaluate its adherence to the general responsivity principle.

One, the program should clearly and explicitly demonstrate to the offender the causal link between thoughts and behaviour. One of the key notions to a cognitive-behavioural approach is that thoughts are a primary determinant or cause of behaviour. Implicitly or explicitly implying that external triggers (i.e., antecedent stimuli) are major determinants of behaviour does an injustice to a cognitive-behavioural approach as it provides another 'excuse' for the offender's behaviour or reaction. Cognitive-behavioural models view these outside stimuli as context to the cognitions. Like other external behaviours, these internal behaviours or cognitions are not caused by the external environment; but are a product of an individual's learning history as reflected in internal behaviours commonly referred to as attitudes, values, beliefs and expectancies. The notion of responsibility and accountability for one's behaviour must not stop at the externally

observed behaviours, but must also extend to the person's internal behaviours (thoughts, emotions). This is a crucial aspect to cognitive-behavioural interventions, noting the difference between cause of behaviour (something that is inside an individual's head) and maintenance of behaviour (both external and internally provided consequences following the behaviour). In our opinion, it is here that many of today's cognitive-behavioural programs can improve upon.

Two, after this thought-behaviour link has been learned, high quality cognitive-behavioural programs identify the individual's personal thinking patterns that relate to their criminal behaviour. Once these are identified, high quality cognitive-behavioural programs begin to tackle the strongest predictor of reoffending, antisocial attitudes and cognitions, regardless of the targeted criminogenic need. The third characteristic of high quality cognitive-behavioural programs is the teaching and modeling of concrete and easy to understand cognitive and behavioural skills. This includes teaching the skills to change both *how* and *what* an individual thinks. The fourth characteristic involves interventions and techniques that emphasize the practice of these skills to reinforce what has been learned and to enhance their generalization.

An important global factor affecting a program's adherence to the responsivity principle is the working relationship between the offender and the individual(s) delivering the program or services. There are a number of core correctional practices described by Andrews and Bonta (2006) that point to practices such as staff having a firm but fair approach, using motivational interviewing techniques, and establishing a positive interpersonal relationship with the offender. Considering the variable approaches, methods and styles with which front line staff interact and develop relationships with the offenders coupled with specific responsivity issues (e.g., gender, Aboriginal), it is not surprising that it is so difficult to ensure or measure a program's adherence to the responsivity principle. Program designers, evaluators, and researchers must make more efforts to operationalize, measure, and demonstrate how a particular program or service actually adheres to the diverse components of the responsivity principle.

Finally, much of the responsibility of moving forward must rest on the research community, particularly those who conduct program evaluations and systematic reviews of the literature. Researchers must explicitly state how and what efforts were made to adhere to these principles and provide methods that can be replicated and tested. All too often, few if any details are provided to allow the reader to evaluate the quality of a treatment program. Readers must infer the degree of adherence to the principles, rarely having ‘hard evidence’, particularly when it comes to the need and responsibility principle. If those reporting on treatment effectiveness were to include specific and explicit information of the treatment quality (e.g., how each principle is translated into practice) and include sufficient evidence that supports adherence, then we as a profession may be able to expand and refine the knowledge about what works *and* how to make it work.

In summary, almost 20 years after that first clinically relevant and psychologically informed meta-analysis, the RNR principles of effective correctional treatment continue to provide the corrections and criminal justice field with an empirically based paradigm to assist in developing evidence-based policies and services. It has also proved fruitful as a framework through which research efforts and empirical knowledge continues to expand. Translating these principles into everyday practice is presently one of our profession’s most significant challenges. In this article, it has been argued that it is our profession’s responsibility to better describe, define, measure and demonstrate how these principles can be translated to the ‘real world’ of program design, implementation, delivery and management, and why these principles of correctional treatment quality must garner greater emphasis in our research efforts. It is believed that through these types of efforts we can make our next bold step forward, speaking not just of empirically based principles and ‘text-book’ prescriptions, but to provide leadership through concrete, explicit and empirically based methods of translating these principles into practical, sustainable, and effective high quality correctional agency programs and services.

References

- Andrews, D. A. (2008). Extensions of the Risk-Need-Responsivity (RNR) model of assessment and correctional treatment. *The Proceedings of the North American Correctional and Criminal Justice Psychology Conference*. Ottawa, Public Safety Canada.
- Andrews, D. A. (2006). Enhancing adherence to risk-need-responsivity: making quality a matter of policy. *Criminology & Public Policy*, 5, pp 595-602.
- Andrews, D. A. & Bonta, J. (2006) *Psychology of Criminal Conduct. 4th Edition*. Newark, NJ: LexisNexis.
- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P. & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28: 369-404.
- Bonta, J. & Andrews, D.A. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. User Report 2007-06. Public Safety Canada, Ottawa, Canada.
- Bourgon, G. & Armstrong, B. (2005). Transferring the principles of effective correctional treatment into a “Real World” prison setting. *Criminal Justice & Behavior*, 32: 3-25.
- Dowden, C. & Andrews, D. A. (1999). What works for female offenders: A meta-analytic review. *Crime & Delinquency*, 45, 438-452.
- Dowden, C. & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*. 449-467.
- Dowden, C. & Andrews, D. A. (2003). Does family interventions work for delinquents? Results of a meta-analysis. *Canadian Journal of Criminology & Criminal Justice*. 327-342.
- Gendreau, P. & Andrews, D. A. (2001). *The Correctional Program Assessment Inventory – 2000 (CPAI 2000)*. Saint John; University of New Brunswick.
- Hanson, R. K. & Bourgon, G (2008). A psychologically informed meta-analysis of sex offender treatment outcome studies. *The Proceedings of the North American Correctional and Criminal Justice Psychology Conference*, Ottawa: Public Safety Canada.
- Lipsey, M. W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works: Reducing Reoffending* (63-78) Chichester, UK: Wiley.
- Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adherence to the principles of effective interventions. *Criminology & Public Policy*, 5, pp 575-594.
- Marlowe, D. B. (2006) When “What Works” never did: dodging the “scarlet M” in correctional rehabilitation. *Criminology & Public Policy*, 5, pp 339-346.
- Wilson, J. A. & Davis, R. C. (2006) Good intentions meet hard realities: an evaluation of the project Greenlight

reentry program. *Criminology & Public Policy*, 5, pp 303-338.

Wormith, J. S., Althouse, R., Simpson, M., Reitzel, L. R., Fagan, T. J., Morgan, R. D. (2007). The reintegration and rehabilitation of offenders: the current landscape and some future directions for correctional psychology. *Criminal Justice and Behavior*, 34, 879-892.

Does Treatment Make Psychopaths Worse? A Meta-Analytic Review

Carrie L. Tanasichuk & J. Stephen Wormith
 University of Saskatchewan, Saskatoon, SK
 E-mail: carrie.tanasichuk@usask.ca

Psychopathy is a clinical construct reflecting a constellation of affective, interpersonal, and behavioural characteristics. Affectively, psychopaths are generally short-tempered, unable to form strong emotional bonds to others, and lack empathy, guilt, and remorse. Interpersonally, psychopaths tend to be grandiose, callous, arrogant, dominant, superficial, and manipulative. Behaviourally, they are often deviant in their violation of social norms (which may or may not involve crime), irresponsible, and impulsive (Cleckley, 1976; Hare, 1996, 2006). Not surprisingly, psychopathy has predicted criminal recidivism in diverse samples of offenders (Edens, Buffington-Vollum, Colwell, Johnson & Johnson, 2002; Gretton, Hare, & Catchpole, 2004; Verona & Vitale, 2006; Walters, 2003). Psychopaths constitute approximately one percent of the general population, but approximately one quarter of the offender population (Hare, 1996). Many different types of treatment have been attempted with psychopaths, including electroconvulsive therapy (Cleckley & Beard, 1942), psychodrama (Corsini, 1958), psychoanalysis (Beacher, 1962), therapeutic communities (Rice, Harris, & Cormier, 1992), rational therapy (Ellis, 1961), medication (Kristiansson, 1995), and cognitive behavioural therapy (CBT; e.g., Catchpole, 2000). However, none of the above methods has been overly successful.

In a landmark study by Rice et al. (1992), it was found that treatment in a therapeutic community appeared to make psychopaths *worse*. When compared to a matched control group of routinely incarcerated psychopaths, treated psychopaths had higher rates of violent recidivism. However, it has been speculated that this may have been due to the kind of therapeutic community that was used and not treatment in general (Harris, Rice & Cormier, 1994). The evaluated therapeutic community involved intensive group therapy that was mainly directed by the patients themselves. Typically, psychopaths are manipulative and can

be quite charismatic when it suits their purposes. As such, psychopaths in the therapeutic community were more likely to be placed in leadership roles. Through interacting with the other offenders, it is quite possible that the psychopaths learned how to manipulate people more effectively and, consequently, became “better” psychopaths (Harris et al., 1994).

Following the often cited study by Rice et al. (1992) showing negative treatment effects, numerous studies have concluded treatment with psychopaths is ineffective, finding neither harmful nor positive treatment effects. Typically these studies have used simple binary measures of outcome concerning recidivism. However, Wong, Witte, Gordon, Gu, and Lewis (2006) found a positive treatment effect for psychopaths when the *severity* of re-offending was taken into consideration. Although treatment did not reduce the time to re-offend or the frequency of re-offending, treated psychopaths were committing less serious offences once released into the community (as measured by sentence length of the recidivistic offence).

Recently, Salekin (2002) performed a meta-analysis on 42 studies of treatment effectiveness with psychopaths. An overall positive treatment effect was found, with CBT, psychoanalysis, and eclectic therapy being the most effective. It was also found that effectiveness was related to more intense treatment (defined as four or more sessions per week). However, as outlined in Harris and Rice (2006), there are several methodological flaws in this meta-analysis. Salekin (2002) included several studies that did not use an objective measure of psychopathy, but merely stated that the patients were “psychopathic” without explaining how the authors arrived at this conclusion. Similarly, several studies did not report on a quantitatively measured outcome variable; the outcomes were mainly based on the therapist’s perceptions, which calls into question the validity of the effect sizes

used in this particular meta-analysis. Third, Salekin (2002) did not distinguish between outcomes (e.g., recidivism, hostility, employment status); all outcomes were combined and defined as “improvement.” Because of these methodological short-comings, some researchers have argued that little may be concluded from this meta-analysis (Harris & Rice, 2006). Consequently, the present meta-analysis was conducted in order to replicate and extend Salekin’s findings while employing more stringent inclusion criteria. Because many studies of treatment effectiveness compare the performance of psychopaths versus non-psychopaths, these studies were also included but analyzed separately.

Method

A computerized search was conducted in order to locate studies which examined treatment effectiveness with psychopaths. Databases searched included *PsycINFO*, *Criminal Justice Abstracts*, *Sociological Abstracts*, *Medline*, *ProQuest Dissertations and Theses*, and *Embase*. The following search terms were used: “psychopath,” “psychopathy,” “treatment,” “treatment outcome,” and “recidivism.” Additional sources of articles were references in previous studies and existing reviews (e.g., Salekin, 2002), as well as the list of psychopathy references on Hare’s (2007) public website.

Inclusion Criteria

In order to increase generalizability of the results, the inclusion criteria were kept minimal. To be included in the meta-analysis, the study must have used an objective measure of psychopathy (e.g., PCL and its derivatives, Cleckley’s criteria) and reported a quantitatively measured outcome variable. These criteria excluded studies that merely relied on a therapist’s subjective opinions of psychopathy and whether or not the patient was improving due to treatment. This was done because the use of subjective data was a serious criticism of Salekin’s (2002) meta-analysis (Harris & Rice, 2006). Furthermore, case studies were excluded. This procedure resulted in 21 studies (15 published articles, 4 theses and dissertations, and 2 conference proceedings) with

a total sample size of 5,550 participants and 50 effect sizes to be included in the meta-analysis.

Coding

All studies were coded by the first author. Information extracted from each study included background information (e.g., year/type of publication, author’s discipline), study information (e.g., sample size, measurement of psychopathy), treatment information (e.g., treatment modality, dosage), as well as type of outcome measured. Five types of outcome measures were utilized in this meta-analysis: recidivism (general, violent, and sexual), anti-social behaviour (e.g., institutional infractions, misconduct), substance use, criminal thinking and personality (e.g., hostility), and treatment behaviour (e.g., attendance, participation, effort).

Calculation of Effect Sizes

Cell frequencies were calculated and transformed into point-biserial correlation coefficients (r_{pb}). If this information was unavailable, other statistical tests (e.g., t tests) were transformed into r_{pb} . The mean effect size was calculated using a random effects model. The random effects model has the assumption that each effect size differs from the population mean by both sampling error and other sources of variability assumed to be randomly distributed (e.g., differences between setting, procedures, and methodology; Lipsey & Wilson, 2001). As such, this model is more appropriate for use with analyses of treatment effectiveness in correctional settings.

Results

Recidivism

Compared to treated non-psychopaths, treated psychopaths had a higher rate of general recidivism ($k = 7$), $r = 0.24$, 95% Confidence Interval (CI) 0.18 to 0.30, $p < .001$. When the studies with control groups of untreated psychopaths were analyzed, no treatment effects regarding general recidivism were found ($k = 3$), $r = -0.10$, 95% CI -0.23 to 0.03, $p = .13$. The rates of general recidivism among treated psychopaths were not significantly different from untreated psychopaths.

Similarly, compared to treated non-psychopaths, treated psychopaths also had a higher rate of violent recidivism ($k = 7$), $r = 0.27$, 95% CI 0.20 to 0.33, $p < .001$. When the studies with control groups of untreated psychopaths were analyzed, no treatment effects regarding violent recidivism were found ($k = 3$), $r = 0.03$, 95% CI -0.19 to 0.25, $p = .79$. The rates of violent recidivism among treated psychopaths were not significantly different from untreated psychopaths.

Regarding sexual recidivism, treated psychopaths were more likely to recidivate than treated non-psychopaths ($k = 4$), $r = 0.17$, 95% CI 0.08 to 0.26, $p < .001$. No studies could be found where rates of sexual recidivism for treated psychopaths were compared to untreated psychopaths.

Anti-Social Behaviour

Compared to non-psychopaths, psychopaths exhibited more violent and aggressive behaviour and misconduct, incurred more institutional infractions, and spent more time in seclusion ($k = 7$), $r = 0.25$, 95% CI 0.12 to 0.36, $p < .001$. These outcomes were measured both during and after treatment.

Substance Use

Compared to non-psychopaths, psychopaths were found to use more substances after treatment, ($k = 3$), $r = 0.20$, 95% CI 0.07 to 0.31, $p < .01$. This included studies that utilized urinalysis as well as one which used a self-report measure of substance use.

Criminal Thinking and Personality

Treatment did not significantly alter criminal thinking and personality among psychopaths ($k = 3$), $r = 0.05$, 95% CI -0.20 to 0.29, $p = .72$. For this analysis, studies that used pre-post measures as well as those which employed control groups were combined. Measures included the *Criminal Career Profile* (CCP; Wong, Templeman, Gu, Andre, & Leis, 2006), hostility, as well as criminal personality.

Treatment Behaviour

Compared to non-psychopaths, psychopaths spent significantly less time in treatment ($k = 5$), $r = -0.14$, 95% CI -0.23 to -0.05, $p < .01$. However, other aspects of treatment behaviour analyzed did

not differ between these two groups. There were no significant differences between attendance, participation, and effort ($k = 5$), $r = -0.17$, 95% CI -0.43 to 0.11, $p = .237$, or between clinical improvement versus no improvement ($k = 2$), $r = -0.41$, 95% CI -0.72 to 0.05, $p = .08$. However, with the small k and the near significant result, this finding merits further consideration.

Discussion

Even after treatment, psychopaths still fare worse than non-psychopaths in terms of recidivism, anti-social behaviour, and substance use. Although psychopaths attend treatment for shorter periods of time than non-psychopaths, their performance in treatment does not appear to significantly differ. Simply comparing psychopaths to non-psychopaths, however, does not inform us whether or not psychopaths are benefiting from treatment. It is possible that even though they are faring worse than non-psychopaths, it is still better than if they had not received treatment at all. In the present meta-analysis, this was not the case. There were no significant treatment effects regarding general or violent recidivism, nor for reducing aspects of criminal thinking and personality. However, contrary to the findings of Rice et al. (1992), there was no evidence to suggest that treatment made psychopaths worse.

There were several methodological limitations in the current study. The results of the analyses comparing treated psychopaths to untreated psychopaths must be interpreted with caution as the number of studies was extremely small. It is possible that an overall treatment effect would be found if more studies included control groups of untreated psychopaths. Second, this study has included only a limited search of "grey" literature (i.e., unpublished studies). This can be problematic, as relying only on published studies does not provide an accurate estimate of all research conducted (Rosenthal, 1991).

Effective Treatment

We cannot conclude that treatment in general is ineffective with psychopathic offenders. All we are able to conclude is that the treatments attempted to date do not appear to be effective. It is possible that the effective treatment method has simply not yet been implemented and/or

evaluated. It is important that researchers do not simply “throw in the towel” with this population. This therapeutic pessimism may interfere with efforts to develop and implement treatment programs designed for this population (Salekin, 2002). Not all researchers are pessimists when it comes to the treatment of psychopaths. Tennet, Tennet, Prins, and Bedford (1993) surveyed members of the Royal College of Psychiatrists (United Kingdom), members of the Criminological and Legal division of the British Psychological Society, and members of the Leicestershire Probation Service, and found that over 90% of respondents felt that psychopathy was treatable. However, the majority were unsure as to what would constitute an effective treatment program.

Guidelines for an effective treatment program have recently been put forth by Wong and Hare (2005). Briefly, these guidelines state that such treatment programs should combine cognitive-behavioural therapy with relapse-prevention. Instead of attempting to increase empathy in psychopaths, more attention should be given to convincing the offenders that there are more prosocial ways of utilizing their abilities to satiate their needs. Diligent supervision and control is crucial, both in the institution and once the offender is released into the community. Furthermore, the treatment program should be empirically evaluated to determine its effectiveness (Hare, 2006; Wong & Hare, 2006).

Another important issue related to the treatment of psychopathic offenders is treatment retention. The present study found that psychopaths spent significantly less time in treatment than nonpsychopaths. As psychopathic offenders are a high risk population, they should be receiving the most intensive treatment interventions (Andrews et al., 1990). There is some evidence to suggest that psychopaths may benefit from high dosage treatment that is delivered four or more times per week (Salekin, 2002). In order to decrease attrition, Wong and Hare (2005) suggest finding innovative and stimulating methods for keeping psychopaths interested in attending treatment, as well as improving the working alliance with the patient. Additionally, psychopaths should not be discharged from treatment for being aggressive and manipulative, as that is the often very reason

they are referred to treatment (Wong & Hare, 2005).

Directions for Future Research

As has been noted in the literature, most of the primary research in this area is simply not designed to address the effectiveness of treatment with psychopaths. Many studies are lacking appropriate control groups. Simply comparing the outcome of treated psychopathic offenders to treated nonpsychopathic offenders sheds little light on the question of treatment effectiveness with this specialized population. Beyond including matched controls in research, it is also crucial that we evaluate treatment effectiveness more carefully than with a binary outcome variable such as recidivism. Although the rates of recidivism did not differ significantly between treatment and control groups, Wong et al. (2006) found that treated psychopaths in one particular program committed less serious offences after their release than the control group. This kind of study should be replicated.

It is crucial that treatment programs are developed with the characteristics of psychopathy in mind. If these characteristics are not taken into account, the program is unlikely to be successful (Hare, 2006). This is in accordance with the *responsivity* principle of effective correctional treatment: the treatment should be tailored to the client’s need and learning style (Andrews et al., 1990). If we are to find an effective correctional treatment with this population, we must first implement programs designed for these unique individuals and evaluate such programs with methodological rigour. At this point we cannot conclude that there is a positive treatment effect for psychopaths anymore than we can conclude that there is a negative or no treatment effect for them.

References

- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- Beacher, A. I. (1962). Psychoanalytic treatment of a sociopath in a group situation. *American Journal of Psychotherapy*, 16, 278-288.
- Catchpole, R. (2000). Psychopathy and recidivism following treatment among previously violent youth. *Masters*

- Abstracts International*, 44 (01), 552. (UMI No. MR03351)
- Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Lois, MO: C. V. Mosby.
- Cleckley, H., & Beard, B. (1942). Electric shock therapy in personality disorders. *Journal of the Medical Association of Georgia*, 31, 303-308.
- Corsini, R. (1958). Psychodrama with a psychopath. *Group Psychotherapy*, 11, 33-39.
- Edens, J. F., Buffington-Vollum, J. K., Colwell, K. W., Johnson, D. W., & Johnson, J. K. (2002). Psychopathy and institutional misbehaviour among incarcerated sex offenders: A comparison of the Psychopathy Checklist-Revised and the Personality Assessment Inventory. *International Journal of Forensic Mental Health*, 1, 49-58.
- Ellis, A. (1961). The treatment of a psychopath with rational psychotherapy. *Journal of Psychology*, 51, 141-150.
- Gretton, H. M., Hare, R. D., & Catchpole, R. (2004). Psychopathy and recidivism in adolescent offenders: A ten year follow up. *Journal of Clinical and Consulting Psychology*, 72, 636-645.
- Hare, R. D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior*, 23, 25-54.
- Hare, R. D. (2006). Psychopathy: A clinical and forensic overview. *Psychiatric Clinics of North America*, 29, 702-724.
- Hare, R. D. (2007). Key references related to psychopathy and the Hare scales. Retrieved January 31, 2007 from <http://www.hare.org/references>
- Hare, R. D., & Hart, S. D. (1993). Psychopathy, mental disorder, and crime. In S. Hodgins (Ed.) *Mental disorder and crime* (pp. 104-115). Newbury Park, CA: Sage.
- Harris, G. T., & Rice, M. E. (2006). Treatment of psychopathy: A review of empirical findings. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 555-572). New York: Guilford Press.
- Harris, G. T., Rice, M. E., & Cormier, C. A. (1994). Psychopaths: Is a therapeutic community therapeutic? *Therapeutic Communities*, 15, 283-299.
- Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology*, 3, 139-170.
- Hemphill, J. F., & Hart, S. D. (2002). Motivating the unmotivated: Psychopathy, treatment, and change. In M. McMurrin (Ed.) Chichester, UK: Wiley.
- Kristiansson, M. (1995). Incurable psychopaths? *Journal of the American Academy of Psychiatry and Law*, 23, 555-562.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Rice, M. E., Harris, G. T., & Cormier, C. A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior*, 16, 399-412.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Beverly Hills, CA: Sage.
- Salekin, R. T. (2002). Psychopathy and therapeutic pessimism: Clinical lore or clinical reality? *Clinical Psychology Review*, 22, 79-112.
- Tennet, G., Tennet, D., Prins, H., & Bedford, A. (1993). Is psychopathic disorder a treatable condition? *Medicine, Science, and Law*, 33, 63-66.
- Verona, E., & Vitale, J. (2006). Psychopathy in women: Assessment, manifestations, and etiology. In C. J. Patrick (Ed.) *Handbook of psychopathy* (pp. 415-436). New York: Guilford Press.
- Walters, G. D. (2003). Predicting institutional adjustment and recidivism with the Psychopathy Checklist Factor Scores: A meta-analysis. *Law and Human Behavior*, 27, 541-557.
- Wong, S. C. P., & Hare, T. D. (2005). *Guidelines for a psychopathy treatment program*. Toronto, ON: Multi-Health Systems.
- Wong, S. C. P., Templeman, R., Gu, D., & Andre, G., & Leis, T. (1996). *Criminal Career Profile: A quantitative index of past criminal convictions*. Saskatoon, Canada: Regional Psychiatric Centre (Prairies).
- Wong, S. C. P., Witte, T. D., Gordon, A., Gu, D., & Lewis, K. (2006). Can a treatment program designed primarily for violent risk reduction reduce recidivism in psychopaths? *Canadian Psychology*, 47, 211.

A Psychologically Informed Meta-Analysis of Sex Offender Treatment Outcome Studies

R. Karl Hanson & Guy Bourgon
Public Safety Canada
E-mail: karl.hanson@ps-sp.gc.ca

Does treatment work for sexual offenders? The debate in the scientific literature remains divided, with some reviews concluding that psychological treatment reduces the recidivism risk of sexual offenders (Hall, 1995; Gallagher, Wilson, Hirschfield, Coggeshall, & MacKenzie, 1999; Hanson et al., 2002; Lösel & Schmucker, 2005) whereas other reviews concluding that it does not - or that there is insufficient evidence (Furby, Weinrott, & Blackshaw, 1989; Harris, Rice, & Quinsey, 1998; Kenworthy, Adams, Brooks-Gordon, & Fenton, 2004; Rice & Harris, 2003).

Typical of the supportive reviews is the meta-analysis conducted by Lösel and Schmucker (2005), who compared the recidivism rates of 9,512 treated sexual offenders to 12,669 untreated sexual offenders. They concluded that there was a positive treatment effect on sexual and other recidivism, and that cognitive-behavioural programs were more effective than other psychosocial approaches. In contrast, Kenworthy et al.'s (2004) review of nine random-assignment studies concluded that "the ethics of providing this still-experimental treatment to a vulnerable and potentially dangerous group of people outside of a well-designed evaluative study are debatable" (p. 2). It is important to note, however, that there was little overlap in the studies considered in these reviews. Most of the Kenworthy et al. studies did not examine recidivism. All reviews have concluded that more and better studies are needed.

The treatment of sexual offenders can be considered a special case of the treatment of offenders in general. The human service interventions that are most effective for general offenders are those that follow the principles of risk, need and responsivity (Andrews, Bonta & Hoge, 1990; Bonta & Andrews, 2007; Bourgon, Hanson & Bonta, this issue). Briefly stated, treatments are most likely to be effective when they treat offenders who are likely to reoffend (moderate or higher risk), target characteristics

that are related to reoffending (criminogenic needs), and match treatment to the offenders' learning styles and cultures (responsivity). The same results have been observed for high quality random assignment studies and non-random assignment studies, and the basic pattern of results have been replicated through meta-analyses by independent groups (Andrews & Bonta, 2006; Landenberger & Lipsey, 2005; Wilson, Bouffard & Mackenzie, 2005). The basic question addressed in the current meta-analysis is whether the same principles apply to the treatment of sexual offenders.

Much of the debate concerning the effectiveness of treatment for sexual offenders has concerned the quality of the available research studies. Few studies have used strong research designs (i.e., random assignment) and even fewer of these have examined interventions consistent with contemporary standards. Consequently, reviewers are forced to consider which of the less-than-ideal studies are good enough. For the current review, decisions concerning study quality were based on the guidelines of the Collaborative Data Outcome Committee (2007a, 2007b; Helmus, this issue).

Of the 118 studies we found that examined the recidivism rates of a treated group of sexual offenders to a comparison group, only 27 met the minimum criteria for study quality: 91 rejected, 22 weak, 4 good and 1 strong. The review used 24 studies that examined sexual recidivism as the outcome criteria (one study was excluded as it used only general recidivism), examined adult or adolescent offenders (one study was excluded as it examined children), and compared offenders assigned to a treatment program to offenders who received no treatment or a treatment that was expected a priori to be inferior (one study was excluded as it examined two treatments expected to be equivalent). These studies examined treatments delivered between 1966 and 2004, and the reports were produced between 1983 and 2006 (50% published, 23 English, 1 French; country of

origin: Canada – 13, USA – 5, England – 3, New Zealand – 2 and Netherlands – 1). The treatment was given either in institutions (12) or the community (11) or both (1).

On average the recidivism rates for the treated offenders were lower than the recidivism rates for the comparison groups (odds ratio of .72, 95% confidence interval of .62 to .84, $k = 24$; $n = 7,751$). There was, however, significant variability in the findings across studies, $Q = 59.58$, $p < .001$. The 4 strongest research designs indicated no overall effect of treatment (odds ratio of .99, 95% confidence interval of .78 to 1.27, $k = 4$; $n = 1,542$), but there was significant variability across studies: $Q = 7.46$, $p < .01$. Some of the treatment programs examined in the strongest research designs would have been expected to have little impact (e.g., unstructured self-help groups; Meyer & Romero, 1980).

In order to examine treatment quality, each intervention was rated by the two authors for the extent to which it conformed to the principles of risk/need/responsivity. The interventions were primarily cognitive-behavioural and made reasonable efforts to engage offenders in treatment. Consequently, most were rated as meeting the responsivity principle (19 out of 24). Half the programs (12 of 24) addressed criminogenic needs (e.g., impulsivity, sexual preoccupations, intimacy deficits) more than non-criminogenic needs (e.g., victim empathy, denial, remorse). Only 3 of the 24 programs explicitly focussed their interventions on higher risk offenders (risk principle).

The extent of adherence to the principles of risk/need/responsivity was strongly related to the observed treatment effects (see Table 1; $r = .45$, $df = 23$, $t = 3.45$, $p < .01$, using fixed effect weighted regression, Hedges, 1994). For the programs adhering to at least two of the principles, there was a significant reduction in the sexual recidivism rates (odds ratio of .57, 95% C.I. of .46 to .70) with significant variability across the programs ($Q = 32.67$, $k = 13$, $p < .001$). The programs following none or one of the principles consistently had no effect (odds ratio of .97, 95% C.I. of .77 to 1.23, $Q = 15.64$, $k = 11$, $p > .10$). There was also evidence that current programs are

more effective than older programs ($r = .49$, $df = 21$, $t = 3.65$, $p < .002$).

Table 1. Reductions in sexual recidivism rates for interventions according to the extent that they conformed to the principles of risk, need and responsivity

	Odds ratio	95% C.I.	N (k)
Not at all	1.05	.79 to 1.41	1,200 (4)
One principle	0.84	.57 to 1.25	1,298 (7)
Two principles	0.57	.46 to .70	5,147 (12)
All three principles	0.51	.13 to 1.97	106 (1)

These results suggest that the same “what works” principles for general offenders should also guide interventions with sexual offenders. This does not mean that all offenders should receive the same interventions. Sexual offenders would be expected to have important criminogenic needs not shared with other offenders (e.g., sexual preoccupations, deviant sexual interests) (Hanson & Morton-Bourgon, 2004, 2005) and specialized scales may be needed for sexual offender risk assessment (Hanson & Morton-Bourgon, 2007). The findings do suggest that those concerned with the treatment of sexual offenders could benefit from carefully considering the much larger literature on the treatment of general offenders. In addition to considering the risk, need, and responsivity principles, the general offender literature indicates the importance of selecting staff based on relationship skills, using treatment manuals, training staff, and starting small (Andrews & Bonta, 2006, chapter 10).

References⁵

- Andrews, D.A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Cincinnati: Anderson.
- Andrews, D.A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19-52.

⁵ Studies included in meta-analysis available upon request.

- Bonta, J., & Andrews, D.A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation*. Corrections Research User Report 2007-06. Ottawa: Public Safety Canada.
- Bourgon, G., Hanson, R. K., & Bonta, J. (2008). Risk, need, responsivity: A heuristic for evaluating the "quality" of offender interventions. *Proceedings of the North American Correctional and Criminal Justice Psychology Conference*. Public Safety Canada.
- Collaborative Outcome Data Committee. (2007a). Sexual offender treatment outcome research: CODC guidelines for evaluation: Part 1: Introduction and overview. Corrections Research User Report 2007-02. Ottawa: Public Safety Canada.
- Collaborative Outcome Data Committee. (2007b). Sexual offender treatment outcome research: CODC guidelines for evaluation: Part 2: CODC Guidelines. Corrections Research User Report 2007-02. Ottawa: Public Safety Canada.
- Furby, L., Weinrott, M.R., & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin*, *105*, 3-30.
- Gallagher, C. A., Wilson, D. B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual reoffending. *Corrections Management Quarterly*, *3*, 19-29.
- Hall, G.C.N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology*, *63*, 802-809.
- Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J. K., Murphy, W., Quinsey, V. L., & Seto, M. C. (2002). First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment of sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, *14*, 169-194.
- Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. Corrections User Report No. 2004-02: Public Safety and Emergency Preparedness Canada.
- Hanson, R.K., & Morton-Bourgon, K.E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73*, 1154-1163.
- Hanson, R.K., & Morton-Bourgon, K.E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis*. Corrections User Report No 2007-01. Ottawa: Public Safety and Emergency Preparedness Canada.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1998). Appraisal and management of risk in sexual aggression: Implications for criminal justice policy. *Psychology, Public Policy, and Law*, *4*, 73-115.
- Hedges, L. V. (1994). Fixed effects models. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 285-299). New York: Russell Sage Foundation.
- Helmus, L. (2008). Structured guidelines for evaluating study quality. *Proceedings of the North American Correctional and Criminal Justice Psychology Conference*. Public Safety Canada.
- Kenworthy, T., Adams, C.E., Brooks-Gordon, B., & Fenton, M. (2004). Psychological interventions for those who have sexually offended or are at risk of offending (Cochrane Review). Cochrane Library, Issue 3. Chichester, UK: John Wiley & Sons.
- Landenberger, N.A., & Lipsey, M.W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, *1*, 451-476.
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, *1*, 117-146.
- Meyer, L.C., & Romero, J. (1980). *A ten year follow-up of sex offender recidivism*. Joseph J. Peters Institute. Pennsylvania, PA.
- Rice, M. E., & Harris, G. T. (2003). The size and sign of treatment effects in sex offender therapy. *Annals of the New York Academy of Sciences*, *989*, 428-440.
- Wilson, D.B., Bouffard, L.A., & Mackenzie, D.L. (2005). A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Criminal Justice and Behavior*, *32* (2), 172-204.

The Vermont Cognitive Self-Change Program: The Case for Risk-Adjusted Classification

Christine M. Sadler & Thomas A. Powell
Castleton State College
E-mail: tpowell@vtfa.com

Over the past thirty years, the correctional treatment literature has provided evidence and some measure of support for treatment programs designed to reduce offender recidivism. Successful correctional treatment programs are generally rooted in cognitive-behavioral theory (Ross & Fabiano, 1985). A study conducted in Washington State found an average 8.2% reduction in criminal recidivism when programs applied and followed cognitive-behavioral principles of change. As noted by the authors, "... even relatively small reductions in recidivism can be quite cost-beneficial... a [five] percent reduction in the reconviction rates of high risk offenders can generate significant benefits for taxpayers and crime victims" (Washington State Institute for Public Policy, 2006, p. 4). There is additional evidence that the application of certain principles of effective correctional intervention yields favorable results (Andrews, Bonta & Hoge, 1990). Specifically, the risk principle promotes reserving higher levels of service and treatment to higher risk offenders; lower risk offenders should receive fewer, if any, services. The needs principle stresses the importance of targeting criminogenic risk factors that are both changeable and related to criminal risk. Responsivity refers to the provision of treatment consistent with the learning style and abilities of the individual offender. Finally, the use of professional override underscores the importance of discretion in the application of these principles to individual cases (Andrews & Bonta, 2003a).

Regrettably, correctional agencies and departments throughout North America have not, for the most part, incorporated behavioral science research findings into actual practice (Boothby & Clements, 2000; Haney & Zimbardo, 1998). In fact, criminal justice policy has demonstrably followed a course of "knowledge destruction" by its ideologues and dismissal of the notion that corrections is a behavioral science that might benefit by adherence to empirical findings

(Andrews & Wormith, 1989). The mendacious actions of politicians toward corrections and the justice system, evident in sex offender legislation, the abolition of parole, and abandonment of rehabilitative ideals, have led to a sadly cynical conclusion: you'll never lose a vote being mean to inmates.

Against this bleak background, we see efforts by some jurisdictions to incorporate research into practice with some promising results. Once such program, the Vermont Cognitive Self-Change, was initiated twenty years ago as a cognitive-behavioral restructuring intervention for male inmates convicted of crimes of violence. It has since spread throughout Vermont's correctional system⁶, including women's institutions and probation/parole settings, and it occupies a central feature of the case planning for many offenders. Preliminary published findings have provided some degree of optimism about the program's impact on recidivism (Henning & Frueh, 1996) and institutional conduct (Baro, 1999). Longer-term outcome studies have not been conducted, although there is recent evidence of support for a comparable program model. Bourgon and Armstrong (2005) found that recidivism was attenuated when an offender's risk was matched with a low, medium or high level of a cognitive-behavioral treatment program. During a one-year follow-up period, the participants who completed the program were less likely to recidivate than a matched comparison group.

The intent of the present study was to determine if the Cognitive Self-Change program, delivered to inmates charged with violent offenses, would result in lower levels of recidivism. Specifically we were interested in whether the "dosing level" of treatment (low, medium or high) influenced

⁶ It has been adopted by many other states, and a slightly modified version is disseminated by the National Institute of Corrections. Jack Bush, Ph.D. is widely credited with its development and implementation.

offenders' rates of recidivism. Additionally, we sought confirmation of the risk principle: that higher risk inmates would benefit more from higher levels of intervention, whereas lower risk inmates would show little improvement regardless of intervention level. We predicted between 65 and 150 hours⁷ of the Cognitive Self-Change program would be the optimal "treatment dosage" after controlling for risk.

Method

The Vermont Department of Corrections (DOC)'s policy manual requires participation in the institutional phases of programs for most inmates, and completion of referred treatment programs is a contingency upon which release is primarily based. Furlough and parole decisions are heavily influenced by inmates' participation and assumed benefit in treatment. It is seen as a proxy for reduced risk of recidivism, often without reference to the presence or absence of underlying risk factors. Individual differences and professional override principles are commonly overlooked. Program referrals are based almost exclusively on the index offense. The net effect of this policy is that the vast majority of inmates sentenced to more than one year are referred to treatment groups such as Cognitive Self-Change (CSC) or the department's sex offender program, depending on the nature of the inmate's offense history. Control groups are unavailable to draw from this population, since virtually all eligible inmates spend at least some time in the program of interest.

The participants for this study were a convenience sample of 387 male inmates in Vermont correctional facilities between 2000 and 2003 who participated in the CSC program for violent offenders. Most (91.1%) were charged with a violent offense. The sample was predominantly Caucasian (90.2%) and their ages ranged from 21 to 68 years upon release. The average LSI-R total score was 26.5 ($SD = 6.678$) with a range of 7 to

46, slightly higher than the average for United States inmates (Andrews & Bonta, 2003b).

The program is designed to change criminal thinking patterns (see Yochelson & Samenow, 1977). Offenders attend groups managed by trained clinicians and prison caseworkers three times per week, participate in role-play, and keep daily journals. Through the CSC program, offenders learn how their own thinking leads to criminal behavior and recognize that crime is not the product of external forces, but rather their own thoughts, beliefs, and actions. Furthermore, offenders recognize that their thinking is within their own control and that they can redirect their thoughts toward prosocial endeavors (Powell, Bush & Bilodeau, 2001).

The Vermont DOC's correctional system is a "hybrid," wherein correctional facilities incorporate features of both jails and prisons, and community supervision includes parole, probation, and conditional re-entry inmates. The typical method of measuring recidivism in risk assessment literature (e.g. whether an offender is charged with a new crime or re-incarcerated over a one or two year follow-up period) is complicated by a number of extraneous local variables, such as brief furlough suspension practices. A more manageable and sensitive criterion measure was developed for the study: incarceration days - the number of days an offender was incarcerated, post-release, in a correctional facility. A second variable was the number of new charges an offender incurred after being released from the facility.

Results

A Chi-Square for Independence was performed to analyze the relationship between CSC dosage (low, medium, or high) and the number of charges returned, post-release (categorized into none, one charge, two charges or more charges). A significant relationship was found: $\chi^2 = (4, 384) = 10.80, p = .029$. Fewer offenders had new charges in the medium and high dosage groups (67.7% and 75.4%, respectively) than in the low dosage group (60.3%). Additionally, more offenders had two or more charges against them in the low dosage group (26.7%) than in the medium and high groups (15.4%, 15.1%) (See Table 1)

⁷ While Lipsey (1995) suggested at least 100 hours in his meta-analysis of juvenile rehabilitation, there is little research indicating the "optimal" dosage for correctional treatment programs.

Table 1. CSC Dosing and Number of New Charges, Post Release

	No New Charges	One New Charge	Two or More Charges
Low n =131	79 (60.3%)	17 (13%)	35 (26.7%)
Medium n =130	88 (67.7%)	22 (16.9%)	20 (15.4%)
High n = 126	95 (75.4%)	12 (9.5%)	19 (15.1%)

Note. $\chi^2(4, 384) = 10.804, p = .029$

The relationship between CSC dosage and recidivism was examined via a one-way between-subjects factorial MANOVA with CSC dosing level (low, medium or high) as the independent variable and number of days incarcerated during the first year, number of days incarcerated during the second year, and number of new charges as dependent variables. A type IV sum of squares was employed due to unequal cell sizes. Bivariate correlations among dependent variables ranged from .28 to .57, $p < .01$.

A significant main effect was found for CSC level; Wilks' $\Lambda = .941, F(1, 386) = 3.94, p = .001$, partial $\eta^2 = .03$. Univariate follow-ups found an effect for the first year, $F(2, 385) = 5.35, p = .005$, partial $\eta^2 = .027$ with offenders in the high dosage category spending fewer post-release days in jail ($M = 42.38, SD = 72.4$) than the medium dosage category ($M = 77.98, SD = 95.03$) (see Table 2). There was a significant effect for the second year post-release $F(2, 386) = 3.210, p = .041, \eta^2 = .016$, but post-hoc tests did not uncover differences among means.

A one-way between subjects MANCOVA was performed, controlling for LSI-R score, to examine the added contribution of offender risk to incarceration days. The covariate was significant; Wilks' $\Lambda = .952, F(3, 384) = 6.391, p < .0001$, partial $\eta^2 = .05$. LSI-R score was significant in the first year, $F(1, 386) = 18.54, p < .0001$, partial $\eta^2 = .05$, and in the second year, $F(1, 386) = 9.14, p = .003$. No significant differences were found for new charges when controlling for risk. Univariate follow-ups for CSC dosage revealed a significant main effect for the first year $F(2, 385) = 3.819, p$

Table 2. Means and Standard Deviations for Incarceration Days Post Release, during the first year, according to CSC dosage

	Number of incarceration days post-release	
	Mean	SD
Low dosage	59.64	91.69
Medium dosage	77.98 _a	95.03
High dosage	42.38 _a	72.37

Note. Subscripts denote statistically significant differences (Tukey HSD $p = .005$)

$= .023, \eta^2 = .02$. Examinations among the means found that offenders in the high dosage group ($M = 42.38, SD = 72.37$) were significantly less likely to spend time in jail than those in the medium dose group ($M = 77.98, SD = 95.03$).

Discussion

The purpose of this research was to examine relationships between dosing levels in the Cognitive Self-Change program and recidivism rates. It was hypothesized that offenders who received a moderate dosing level of CSC (65-150 hours) would spend the fewest days in jail after being released. This prediction was based on the presumption that inmates who are retained in treatment beyond 150 hours are either recidivists required to “recycle” through the program or individuals unable to proceed at the expected rate through the program for a variety of other reasons. However, findings revealed that offenders with the highest level of CSC exposure (150 hours or more) were the least likely to return to a correctional facility, albeit during the first year only. When risk level was controlled, results were unchanged. This effect was demonstrated in the first year post-release, suggesting that CSC's impact may attenuate over time, at least as the program is currently implemented in Vermont.

Because LSI-R data are not meaningfully applied to classification decisions in the Vermont DOC, there is no relationship between assessed offender risk level and treatment program assignment, including length of treatment. The present findings suggest that if the Vermont DOC is to continue using assessment tools such as the LSI-

R, it is imperative that it employ a set of procedures, including standardization of the instrument using Vermont inmate norms to determine cut-off scores for purposes of program referral, based on observed recidivism risk for the differential risk groups. It appears that present LSI-R utilization is limited to preliminary classification screening, based on an unvalidated dichotomous decision scheme, with an automatic “upward” override for any case presumed to be violent or sexual in nature by virtue of the statute violated.

Several limitations in this study are noted. There was no control or comparison group with which to contrast treatment subjects who received low, medium and high dosages of CSC due to the “universal access to programs” policy of the department. Additionally, it is unknown why some subjects in the low dosage group received a very small number of treatment hours. A variety of reasons could explain this phenomenon, including transfer or suspension from the treatment program. A further question concerns the reliability of data entry of CSC hours. The racial and ethnic demographics of this sample are not representative of most other U.S. offender populations; further research on programs such as CSC is needed with a more racially diverse offender populations (Schlager & Simourd, 2007).

The classification practices of the Vermont Department of Corrections appear to compromise the risk and override principles of effective correctional practice. The placement of lower risk offenders in a high intensity program did not yield lower incarceration days as a function of dosing level. The practice of universally assigning inmates to CSC primarily on the basis of an index offense associated with violence without regard to objective risk factors or override considerations is inconsistent with an evidence-based classification approach. It saturates overcrowded prisons with lower risk inmates who have treatment requirements that may extend their term or subject them to the potentially iatrogenic effects of prison treatment.

The systematic use of available data from objective assessment measures such as the LSI-R provides the foundation for a more individually attuned correctional rehabilitation system

(Andrews & Bonta, 1990). Administrators may benefit from considering the cost benefits, as well as the program and public safety implications, of reserving Cognitive Self-Change program slots for high-risk offenders, who appear to benefit most from this intervention.

References

- Andrews, D.A. & Bonta, J. (2003a). *The Psychology of Criminal Conduct*. (3rd ed). Anderson Publishing: Cincinnati: OH.
- Andrews, D.A. & Bonta, J.L. (2003b). *The Level of Service Inventory-Revised User's Manual: U.S. Norms Manual Supplement*. North Towanda, New York: Multi-Health Services.
- Andrews, D.A. & Bonta, J. L. (2001). *The Level of Service Inventory-Revised User's Manual*. Toronto: Multi-Health Services.
- Andrews, D.A., Bonta, J. & Hoge, R.D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 7, 19-52.
- Andrews, D.A. & Wormith, J. S. (1989). Personality and crime: Knowledge destruction and construction in criminology. *Justice Quarterly*, 6 (3), 289-309.
- Baro, A. (1999). Effects of a cognitive restructuring program on inmate institutional behavior. *Criminal Justice and Behavior*, 26, 466-484.
- Boothby, J. L. & Clements, C.B. (2000). A national survey of correctional psychologists. *Criminal Justice and Behavior*, 27 (6), 716-732.
- Bourgon, G. & Armstrong, B. (2005). Transferring the principles of effective treatment into “real world” prison setting. *Criminal Justice and Behavior*, 32, 3-25.
- Haney, C & Zimbardo, P. (1998). The past and future of U.S. prison policy: Twenty-five years after the Stanford Prison Experiment. *American Psychologist*, 53 (7), 709-727.
- Henning, K.R. & Frueh, C. (1996). Cognitive-behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Corrections Cognitive Self-Change program. *Criminal Justice and Behavior*, 23, 523-541.
- Lipsey, M.W. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), *What works; reducing reoffending* (pp. 63-78). Chichester, UK: Wiley.
- Powell, T., Bush, J., & Bilodeau, B. (2001). Vermont's Cognitive Self-Change Program: A 15 year review. *Corrections Today*, 63, 116-120.
- Ross, R. & Fabiano, E. (1985). *Time to think: A cognitive model of delinquency prevention and offender rehabilitation*. Institute of Social Sciences and Arts, Ince: Johnson City: TN.

Schlager, M.D. & Simourd, D.J. (2007). Validity of the Level of Service Inventory-Revised (LSI-R) among African American and Hispanic male offenders. *Criminal Justice and Behavior*, 34, 545-554.

Washington State Institute for Public Policy (2006). Evidence-based adult corrections programs: What works and what does not. Olympia, WA: Author.

Yochelson, S. & Samenow, S. (1977). *The Criminal Personality: The Change Process*. Northvale, NJ: Jason Aronson.

Prediction of Program Dropout in a High-Intensity Violent Offender Program in Canada

Renée Gobeil & Ralph Serin
Carleton University
E-mail: gobeil@gmail.com

In the last three decades, a large body of literature has accumulated supporting the utility of evidence-based correctional programs (e.g., Andrews & Bonta, 2006). By targeting offenders' criminogenic needs, correctional programs can lead to reductions in dynamic risk, and ultimately, in rates of recidivism. Given this body of literature, it is not surprising that there is ample evidence that rates of recidivism are higher for treatment non-completers than for completers across a range of correctional programs, including domestic violence programs and sex offender programs (e.g., Buttell & Pike, 2002; Hanson et al., 2002).

It is unclear, however, whether differential recidivism rates are attributable to underlying differences between completers and non-completers, or whether it is program participation, in and of itself, which explains this difference. The truth is likely a combination of the two, in that some non-completing offenders have the potential to benefit from program participation, while some would benefit less.

Regardless of where exactly the balance between these two possibilities lies, it is clear that there are important consequences associated with program attrition. In addition to the higher rates of recidivism typical of those who drop out or are expelled from programs (e.g., Hanson et al, 2002), program attrition has other consequences. When offenders fail to complete programs to which they have been assigned, the cost per completing offender increases; this can be problematic in an era where those involved in decisions to offer or not offer programs must weigh financial considerations. Moreover, when offenders who ultimately drop out are assigned to programs, others may be denied a place in the program due to lack of space. Within this context, a full understanding of the factors associated with attrition would allow for the development of interventions aimed at decreasing program drop-

out and expulsion (e.g., Beyko & Wong, 2005).

Correlates of Attrition

An increased understanding of program attrition is of obvious value. A first step is an examination of the correlates of attrition. These correlates have been identified in previous research by contrasting program completers and their counterparts who drop out, are expelled, or who fail to complete for another reason. Much of this research has focused on sex offenders and domestic violence offenders. Commonly identified factors can be grouped in three areas: Demographic factors include age, ethnicity, socioeconomic status, education level achieved, and relationship status (e.g., Garfield, 1994; Hamberger & Hastings, 1989); risk and need variables include Statistical Information on Recidivism scores and Community Intervention Scale ratings (e.g., Nunes & Cortoni, 2006; Wormith & Olver, 2002); and responsivity variables include denial and motivation (e.g., Gondolf & Foster, 1991; Scott, 2004).

The Current Study

As previously mentioned, much of this research has been conducted with sex offender and domestic violence programs. As such, the present analysis aimed to examine the differences between program completers and non-completers of two violent offender programs, as well as to examine the extent to which any identified differences allow for the prediction of program completion. Analyses included both completers and non-completers of two programs for violent offenders, the Persistently Violent Offender program and Anger and Emotion Management.

Method

Participants

Participants were a total of 256 violent offenders who were incarcerated in medium security federal institutions. Two-hundred-and-eight offenders

completed one of the programs, with 70 completing the Persistently Violent Offender program, and a total of 138 completing Anger and Emotion Management. These offenders were contrasted with 48 offenders who failed to complete either program due to refusal, drop-out, or discharge. Of these, three quarters ($n = 35$) had originally been enrolled in the Persistently Violent Offender program.⁸

Measures

Data for this examination were originally gathered as part of an evaluation of the Persistently Violent Offender program (Serin, Gobeil, & Preston, in press). Variables of interest included background information (educational level, employment level, age, intelligence, marital status, as well as the presence of alcohol and drug problems, the number of criminals in the offender's family, and the quality of community support), motivation for treatment (Readiness for Change scale [Prochaska & DiClemente, 1983]), measures of risk (scores on the Statistical Information on Recidivism [SIR-R1; Nuffield, 1982] and on the Psychopathy Checklist Revised [PCL-R; Hare, 1991], and a diagnosis of antisocial personality disorder). Baseline scores on a number of measures used in the Persistently Violent Offender program evaluation were also used: the Aggression Questionnaire (Buss & Perry, 1992), the I-7 Impulsivity Questionnaire (Eysenck, Pearson, Easting, & Allsop, 1985), the Interpersonal Reactivity Index (Davis, 1980), and the Reactions to Provocation (Novaco, 1994).

Procedure

In order to examine the predictive ability of these measures, a series of bivariate analyses were first completed to determine which measures were significantly able to differentiate between program completers and program non-completers. Those which were identified in these analyses

were then entered into a hierarchical binary logistic regression to predict program completion status. With the rationale that those measures relating to program treatment targets – that is, those relating to anger, aggression, impulsivity, and empathy – are dynamic constructs, these were entered as a second block in the analyses. This procedure allowed for determination of the predictive ability of scores on these measures above and beyond that shared with the demographic and risk variables.

Results

A series of chi-square tests of independence and independent sample t-tests revealed that only three measures were significantly able to differentiate between program completers and non-completers: PCL-R score, $t(212) = 2.96, p < .01$; SIR score, $t(227) = 2.84, p < .01$; and impulsivity, as measured by the I-7 Impulsivity Questionnaire, $t(255) = 2.68, p < .01$.⁹ Between-group differences in aggression, as measured by the Aggression Questionnaire, $t(255) = 1.89, p = .06$, and in empathy, as measured by the Interpersonal Reactivity Index scores, $t(255) = 1.86, p = .06$, also approached significance. In each case, non-completers had more problematic ratings on each of these variables than did their counterparts who completed one of the programs. Notably, neither the background measures nor the motivation rating significantly differentiated the two groups.

These identified correlates were entered into a hierarchical binary logistic regression. The significant treatment target measures (i.e., impulsivity, aggression, empathy) were entered in a second block in order to examine whether they offered any predictive ability above and beyond that offered by the other measures. Table 1 presents the results of this analysis. Because most offenders completed the program, the percentage of cases wherein non-completion was successfully predicted was quite high even for the base model, and did not improve importantly with the two additional models. The addition of each group of predictors in this model, however, significantly

⁸ The difference in attrition for the PVO and A&EM programs is worthy of note – it should be acknowledged that because the PVO program was being piloted at the time of the study, whereas the A&EM program had been ongoing for quite some time, this difference is likely at least partially attributable to some scepticism on the offenders' part toward the program. Alternatively, it may be attributable to the characteristics of the slightly different clienteles served by each group.

⁹ There is some variability in the degrees of freedom for each test as not all offenders completed each measure. For example, the SIR scale is not administered to Aboriginal offenders.

Table 1. Logistic Regression of Program Completion on Correlates of Attrition.

Model		χ^2	Nagelkerke R^2	% Correctly Predicted
Base		-	-	91%
Model (PCL-R, SIR)	1	4.83*	.17	91%
Model (PCL-R, SIR, Impulsivity, Aggression, Empathy)	2	14.26**	.33	92%

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

increased the proportion of variance accounted for in prediction, and the final model, which included all the variables which were significant at the bivariate level, had a moderate Nagelkerke R-squared value. This finding demonstrates that the model had noteworthy predictive ability and that scores on the measures of impulsivity, aggression, and empathy improved prediction of program completion. In considering the predictors individually, however, only SIR score remained uniquely predictive above and beyond the variance shared by the other predictors, with increases in level of risk (represented by lower scores) being associated with increasing odds of program non-completion (see Table 2).

Discussion

The objectives of the present study were to identify correlates of attrition and to examine the extent to which these correlates were predictive of program completion in a sample of offenders enrolled in two treatment programs focused on violence. Altogether, analyses demonstrated relatively few significant differences between program completers and non-completers, and modest ability to predict program completion based on knowledge of background, risk, and treatment target variables.

Though it was initially surprising that no between-group differences were found with regards the demographic factors identified as correlates of attrition in previous research (e.g., Garfield, 1994; Hamberger & Hastings, 1989) – that is, on ethnicity, age, educational level and occupational level – the present sample may have been too

Table 2. Logistic Regression of Program Completion on Model 2 Variables

Predictor	<i>B</i>	SE <i>B</i>	Odds Ratio
SIR	-0.13*	0.08	0.88
PCL-R	0.07	0.08	1.12
Impulsivity	0.80	0.53	2.22
Aggression	0.02	0.03	1.02
Empathy	-0.71	0.46	0.49
Constant	-3.93*	2.37	-

Note. Reference category in calculating odds ratios is “non-completion”. * $p < .05$.

small to allow for identification of such factors. Moreover, a significant proportion of the literature identifying these demographic factors as predictive is American (where different ethnic proportions appear) and / or related to court-mandated programming, especially domestic violence programming. This latter point may be particularly important; though it has been suggested that incarcerated offenders are unable to choose not to participate in programs and otherwise comply with their correctional plans due to security classification and conditional release implications (e.g., Taylor, 2001), there is still perhaps more freedom regarding program participation among this group than among offenders for whom the choice is between program completion and incarceration. As such, those offenders who would, in a court-mandated

sample, have failed to participate, may, in the current sample, have originally refused enrolment.

Another point of difference in contrasting the present findings with previous literature is that scores on the measure of motivation for treatment were not able to differentiate between program completers and non-completers. This point was surprising, as differences in motivation have been consistently identified elsewhere (e.g., Nunes & Cortoni, 2006; Scott, 2004). Finally, however, findings relating to risk were consistent with the previous literature (e.g., Nunes & Cortoni, 2006; Wormith & Olver, 2002), in that risk, as measured by the SIR scale, was the biggest predictor of program completion of the variables considered.

One of the key findings of the present study was that scores on some of the measures of treatment targets - most notably impulsivity, but also aggression, empathy, and anger - were able to differentiate completion status. For each of these variables, it was those with the more problematic profiles which were more likely to be non-completers. This finding suggests that it is precisely those who most need the program who are most likely not to see it through to completion. Though this is consistent with the body of literature which indicates that there is considerable similarity in the variables which are predictive of attrition and of recidivism (e.g., Andrews & Bonta, 2006; Gendreau, Little, & Goggin, 1996), it is worthy of further attention. These findings suggest that determining ways to improve these and other programs in such a way as to promote greater engagement and interest from offenders would be beneficial. As has been argued elsewhere (Beyko & Wong, 2005), this approach is more ethically appropriate than choosing to deny treatment to offenders who exhibit high scores on measures associated with attrition

Similarly, the present findings demonstrate that one of the strongest predictors of program non-completion among this sample of violent offenders was high impulsivity. Though this characteristic may lead to expulsion due to inappropriate behaviour, this finding also suggests the possibility that in some cases, program drop-out may be an impulsive rather than thoroughly thought-out choice. Considering that offenders

being assigned to violent offender programming often have difficulties with emotions and self-management, this possibility seems likely. As such, it might be worthwhile to develop a procedure to approach offenders after they have dropped out to have them consider returning to the program. It is possible that some may choose to do so, and therefore potentially reap the benefits of program participation.

Altogether, the present results indicate that we were only moderately able to predict program completion. Nonetheless, we identified a number of interesting findings which have, in turn, suggested areas where development of new procedures to engage, interest, motivate, and retain offenders may lead to decreases in attrition.

References

- Andrews, D.A. & Bonta, J. (2006). *The psychology of criminal conduct (4th ed.)*. Cincinnati, OH: Anderson Publishing.
- Beyko, M.J. & Wong, S.C.P. (2005). Predictors of treatment attrition as indicators for program improvement not offender shortcomings: A study of sex offender treatment attrition. *Sexual Abuse: A Journal of Research and Treatment, 17*, 375-389.
- Buss, A.H. & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology, 63*, 452-459.
- Buttall, F.P. & Pike, C.K. (2002). Investigating predictors of treatment attrition among court-ordered batterers. *Journal of Social Service Research, 28*, 53-68.
- Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10* (4), 85.
- Eysenck, S.B.G., Pearson, P.R., Easting, G. & Allsop, J.F. (1985). Age norms for impulsiveness, venturesomeness, and empathy in adults. *Personality and Individual Differences, 6*, 613-619.
- Garfield, S.L. (1994). Research on client variables in psychotherapy. In Bergin, A.E. & Garfield, S.G. (Eds.), *Handbook of Psychotherapy and Behavior Change (4th ed.)*, pp. 190-228. New York: John Wiley and Sons.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology, 34*, 575-607.
- Gondolf, E.F. & Foster, R.A. (1991). Pre-program attrition in batterer programs. *Journal of Family Violence, 6*, 331-349.
- Hamberger, L.K. & Hasting, J.W. (1989). Counseling male spouse abusers: Characteristics of treatment completers and dropouts. *Hospital and Community Psychiatry, 39*,

763-770.

- Hanson, R.K., Gordon, A., Harris, A.J.R., Marques, J.K., Murphy, Q., Quinsey, V.L., et al. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 14*, 169-194.
- Novaco, R.W. (1994). Anger as a risk factor for violence among the mentally disordered. In J.
- Monahan and H.J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 21-59). Chicago: University of Chicago Press.
- Nunes, K.L. & Cortoni, F.A. (2006). *The heterogeneity of treatment non-completers*. Report R-176. Ottawa, ON: Correctional Service of Canada.
- Prochaska, J.O. & DiClemente, C.C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51*, 390-395.
- Scott, K.L. (2004). Stage of change as a predictor of attrition among men in a batterer treatment program. *Journal of Family Violence, 19*, 37-47.
- Serin, R.C., Gobeil, R., & Preston, D.L. (in press). Evaluation of the Persistently Violent Offender Treatment Program. *International Journal of Offender Therapy and Comparative Criminology*.
- Taylor, G. (2001). The importance of developing correctional plans for offenders. *Forum on Corrections Research, 13*(1), 14-17.

Evaluating Sex Offender Treatment Efficacy: A Review of Research Findings from a High Intensity Sex Offender Program

Mark E. Olver
University of Saskatchewan
E-mail: mark.olver@usask.ca

Stephen C. P. Wong
Regional Psychiatric Centre,
University of Saskatchewan

The adverse consequences of sexual abuse and the widespread and underreported nature of this social problem are well documented (Campbell & Wasco, 2005). Canada's 1993 national Violence Against Women Survey revealed that nearly 40% of women reported having experienced some form of sexual assault since the age of 16 (Johnson & Sacco, 1995) and over half of the reported sexual assaults in Canada in 2003 involved children or youth under age 18 (Canadian Centre for Justice Statistics, 2004). Offenders who go on to sexually re-offend, in addition to claiming further victims, incur additional fiscal costs associated with their re-incarceration. As such, correctional service providers face an important task to identify individuals at high risk for sexual violence and to target these individuals for treatment.

"What Works" with Sexual Offenders: A Review

An important practical question concerns the effectiveness of such programs for reducing sexual offense recidivism; that is, does treatment for sex offenders work? Perhaps the most common methodology for exploring this question is the single treatment outcome study, in which the recidivism rates of a group of treated sex offenders are compared to a comparable control group of sex offenders who have not received such services. A significantly lower base rate of sexual offense recidivism in the treated offenders relative to the controls would seem to provide support for the efficacy of treatment in reducing recidivism. In recent years, there has been an increase in meta-analytic reviews of single sex offender treatment outcome studies as a means of providing empirically informed aggregate findings about the effectiveness of sex offender programs in general (e.g., Hall, 1995; Gallagher, Wilson, Hirschfield, Coggeshall, & MacKenzie 1999; Hanson et al., 2002).

However, as Hanson, Broom, and Stephenson (2004) rightly point out, the value of meta-

analysis hinges on the quality of the individual studies that go into them. Early attempts to provide a broad quantitative evaluation of the efficacy of sex offender treatment were not particularly promising (e.g., Furby, Weinrott, & Blackshaw, 1989), in part, because the majority of the single treatment outcome studies included in the meta-analytic reviews were methodologically weak. As well, older studies were more likely to be based on less well conceptualized programs. Recent studies tend to use more vigorous methodologies, for instance through matching treatment and control groups on risk-related variables (Nicholaichuk et al., 2000), statistically controlling for risk, offense, history and age (Friendship, Mann, & Beech, 2003), imposing uniform and lengthier follow-up times (Hanson et al., 2004), and in one exceptional case, comparing treated offenders to matched randomly assigned control groups (Marques, Wienderanders, Day, Nelson, & van Ommeren, 2005). There has also been a substantial growth in knowledge about what treatment approaches "work" with offenders in general (e.g., Andrews et al., 1990) and sex offenders in particular (e.g., Marshall, Anderson, & Fernandez, 1999) for reducing recidivism. Other questions abound, for instance, how to classify sex offenders who drop out of treatment; some researchers analyze dropouts separately (e.g., McGrath, Cumming, Livingston, & Hoke, 2003), others include dropouts among the completers if they have completed a minimum portion of the treatment program (e.g., 50% or more) (e.g., Marques et al., 2005), and others have employed an intent-to-treat design in which all offenders to attempt treatment are compared to controls, who remain, strictly speaking, untreated (e.g., Nicholaichuk et al., 2000).

It seems as sex offender treatment programs have evolved and single treatment outcome studies have improved in quality, quantitative reviews of sex offender treatment have yielded more

encouraging findings. For instance, Hall (1995) obtained a small positive effect ($r = .12$) in his meta-analysis of 10 sex offender outcome studies, although this was offset by the fact that several of the studies included comparison groups of offenders made up of treatment refusers or dropouts. Gallagher et al.'s (1999) quantitative review of 25 sex offender outcome studies also generated evidence for the effectiveness of cognitive behavioral and relapse prevention based approaches for reducing sexual recidivism. Moreover, Hanson et al.'s (2002) meta-analysis of 43 sex offender treatment outcome studies found evidence for lower recidivism rates among treated offenders (12%) relative to comparison controls (17%) and stronger support among the "best" studies that examined "current treatments." Most recently, Hanson and Bourgon (2007) reported the findings from an updated meta-analysis of sex offender treatment outcome as a function of the quality of individual studies. Retaining 24 studies that met a set of minimal standards for methodological rigor (ranging from "weak" to "strong"), an odds ratio (OR) of .70 was obtained (i.e., interpreted to mean for every 100 untreated offenders who sexually re-offend, 70 treated offenders would re-offend).

In summary, while the extant literature, broadly analyzed, suggests that sex offender treatment has the potential to reduce sexual recidivism, the findings provide stronger support for the efficacy of certain treatment approaches (i.e., evidence based) and for programs adhering to the principles of risk, need, and responsivity.

Findings from a High Intensity Sex Offender Treatment Program

The Clearwater Sex Offender Program is a high intensity sex offender treatment program located at a multi-level security federal psychiatric facility (the Regional Psychiatric Centre; RPC) in Saskatoon, Saskatchewan. The Clearwater Program has been providing treatment services to federal sex offenders since the early 1980's. Over the years, the program content and structure has undergone changes as sex offender treatment and research has advanced. Typically, the program has been about 8 months in duration, cognitive behavioral in orientation, and has targeted moderate to high risk sex offenders. Actuarial-

conceptual risk measures (see Hanson & Morton-Bourgon, 2007), such as the Stable 2007 (Hanson, Harris, Scott & Helmus, 2007) and the Violence Risk Scale-Sexual Offender version (VRS-SO; see Olver, Wong, Nicholaichuk, & Gordon, 2007; Wong et al., 2003) are used to appraise risk and identify criminogenic needs to be targeted for treatment. Special efforts are made to be sensitive to responsivity issues, for instance, through providing additional culturally-based treatment services and programmatic adaptations for low functioning sex offenders.

Nicholaichuk et al. (2000) conducted an initial outcome evaluation of the Clearwater Program in a sample of 296 sex offenders who had attended treatment services and followed up approximately 6 years in the community. A large sample of 283 sex offenders from the Prairie Region were assembled as an untreated control group and efforts were made to match the groups closely on age and criminal history. The untreated offenders, in turn, were followed an average of 7.3 years post-release. Significantly fewer treated offenders were convicted for a new sex offense (14.5%) over the follow-up period than the matched untreated controls (33.2%). Treated repeat sex offenders also had a significantly lower sexual reconviction rate (23.5%) than repeat offenders in the untreated cohort (43%).

The Nicholaichuk et al. (2000) study provided further support for the efficacy of sex offender treatment, and it had some noted methodological strengths such as a fairly lengthy mean follow-up time and large samples of treated and untreated offenders matched on important risk-related variables. However, some methodological caveats are worth noting. First, the treated group had a shorter mean (and more variable) follow-up time or at-risk period than the controls, which might have created biases in favor of obtaining a positive treatment outcome. Second, in their commentary on the Nicholaichuk et al. (2000) study, Hanson and Nicholaichuk (2000) cautioned that the control sample contained a collection of untreated offenders who had quite old release dates and criminal records that pre-dated the treated sample. As the RCMP usually purge old criminal records unless offenders remain criminally active, the group of individuals with old release dates might contain a disproportionate

number of sexual recidivists. Although this could have the net result of exaggerating the treatment effect, a re-analysis of the Nicholaichuk et al. (2000) data, minus these individual cases, still demonstrated a strong treatment effect (OR = .47) (Hanson & Nicholaichuk, 2000).

With these caveats in mind, Olver, Wong, and Nicholaichuk (in press) conducted a subsequent follow-up treatment outcome evaluation of the Clearwater Program on a sample of 472 sex offenders who had attended treatment services. The treatment condition included 176 new treatment admissions that had subsequently been released and followed up a minimum of two years in the community. The same control sample was used as in Nicholaichuk et al. (2000) and the CPICs were updated and follow-up times extended for both offender groups from the earlier investigation; however, offenders from the control sample who had earlier release dates than the earliest released treated offender were removed from the study ($n = 17$).

Adding the new treatment participants also necessitated controlling for some new potential confounds since the treated and control groups were no longer matched as in the original study. In the more recent follow-up study, the comparison group had a substantially longer average follow-up time (14 vs. 8 years), were significantly younger at time of release (mean age 30.5 vs. 35.8 years), and had a smaller proportion of offenders who had at least one prior sexual conviction (38.2% vs. 48.5% respectively). To control for these differences, Cox regression survival analysis (to account for differences in follow-up time) was performed to examine sexual reconviction rates between the treatment and comparison groups across time while statistically controlling for age and sexual offending history. After imposing these controls the binary treatment variable remained a significant predictor of sexual recidivism (as did age and offending history), with treated offenders demonstrating lower rates of sexual reconviction over the follow-up period.

Conclusions

Treatment outcome evaluations of sex offender programs have demonstrated important methodological advances over time since the

Furby et al. (1989) review together with or perhaps in response to advances in knowledge and clinical practice about “what works” in sex offender treatment. The current paper provided a brief overview of some of these research developments and a more detailed review of the treatment efficacy of the Clearwater Sex Offender Treatment Program. In short, the results demonstrated a consistent significant treatment effect for the Clearwater Program in reducing sexual offense recidivism across two separate investigations employing different methodologies (matching vs. statistical controls) to control for potential confounds, large samples, and quite long follow-up time.

The Clearwater Program was developed and has been in existence at the Regional Psychiatric Centre since the early 1980s¹⁰. It was one of the first sex offender programs in the Correctional Service of Canada. The program has evolved and changed over the years: more medium and high risk sex offenders were admitted, program content was modified to reflect developments in the sex offender literature and best practice principles, program delivery became more structured, and the program was accredited twice by international expert panels after making significant changes. The leadership of the program also had changed many times together with many complete changeovers of line staff. The program also had its share of ups and downs and had to weather many challenges over the years; not at all unexpected within a large bureaucracy such as the CSC.

What had remained quite consistent from program inception to the present (other than the program name!) and had become the accepted treatment philosophy of Clearwater is the unwavering focus on risk reduction, intervention using cognitive-behavioral or behavioral treatment approaches, and the targeting of criminogenic factors in treatment; in essence, what the “What Works” principles advocate. Despite the coming and going of program directors and clinical leads with varying trainings, skills, professional backgrounds, personalities and management styles, the overall program philosophy remained

¹⁰ Dr. Arthur Gordon, Ph.D. was the psychologist instrumental in the development of the program.

fairly constant, withstood the challenges of time and had become the backbone of the Clearwater program. As reported above, two large-scale reviews of the Clearwater Program both with very substantial sample sizes and follow up periods have reaffirmed its treatment efficacy: quite comforting to know especially for the management of the program and the Centre. Can we attribute the “successes” of the program to Clearwater’s constant program philosophy? No, we cannot; the available empirical evidence does not allow for such extrapolation. However, what we can say is that the positive outcomes of the program are entirely consistent with our understanding of the “What Works” principles on which the treatment philosophy of the Clearwater program is more or less aligned. In our not so final analyses, it appears that “What Works” can work in an institutional based program for quite high risk repeat sex offenders over the long term in spite of the many trials and tribulations that institutional treatment programs such as Clearwater must endure. As Clearwater is again making changes to accommodate more developmentally delayed and cognitively impaired sex offenders with mental health concerns, the staff can make programmatic adjustments with the understanding and confidence that they are making changes on top of a very solid foundation. In no small part, the tribute for the success of Clearwater must go to the participants and staff of the program over the years; for the participants, in their beliefs that they can make changes under the right circumstances, and, for the staff, to their dedication in providing the right circumstances.

References

- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically-relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- Campbell, R., & Wasco, S. M. (2005). Understanding rape and sexual assault: 20 years of progress and future directions. *Journal of Interpersonal Violence*, 20, 127-131.
- Canadian Centre for Justice Statistics (2004). Canadian crime statistics 2003.
- Friendship, C., Mann, R. E., & Beech, A. R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence*, 18, 744-759.
- Furby, L., Weinrott, M. R., & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin*, 105, 3-30.
- Gallagher, C.A., Wilson, D.B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual reoffending. *Corrections Management Quarterly*, 3, 19-29.
- Hall, G. C. (1995). Sex offender recidivism revised: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology*, 63, 802-809.
- Hanson, R. K., & Bourgon, G. (2007). *Meta-analysis of sexual offender treatment outcome studies: Distinguishing quality studies from quality treatment*. Presented at the North American Correctional and Criminal Justice Psychology Conference, Ottawa, Ontario.
- Hanson, R. K., Broom, I., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Science*, 36, 87-96.
- Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J., Murphy, W. D., Quinsey, V. L., & Seto, M. (2002). First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14, 169-194.
- Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). *Assessing the risk of sex offenders on community supervision*. (User Report No. 2007-05). Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R. K., & Morton-Bourgon, K. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis*. (User Report No. 2007-01). Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R. K., & Nicholaichuk, T. P. (2000). A cautionary note regarding Nicholaichuk et al. *Sexual Abuse: A Journal of Research and Treatment*, 12, 289-293.
- Johnson, H., & Sacco, V. F. (1995). Researching violence against women: Statistics Canada’s national survey. *Canadian Journal of Criminology*, 37, 281-304.
- Marques, J. K., Wienderanders, M., Day, D. M., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: California’s Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, 17, 79-107.
- Marshall, W. L., Anderson, D., & Fernandez, Y. (1999). *Cognitive behavioural treatment of sexual offenders*. West Sussex, England: Wiley.
- McGrath, R. J., Cumming, G., Livingston, J. A., & Hoke, S. E. (2003). Outcome of a treatment program for adult sex offenders. *Journal of Interpersonal Violence*, 18, 3-17.
- Nicholaichuk, T. P., Gordon, A., Gu, D., & Wong, S. (2000). Outcome of an institutional sexual offender treatment

program: A comparison between treated and matched untreated offenders. *Sexual Abuse: A Journal of Research and Treatment*, 12, 139-153.

Olver, M. E., Wong, S. C. P., & Nicholaichuk, T. (in press). Outcome evaluation of a high intensity inpatient sex offender treatment program. *Journal of Interpersonal Violence*.

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

Wong, S., Olver, M. E., Nicholaichuk, T. P., & Gordon, A. (2003). *The Violence Risk Scale: Sexual Offender version (VRS-SO)*. Regional Psychiatric Centre and University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Care, Control, and Mental Disorder: Comparing Practices and Outcomes in Prototypic Specialty vs. Traditional Probation

Sarah M. Manchak, Jennifer L. Skeem, & Sarah Vidal
E-mail: smanchak@uci.edu

Individuals with mental disorder are disproportionately represented in the criminal justice system (Lurigio, 2001; James & Glaze, 2006). A majority of these offenders are supervised in the community on probation (Glaze & Palla, 2005). Traditional models of supervision may be insufficiently responsive to the pronounced needs and unusual risk factors of probationers with mental disorder (PMDs), as evidenced by their increased rates of probation revocation compared to their non-mentally disordered counterparts (Dauphinot, 1999). In response to this problem, many agencies across the U.S. have implemented specialty caseloads that are reduced in size and supervised by officers with interests and training in mental health. Relative to traditional programs, specialty programs aim to integrate internal and external services, balance care and control in probation officer (PO)-probationer relationships, and emphasize treatment adherence (Skeem, Emke-Francis, & Eno Loudon, 2006).

Despite efforts to initiate and maintain these specialty programs, little is known about whether (a) the differences believed to distinguish specialty from traditional supervision are implemented in practice, and (b) specialty supervision improves PMDs' clinical and criminal outcomes, relative to traditional supervision. The current study seeks to describe how traditional and specialty caseloads differ in their supervision strategies and resource coordination, and how this affects PMDs' outcomes.

Method

Participants

182 PMDs on specialty probation were matched with 176 PMDs on traditional probation.

Results indicate a successful match on age, gender, ethnicity, index offense (person/violent

versus property/non-violent and felony versus misdemeanor), and time spent on probation up to the baseline interview. Our total sample ($n=358$) is approximately 60% male, predominantly Caucasian (38%) and African American (50%), and about 12% of Hispanic ethnicity. Over half are aged 35-65 (57%), and one third (34%) had a violent/person index offense (e.g., assault, robbery). At the time of the baseline interview, approximately 64% had been on probation less than one year, 21% for one to two years, and 14% for over two years.

Procedures

Once consented, all participants complete a baseline, six month, and twelve month interview. The probationers' supervising PO also completes a survey on the probationer at each of the time points. The present analysis examines the differences between specialty and traditional probation supervision on (a) baseline compliance strategies (techniques used by POs to obtain compliance from their probationers), (b) PO boundary spanning (how well the PO steps outside his/her traditional supervisory role to coordinate services and care for probationers), and (c) PO/Probationer relationship quality. Preliminary outcome data include: probationers' perceptions of the effects of probation (as a proximal outcome), probationers' access to services, and probation compliance outcomes between baseline and six months. These results represent all of our baseline data ($n=358$) and about three quarters of our six month data ($n=264$). Data collection for the remaining six month interviews is ongoing. We will also eventually have twelve and eighteen month outcome data to report.

Data Analysis: Propensity Scores

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 73-75. Public Safety Canada.

This paper is also published in *Crime Scene: The Official Organ of Criminal Justice Psychology*, 14(2), 29-31.

Some analyses presented below control for propensity scores, or the likelihood of being on traditional probation given certain criminal (e.g., most serious crime ever and total number of past arrests), clinical (Global Assessment of Functioning score), and personality features (anxiety scale of the Personality Assessment Inventory). Because the present study represents a matched design rather than random assignment to specialty or traditional probation conditions, this statistical technique helps to balance potentially non-equivalent groups to obtain more valid estimates of the effects of traditional vs. specialty supervision on PMDs (Luellen, Shadish, & Clark, 2005).

Results

Preliminary analyses examining differences between sites on supervision practices demonstrate that specialty POs see their probationers slightly more often per month (2.2 times/month vs. 1.5 times/month) and for a longer amount of time (26 minutes vs. 18 minutes) than traditional POs ($t = 6.0, p < .001, t = 5.4, p < .001$, respectively). Factor analyses of supervision strategies yielded three primary factors: problem solving, threats, and sanctions. The sites differ significantly from one another in terms of supervision techniques. Specifically, specialty POs use significantly more problem solving strategies ($t = 6.5, p < .001$) and significantly less threats ($t = -3.6, p < .001$) and less sanctions ($t = -2.1, p < .05$) than their counterparts in traditional probation agencies. These differences remain significant after controlling for propensity scores.

Next, we examine differences in PO boundary spanning. Factor analyses in this realm yielded four primary factors: PO knowledge (e.g., of other systems), PO helpfulness, PO advocacy, and PO direct/“hands on” involvement (e.g., attends treatment with the probationer on his first visit). Specialty POs are perceived by their probationers as more knowledgeable ($t = 4.8, p < .001$), more helpful ($t = 6.5, p < .001$), and more like an advocate ($t = 4.6, p < .001$). Specialty POs are less likely to be perceived as using a more direct “hands on” approach, but this difference is not significant. Only the knowledge factor remains significant after controlling for propensity scores.

The final set of baseline site comparisons involves the quality of the PO-probationer relationship (PPR). Factor analyses on items assessing the PPR elicited a three factor solution: fairness/caring, trust, and toughness. Probationers on specialty caseloads perceive their relationship with their PO to be more fair and caring ($t = 7.9, p < .001$) and more trusting ($t = 5.5, p < .001$), but less tough ($t = -6.0, p < .001$) than traditional probationers’ perceptions of the PPR. The fairness/caring and trust factors remain significant after controlling for propensity scores.

Our next sets of analyses examine probationers’ perceptions of probation’s effect on their lives. Probationers on specialty supervision are significantly more likely to feel that probation has helped them keep treatment appointments and take prescribed medications ($t = 5.0, p < .001$), get and stay well ($t = 5.1, p < .001$), gain more control over their lives ($t = 4.7, p < .001$), and stay out of trouble with the law ($t = 5.4, p < .001$).

Using available six month outcome data, we next examine the amount and quality of services accessed between the baseline and six month interviews. Probationers on specialty supervision are more likely to get mental health treatment ($\chi^2 = 41.2, p < .001$) and medication ($\chi^2 = 14.0, p < .01$). The two sites do not differ significantly in terms of substance abuse or physical health treatment. Only the differences between sites on mental health treatment remain significant after controlling for propensity scores. Factor analysis on items ascertaining the nature of help provided by mental health treatment providers produced four factors: mental health, legal/drug problems, daily life, and housing. There are significant differences between the two sites on help with daily life ($t = -2.4, p < .05$).

Our final set of analyses examines preliminary probation compliance outcomes. Results indicate no significant differences between the sites on probationers’ substance abuse per both self report and official records. There is a significant difference between the two sites on the likelihood for technical violations ($\chi^2 = 33.3, p < .001$). Specifically, probationers on specialty probation are significantly more likely to be violated for treatment non-compliance ($\chi^2 = 54.3, p < .001$), and not reporting to probation ($\chi^2 = 25.9, p < .001$).

These differences remain significant after controlling for propensity scores.

Discussion

This study is the first to suggest that specialty and traditional supervision officers adhere, in practice, to the supervision philosophies that define them. Moreover, it appears that mental health treatment plays an integral role in specialty probation; PMDs on specialty probation are more likely to receive mental health treatment and be violated for treatment non-compliance than PMDs on traditional probation. These results also suggest that there may be more differences between the two supervision approaches than just strategy. There appears to be better PO-probationer rapport and more probationer satisfaction with specialty probation than with traditional probation, but the current state of our data did not allow for analyses to tease apart the predictive effects of a number of variables on future criminal justice and mental health outcomes. This is an ongoing study, and future analyses on a complete data set will help solidify our understanding of the individual and systemic causal mechanisms contributing to probationers' success and non-success on probation.

References

- Dauphinot, L. (1996). *The efficacy of community correctional supervision for offenders with severe mental illness*. Unpublished doctoral dissertation, University of Texas at Austin.
- Glaze, L. E. & Palla, S. (2005, December 6). *Adults on Probation in the United States* (taken from the Annual Probation Data Survey data series, CJ-8). Washington, DC: Bureau of Justice Statistics.
- James, D., & Glaze, L. (2006). *Mental health problems of prison and jail inmates*. Washington, DC: Bureau of Justice Statistics.
- Luellen, J., Shadish, W., & Clark, M. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, 29(6), 530-558.
- Lurigio, A. J. (2001). "Effective services for Parolees with Mental Illness." *Crime and Delinquency* 47: 446-461.
- Skeem, J., Emke-Francis, P., & Eno Loudon, J. (2006). Probation, mental health, and mandated treatment: A national survey. *Criminal Justice and Behavior*, 33(2), 158-184.

Section D

Police & Court Psychology

Clinical Versus Actuarial Geographic Profiling Approaches: A Meta-Analysis

Karla B. Emeno
Carleton University
E-mail: kemeno@connect.carleton.ca

Paul J. Taylor
University of Liverpool

Craig Bennell
Carleton University

Brent Snook
Memorial University of Newfoundland

Geographic profiling (GP) predictions (about where an unknown serial offender lives) can be produced using a variety of strategies. Some predictions are made using an equation or mechanical aid (actuarial approach) while others are made by human judges drawing on experience or heuristic principles (clinical approach). The actuarial approach in this context typically takes the form of computerized GP systems. These systems apply mathematical models of offender spatial behaviour to a series of linked crimes in order to produce a probability surface that indicates the likelihood of an offender residing at particular points in this area of criminal activity. The mathematical model that is applied most often takes the form of a distance decay function, reflecting the fact that the likelihood of an offender residing at a particular location decreases as the distance between that location and a crime site increases. The clinical approach typically involves the use of heuristics that essentially exploit well-known facts about offender spatial behaviour (such as the distance decay phenomenon just described) to make informed GP predictions. There is currently a debate about which of these approaches will result in the most accurate profiles (e.g., Rossmo, 2005; Snook, Taylor, & Bennell, 2004).

Support for the Clinical Approach to GP

On the one hand, Gigerenzer and his colleagues (e.g., Gigerenzer, Todd, & The ABC Group, 1999) have amassed a considerable amount of research demonstrating that heuristics, or simple cognitive strategies that allow people to make quick decisions, can perform as accurately as computationally expensive techniques (e.g., multiple regression). More specifically, this research demonstrates that heuristics work well when they match the structure of the decision environment (Gigerenzer & Selton, 2001;

Martignon & Hoffrage, 1999; Simon, 1956). Thus, if people utilize a heuristic (e.g., search for crime locations in a series that are furthest apart and predict that the offender lives at a point midway between those two locations) that matches the patterns found in serial offender spatial behaviour (e.g., most serial offenders live within their criminal activity space), they are likely to perform well on the GP task.

Support for this view comes from a series of studies that have directly compared the clinical and actuarial approach for making GP predictions (e.g., Snook, Canter, & Bennell, 2002; Snook, Taylor, & Bennell, 2004). For example, Snook, Taylor, and Bennell (2004) considered explicitly the role of heuristics in clinical solutions to GP by examining individuals' explanations of how they reached their predictions before and after training on how to make GP predictions. They found that a substantial portion of people (approximately 50%) utilized appropriate heuristics even before training while others only did so when given instructions about how to make accurate GP predictions. In this case, "appropriate" heuristics related to one of two rules of thumb: (1) that offenders live roughly in the middle of their crime site locations (the circle heuristic) or (2) that offenders live in close proximity to the majority of their crime sites (the decay heuristic). Moreover, participants who used ecologically rational heuristics made predictions that were as accurate as the actuarial GP tool in *CrimeStat*, a commonly used profiling system.

Support for the Actuarial Approach to GP

On the other side of the debate, Rossmo (2005) and Rossmo and Filer (2005) have highlighted several criticisms against the aforementioned GP research. First, they argue that the data selected for analysis in previous studies does not meet the GP assumptions outlined by Rossmo (2000),

which include the fact that GP works best when the crimes are committed by a single offender, no movement of residence has taken place during the crime series, the offender is not a commuter, and the victim backcloth is uniformly distributed around the offender's anchor point. Second, they highlight the fact that the number of crimes used to make GP predictions was lower than what Rossmo (2000) recommends (i.e., some of the maps used in previous studies included three crimes, whereas Rossmo recommends five or more crimes). Third, they point out that GP performance in examinations of the clinical approach have been measured using a metric referred to as error distance (i.e., the straight line distance between predicted and actual home locations) rather than an "area-based" measure such as *hit percentage* (i.e., the proportion of the prioritized search area that needs to be searched before finding the offender's anchor point).

Fourth, Rossmo and his colleague claim that, due to the fact that Snook, Taylor, and Bennell have no experience as police officers, investigators, or geographic profilers, the credibility of the advice emerging from their studies is questionable. Fifth, they suggest that the focus on students as participants in previous studies does not allow the results to be generalized to "real-world" investigative settings, where law enforcement officers typically make the GP predictions. Sixth, they raise the issue that the predictions considered in previous experiments fail to incorporate information that is routinely used by professional geographic profilers (e.g., arterial routes and barriers), thereby making any result invalid. Finally, they state that studies, which have shown support for the use of the clinical GP approach, have not compared participants' performance to commonly used GP systems.

A number of these criticisms have recently been addressed. For example, Paulsen (2006) found that regardless of the length of the crime series under consideration and regardless of the way profile accuracy was measured, predictions made by human judges and spatial distribution methods were as accurate as predictions made by more complex actuarial-based methods, some of which are commonly used by the police. In addition, Bennell, Snook, Taylor, Corey, and Keyton (2007) found that training significantly improved

the predictive accuracy of police officers faced with the GP task, regardless of the number of crime locations used to construct the geographic profile or the amount of topographical detail presented (two variables that were intended to enhance the external validity of the GP task in these studies). The trained participants substantially outperformed *CrimeStat*.

However, despite these studies, debate continues around the issue of whether actuarial or clinical approaches to GP are more effective.

The Current Study

In an attempt to resolve this debate, a meta-analysis of study effects was conducted to assess the overall body of evidence produced by the combination of studies that examined predictive accuracy of clinical versus actuarial GP approaches. The goal in conducting the meta-analysis was to determine, in statistical terms, the extent to which the clinical approach may be thought of as providing predictions that are as accurate as actuarial methods.

Method

Procedure

Eligible studies for the meta-analysis were those that: (1) used an experimental scenario (i.e., contained an independent variable) in which both the crime locations and home location were known to the experimenter, (2) compared the predictive accuracy of the trained participant group to a comparison group of untrained participants, an actuarial GP system, or both, and (3) reported information regarding the relationship between the predicting group and the accuracy of prediction in a (statistical) form that could be converted into a common effect size (r). A search of the GP literature revealed seven potential studies, of which six met the aforementioned criteria (these studies are indicated by an asterisk in the reference section).

Analysis

Pearson correlation coefficients (r) were calculated across the experimental groups. When statistics other than r were reported in the study (i.e., F , t , p), the appropriate formulae were used to convert them to r values. Effect size (ES)

magnitudes were assessed by examining the mean r values and their respective 95% confidence intervals (CI s) for each outcome. Results were also assessed using Rosenthal and Rubin's (1982) Binomial Effect Size Display (BESD). This statistic allows one to examine changes in success rates that are attributable to the predictor variable, assuming a base rate of 50%.

Results

In relation to the effect of heuristic training on predictions, the results indicate that training produces more accurate predictions ($r = .27$). In terms of the BESD, the accuracy rate of trained participants was 63.5%, which compares with a 36.5% accuracy rate for untrained participants. For the human participant-actuarial system comparison, the results indicate that trained participants performed about as well as actuarial systems ($r = -.01$)¹¹. In terms of the BESD, the accuracy rate of human predictions was 50.5% compared to 49.5% for the actuarial methods.

These results suggest that training has a moderate effect on peoples' ability to predict home locations. Moreover, training appears to allow people to make better predictions that are on par with those made by actuarial systems. However, note that the 95% CI s for the two point estimates are greater than .10 units in width (see Table 1), which suggests that both results should be regarded as tentative findings.

Discussion

Based on the results presented in this paper, seven tentative conclusions can be drawn:

Table 1. Effect Sizes for Experimental versus Comparison Groups

Outcome (k)	N	Mean r (SD)	95% CI_r
Pre-Training vs. Post Training (14)	826	.27 (.15)	.18 to .36
Human vs. Actuarial (29)	1005	-.01 (.39)	-.16 to .14

Note. k = number of effect sizes; N = number of participants per predictor; mean r (SD) = mean Pearson correlation coefficient for each predictor with standard deviation; 95% CI_r = confidence intervals about r .

- (1) Many people have access to GP strategies that allow them to make relatively accurate GP predictions.
- (2) Providing people with advice regarding appropriate GP strategies improves their performance.
- (3) Heuristic training does not allow everyone to perform equally well, as trained people in every study exhibited a high degree of variation in their performance.
- (4) Heuristic training often allows people to perform as well as actuarial strategies.
- (5) The above appears to hold true even when predictions are based on different numbers of crime and under other conditions (e.g., such as when topographic information is available for processing).
- (6) Under conditions where people perform poorly as a group, GP systems also appear to perform poorly.
- (7) Under some circumstances, clinical judgments can outperform actuarial methods.

However, even if the above conclusions are accepted as valid, there remain several advantages to using actuarial GP systems. These advantages include the fact that predictions made by actuarial systems will always be more reliable than predictions made by human judges and actuarial GP methods may allow the user to avoid cognitive overload when he/she is inundated with investigative information.

¹¹ Note that, of these 29 ESs, 21 relate to error distance ($r = -.12$), four relate to an area-based measure of accuracy calculated from uniform search strategies ($r = .50$), and four relate to an area-based measure of accuracy calculated from directed (i.e., irregularly shaped) search strategies ($r = 0.06$) (these last two measures deal directly with Rossmo's criticism of our use of error distance to measure accuracy and, in the case of the directed search, suggest that the clinical approach can produce area-based levels of accuracy that are similar to those produced by actuarial systems).

There are several areas of future research that are important to examine. First, replication and expansion of clinical versus actuarial research is required (e.g., looking at different crime types). Second, researchers need to thoroughly explore the assumptions that are supposedly required before GP predictions can be made in an accurate fashion. Third, the qualitative component of GP needs to be examined in order to determine if anything useful is added by this contribution (the qualitative component of GP consists of a profiler manipulating a quantitative GP prediction based on their understanding of case specifics). Finally, GP research needs to be taken into the operational arena where both the clinical and actuarial approach to GP can be evaluated in a more realistic fashion.

Conclusion

People often have access to heuristics that allow them to make accurate GP predictions. For those people who do not have access to these heuristics, it may be feasible to rely on a fast and frugal training exercise that teaches these individuals simple decision rules. However, given their advantages, if future research demonstrates that sophisticated actuarial GP systems are effective in criminal investigations there seems to be little reason not to use them, so long as time, knowledge, and resources are plentiful. The choice then, between whether a clinical versus an actuarial approach to GP is most appropriate in a specific criminal investigation, will likely be context dependent. Fortunately, at the moment, both strategies appear to hold some merit.

References¹²

- * Bennell, C., Snook, B., Taylor, P. J., Covey, S., & Keyton, J. (2007). It's no riddle, choose the middle: The effect of number of crimes and topographical detail on police officer predictions of serial burglars' home locations. *Criminal Justice and Behavior, 34*, 119-132.
- Gigerenzer, G., & Selton, R. (2001). *Bounded rationality*. London, UK: MIT Press.
- Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. New York, NY: Oxford University Press.

- Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work? A case study in ecological rationality. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 119-140). New York, NY: Oxford University Press.
- * Paulsen, D. (2006). Human versus machine: a comparison of the accuracy of geographic profiling methods. *Journal of Investigative Psychology and Offender Profiling, 3*, 77-89.
- Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*, 166-169.
- Rossmo, D. K. (2000). *Geographic profiling*. Boca Raton, FL: CRC Press.
- Rossmo, D. K. (2005). Geographic heuristics or shortcuts to failure?: A response to Snook et al., (2004). *Applied Cognitive Psychology, 19*, 651-654.
- Rossmo, D. K., & Filer, S. (2005, August/September). Analysis versus guesswork: The case for professional geographic profiling. *Blue Line, 24* – 26.
- Simon, H. A. (1956). Rational choice and the structure of environments. *Psychological Review, 63*, 129-138.
- * Snook, B., Canter, D. V., & Bennell, C. (2002). Predicting the home location of serial offenders: A preliminary comparison of the accuracy of human judges with a geographic profiling system. *Behavioral Sciences and the Law, 20*, 109-118.
- * Snook, B., Ennis, A. R., Bennell, C., Johnson, S. D., Taylor, P. J., & Campbell, L. (2006). *The effect of heuristic-training and topographical information on geographic profiling performance*. Manuscript submitted for publication.
- * Snook, B., Taylor, P. J., & Bennell, C. (2004). Geographic profiling: The fast, frugal and accurate way. *Applied Cognitive Psychology, 18*, 105-121.
- * Taylor, P. J., Bennell, C., & Snook, B. (2007). The bounds of cognitive heuristic performance on the geographic profiling task. *Manuscript submitted for publication*.
- ** Van der Kemp, J., van Ruth, E., Blokland, A., & Snook, B. (2005). *X marks the spot: Comparing police officers, students, and geographic profiling software on the accuracy of their predictions*. Poster presented at the 15th European Conference on Psychology and Law, Vilnius, Lithuania.

¹² References marked with an asterisk were included in the current meta-analysis. Reference marked with double asterisks was excluded from the current meta-analysis.

Exploring the *CSI Effect*: Is it Real? If so, What is it?

Steven M. Smith, Marc W. Patry, & Veronica Stinson
 Saint Mary's University, Halifax, NS
 E-mail: steven.smith@smu.ca

There is no question that *CSI* and its related spin-offs are among the most popular shows in television today. Indeed, *CSI* is consistently among the top five shows (often #1 or #2) in North America (www.nielsenmedia.com). Recently, newspapers and television news programs (e.g., *USA Today*, *the Toronto Star*, *CNN*) have been reporting on how television crime dramas such as *CSI*, *Law & Order*, *Bones*, and other related shows may influence how people think about and behave relative to the legal system. The news media has dubbed this the *CSI effect*.

News media coverage often describes the lengths to which prosecutors and other legal professionals are trying to overcome the impact of these programs (e.g., see Patry, Smith & Stinson, in press). However, the social psychological literature is clear that working to overcome a "bias" is only effective if we can know the nature and extent of that bias (e.g., Wegener & Petty, 1997). Indeed, without knowing the nature of the bias one can easily over or under correct, particularly in the legal system (see Wegener, Kerr, Fleming & Petty, 2000). Thus, it has become clear that empirical evidence on the nature and consequences of the so-called *CSI effect* is warranted. In this paper we provide an overview of conceptualizations of the *CSI effect* and review recent work that has explored the effect, including some of our studies.

How is the CSI effect described?

Each episode of *CSI* tells the story of a sensational criminal case wherein forensic investigators use state-of-the-art techniques, sparing no expense (in either cost or personnel) to solve the crime. The techniques presented on these shows are usually plausible, as the show's producers are careful to use technically possible investigative tools, but surveys of police professionals suggest these techniques are far from common (see Stinson, Patry & Smith, in press). Thus, given the relative

"inaccuracy" of these television programs, it is quite possible that viewers (and potential jurors) may be learning incorrect information from televised crime dramas. Importantly, a recurring theme in these programs is that the quality of scientific evidence is so good that the criminal is almost always caught and usually confesses in light of the overwhelming evidence against him (or more rarely, her).

Typically, media reports of the *CSI effect* include references to an undesirable effect exhibited during jury trials which results from juror's reaction to the presence or absence of "appropriate" scientific evidence as trial exhibits (see Tyler, 2006; Podlas, 2006). Essentially, the argument is that watching *CSI* causes jurors to have unrealistic expectations about the quantity, quality, and availability of scientific evidence (see Smith et al., in press; Stinson et al, in press). When the scientific evidence presented at trial fails to meet jurors' television-enhanced expectations, they are more likely to acquit the defendant. How accurate is this conceptualization of the *CSI effect*?

Recent Research on the CSI effect

Although many news reports have documented the perceptions of legal professionals with regard to the *CSI Effect*, there is little empirical research about this topic. One notable exception is an analysis of the perceptions and behaviors of members of the Maricopa County Prosecuting Attorney's Office (2005) which conducted a survey of 102 prosecutors to assess the perceived impact of the *CSI Effect* and how these lawyers are responding to this issue. Clearly, these lawyers believe *CSI* is having an effect, as 38% of attorneys reported they had lost a case because of the *CSI effect*; 45% contended that jurors relied on scientific evidence more than they should; and 72% maintained that *CSI* fans exerted undue influence on other jurors. In terms of solutions to the problem, 70% of prosecutors asked jurors

about television viewing habits, 90% took the time to explain police procedures to jurors, and an amazing 52% plea bargained cases when they thought *CSI*-educated jurors might object to the evidence presented in the case.

Despite legal professionals' beliefs, little evidence exists regarding the impact of *CSI* and related shows on trial outcomes. The only published study on the matter that we are aware of (Podlas, 2006) has found no such relationship. In Podlas' study participants read a scenario of an alleged rape which was based entirely on the credibility of witnesses (no forensic evidence was presented), then rendered a verdict and reported on the basis for their decision. Although the lack of forensic evidence was the reason most often provided for mock jurors' decisions, frequent viewers of *CSI* were not more likely than infrequent viewers of *CSI* to cite the lack of forensic evidence for their not-guilty verdicts.

Considering this relative dearth of evidence, it is not surprising that the argument has been made that the purported *CSI* effect does not exist. Indeed, Tyler (2006) convincingly argues that despite claims that the *CSI* effect constitutes a pro-defense bias, the effect could actually produce a pro-prosecution bias. To support his argument, Tyler notes that in almost all the storylines on *CSI*, the criminal is caught and convicted. Tyler speculated that this consistent conclusive ending may give people unrealistic expectations for real life cases. Tyler argues that complaints about the *CSI* effect may in fact result from disgruntled prosecutors trying to find alternative explanations for their failures.

Our Recent Research on the CSI Effect

In conducting our research, we first wanted to understand how the *CSI* effect was described in the media. Thus our first study focused on an analysis of 250 newspaper articles from media sources. Our analysis suggested that the *CSI* effect tended to be described in the popular media as having one of four impacts: 1) television crime dramas have peaked student interest in Biology, Anthropology, Psychology and other forensically relevant topics leading to increased enrollments in these courses and relevant programs; 2) these shows have actually educated criminals in how to

engage in criminal activity without getting caught; 3) jurors are influenced by *CSI* to expect too much evidence and thus acquit defendants; and 4) influencing how lawyers and other legal professionals behave. Because news reports frequently characterized the *CSI* effect as being negative and assumed that juror expectations were being affected by the unrealistic portrayals, our next study explored how *CSI* portrays crime scene investigation.

Our next study was a content analysis of the first seasons of *CSI* and *CSI: Miami* (see Patry et al., in press; Smith et al., in press). We sought to document the types of forensic procedures portrayed, the frequency of errors, and the frequency with which criminals were caught. We identified over 75 types of forensic evidence. Of all techniques used, 18.9% were DNA based, and 12% used fingerprinting, making them the two most common types of evidence portrayed. As expected, the criminal was caught in 97% of the storylines. Errors were rare and were always caught before any negative consequences could arise.

In three subsequent studies (Stinson, et al., in press) we surveyed legal professionals to assess their views on the *CSI* effect. First we conducted a web-based survey of nine defense lawyers, who typically did not view the *CSI* effect as a problem, but 2/3 said their crime dramas had distorted their client's views of the legal system. Next we surveyed of 127 Death Investigators (i.e., 83 Police, 28 Medical Examiners, 7 Fire/Arson Investigators, and 6 others) who confirmed that crime dramas are less-than-accurate, and have changed the way in which police practice, investigate, and interact with the public. Furthermore, 94% indicated that crime dramas influenced the public's expectations of their profession. We followed this up with a survey of 36 police officers completed a survey similar to that used in the previous study but with additional questions on the extent to which *CSI* and similar shows influenced juries, criminal behaviors, and perceptions of the legal system. Respondents were also asked to estimate the percentage of crimes which are solved on these shows, relative to real life. Although most of the police officers (68%) indicated that *CSI* had no effect on their behavior, almost all (92%) indicated that the shows had

some effect on public expectations. Most respondents indicated that the shows did not represent the court system accurately and that the shows had some effect on jury decision making. Importantly, all respondents felt that *CSI* affected people's *perceived* knowledge of forensic techniques, but most thought that knowledge was inaccurate. Although respondents estimated that the shows depicted a 94% solution rate, they estimated that only 40% of crimes are solved in real life (see Stinson et al, in press). Nonetheless, it is possible that legal professionals are over-reacting to media hype about the CSI effect, thus we have also explored the extent to which crime dramas influence public perceptions evidence.

In our next study, (see Patry, et al., in press) we surveyed 320 jury-eligible adults concerning several types of evidence. Participants clearly demonstrated a preference for scientific over more traditional forms of evidence, with DNA and fingerprint evidence consistently rated as the most reliable and useful evidence. Subsequent to this (Smith et al., in press) we explored how television viewing habits related to beliefs about forensic evidence. Data from 148 participants showed that self-reported viewing of *CSI* and *Law and Order* (0 to 15 hours per week) predicted favorable views toward a number of types of scientific but not non-scientific evidence. However, as correlation does not necessarily indicate causation, we conducted another study to test for a causal relationship between exposure to *CSI* and attitudes toward forensic evidence.

For the causal study (see Smith et al., 2006), we randomly assigned 190 undergraduates to watch zero, four, or eight episodes of *CSI*. Compared to those who did not view *CSI*, participants who watched 4-8 episodes of *CSI* had higher estimates of the reliability of DNA evidence, both accuracy and reliability of DNA and fingerprint analysis, and had more confidence in their judgments about the reliability of DNA analysis. It is quite possible that this effect occurs because DNA and fingerprint analysis are the techniques most commonly portrayed on the show.

In the final study we will present here, we explored the extent to which people's attitudes toward forensic evidence are malleable. Thus, we showed sixty-three participants a video entitled

"Reasonable Doubt", produced by CNN, which has four segments providing a critical examination of the quality of DNA, compositional, fingerprint and fiber evidence. After watching the video (initial attitudes had been recorded earlier in the term in an ostensibly unrelated task) participants rated forensic techniques as less reliable, but it did not influence ratings of more circumstantial evidence (motive, opportunity, confessions, and alibi evidence).

Summary and Conclusions

Based on the studies we have conducted, we feel we can confidently say that the CSI effect does indeed exist, but not necessarily in the form typically found in media reports. In particular, we feel that there is evidence that crime dramas have had an influence on public perceptions of forensic evidence and of professionals working in areas related to criminal law. Further, the evidence we have presented here clearly suggests that legal professionals are changing their behavior in response to these perceived effects. Nonetheless, more research is needed, particularly in order to more clearly delineate the nature and consequences of the CSI effect.

One important question that remains is the extent to which the CSI effect influences trial proceedings. Nonetheless, the fact that legal professionals seem to be changing their behavior in response to the perceived impact of crime dramas on the public is quite interesting. Therefore it is imperative that social science address the issue of juror decision-making in order to properly inform the legal community of the nature of the CSI effect, as understanding the nature and magnitude of any bias is necessary before any intervention is appropriate (see Wegener et al., 2000). Our research suggests that people who watch *CSI* judge forensic evidence to be *more* reliable and accurate than the non-*CSI* watching public. Thus, consistent with the views of Tyler (2006) a pro-prosecution bias may be more likely, assuming the expected evidence is provided at trial.

Ultimately, it is clear that many questions remain unanswered with regard to the CSI effect. More research is needed in order to fully understand the extent to which television crime dramas are

influencing public expectations and interactions with the legal system, legal professionals, their investigative procedures, and finally, trial procedures and outcomes. The current situation where actions are being taken without consideration for the true nature of the CSI effect could very well lead to more problematic outcomes than CSI was believed to have caused in the first place.

References

- Maricopa County (2005). *The CSI Effect and its Real-Life Impact on Justice: A Study by the Maricopa County Attorney's Office* (June 30, 2005). Retrieved on September 20, 2005 from http://www.maricopacountyattorney.org/Press/PDF/CSI_Report.pdf
- Patry, M. W., Stinson, V., & Smith, S. M. (in press). CSI Effect: Is popular television transforming Canadian society? In Greenberg & Elliott (Eds.). *Communications in Question: Canadian Perspectives on Controversial Issues in Communication Studies*.
- Patry, M., Stinson, V. & Smith, S. M. Blurring the line between fact and fiction: Expert opinions about forensic investigation tools represented on CSI. Unpublished Manuscript.
- Podlas, K. (2006). "The CSI Effect": Exposing the media myth. *Fordham Intellectual Property, Media and Entertainment Law Journal*, . 429-465.
- Smith, S. M., Stinson, V., & Patry, M. (in press). But what is the CSI Effect? How crime dramas influence people's beliefs about forensic evidence. *Canadian Journal of Police and Security Services*, 5,
- Stinson, V., Patry, W., & Smith, S.M. (in press). The CSI Effect: Reflections from Police and Forensic Investigators" *Canadian Journal of Police and Security Services*, 5,
- Tyler, T. R. (2006). Viewing CSI and the threshold of guilt: Managing truth and justice in reality and fiction. *The Yale Law Journal*, 115, 1050-1085.
- Wegener, D. T., Kerr, N. L., Fleming, M. A., & Petty, R. E. (2000). Flexible correction of juror judgments. *Psychology, Public Policy, and Law*, 6, 629-654.
- Wegener, D. T.; Petty, R. E. (1997). The flexible correction model: The role of naive theories of bias in bias correction . In Zanna, Mark P. (Ed). *Advances in experimental social psychology*, Vol. 29. (pp. 141-208). San Diego, CA, US: Academic Press.

The CSI Effect: An Examination of the Source of Biases toward Forensic Science

Laura J. Hanby & Craig Bennell
Carleton University, Ottawa, ON
E-mail: lhanby@connect.carleton.ca

The proliferation of crime dramas on television has led to a dialogue in the popular media about the ostensibly named “CSI Effect”. In the media, the CSI Effect is typically referred to as the influence of forensic programming on the public’s perceptions of forensic science (Justis, 2006). News media reports frequently cite the occurrence of the CSI Effect and present isolated cases in which it appears that jurors have been influenced by shows like *CSI* (e.g., by assuming that DNA evidence should be available at all crime scenes). Without any real evidence of its existence, actors within the criminal justice system are already changing the manner in which they carry out their duties to compensate for the CSI Effect (Patry, Stinson, & Smith, in press).

Anecdotal evidence of this phenomenon has led researchers to examine the CSI Effect. For example, in response to the reliance on anecdotal support for the CSI Effect, Patry, Stinson, and Smith (in press) undertook three studies to examine the CSI Effect in terms of its influence on potential jurors. First, jury-eligible adults were surveyed to determine the public’s opinion on scientific and non-scientific evidence. Scientific evidence was clearly preferred, while non-scientific evidence was seen as less useful and reliable. Second, the relationship between television viewing habits and beliefs about forensic evidence were examined through a survey. Participants that watch *CSI* and *Law and Order* programs had more favourable views regarding scientific evidence. Lastly, the causal relationship between viewing crime dramas and attitudes about forensic evidence was tested by randomly assigning undergraduate students to view four, eight, or no episodes of *CSI*. After viewing the episodes, a survey found that participants who watched *CSI* had higher estimates of the accuracy and reliability of DNA and fingerprint analysis than those that did not. The researchers believe that the findings of these

studies provide evidence for the existence of a CSI Effect in Canada.

The series of studies conducted by Patry, Stinson, and Smith (in press) are comprehensive and provide insight into many of the questions that have failed to be addressed previously. However, this and other research is still plagued by the problem of omitting important preliminary questions that need to be answered. In all of the studies of the CSI Effect, the assumption is made that jurors have developed biases as a direct result of television programs such as *CSI*. However, it has not been established whether the portrayal of forensic science on *CSI* is inaccurate. Before research can begin to address if the CSI Effect exists, how it surfaces, and why it occurs, this needs to be clarified. If *CSI* is deemed to be relatively accurate, it is possible that the so-called CSI Effect has been mislabelled – jurors may have biases toward forensic evidence, but the source of that bias is not likely to be *CSI*.

The purpose of this study is to determine whether *CSI* is inaccurate in its presentation of forensic science. Specifically, two research questions are addressed. First, the way in which *CSI* portrays forensic evidence will be examined. The second research question involves a preliminary examination of whether *CSI*’s depiction of the field of forensic science is inaccurate, as presumed in previous research.

Method

Stage 1: Determining the Content of CSI

Procedure and Materials. A content analysis was conducted on the 23 episodes from season three of *CSI: Crime Scene Investigation*. A coding manual was designed to determine how forensic evidence, tools, and procedures are depicted on *CSI*. For each forensic technique, the timeframe in terms of the length of time it took to obtain the results was coded. The conclusions that could be made with each piece of forensic evidence were also

assessed. Lastly, the coding manual focused on how criminal cases on *CSI* are portrayed more generally (e.g., how cases are resolved).

Inter-rater Reliability. The primary investigator coded all of the episodes, while the second rater coded a random selection of five episodes. The level of agreement for each episode ranged from 65 to 86%. The average agreement for each individual item coded was 76%. This level of agreement was deemed to be sufficient to continue with the content analysis.

Stage 2: Assessing the Accuracy of CSI

Participants. The respondents in the questionnaire stage of the study consisted of professionals working in the field of forensic science, belonging to either the American Academy of Forensic Sciences or the Canadian Society of Forensic Science. The response rate to the questionnaire was extremely low, resulting in a sample consisting of six forensic experts (50% male and 50% female).¹ The age of the respondents ranged from 20 to 60 years, with a mean age of 38.5 years ($SD = 18.87$). The fields of work of the sample included Pathology/Biology, Criminalistics, Education, and Forensic Nursing.

Procedure and Materials. The questionnaire was generated from the results of the content analysis and made available to participants on a survey website. The objective of the questionnaire was to assess the accuracy of the content portrayed on *CSI* through a comparison with the experiences of forensic experts. The survey consisted of three sections: personal characteristics, forensic evidence and techniques, and case characteristics. The respondents were asked to rate, on a scale from 1 (strongly agree) to 5 (strongly disagree), the extent to which they agreed with the way in which the evidence and techniques are depicted on *CSI*. The questionnaire also aimed to determine whether the way in which cases were generally portrayed on *CSI* was in line with real world experiences.

¹Due to the method of recruiting participants, it is impossible to estimate the exact response rate. It is unknown how many people were reached by email through the AAFS or viewed the recruitment letter on the CSFS website. Therefore, the response rate cannot be calculated.

Results

The Content of CSI

Case Characteristics. The content analysis of season three of *CSI* determined that murder was the most frequent type of crime depicted (90% of cases). In contrast, sexual assault, property crimes, assaults, accidents, and suicides accounted for only 10% of cases handled by forensic scientists. The typical resolution of all crimes was that the case was solved as a result of scientific evidence (88% of cases). In the remaining 12% of cases, the case was resolved as a result of another reason (e.g., a confession) or was not solved by the end of the episode. In 81% of cases, the speed with which cases were resolved was estimated to be between one and seven days. In the remaining 19% of cases, the speed with which cases were resolved was split approximately equally between less than 24 hours and greater than seven days.

Forensic Evidence, Tools and Procedures. The most frequently appearing type of evidence was DNA evidence. The other commonly portrayed types of evidence were blood, fingerprint, and drug evidence. The types of tools and procedures ranged from simple visual comparisons to the more complex, such as Lansberry's Ridge Builder, a liquid which temporarily "plumps out" ridges on fingerprints. The timeframe to obtain results from the various tools and procedures was typically immediate or less than 48 hours. In most cases, the evidence was either precisely indicative of its source or it gave some information regarding its source. Typically there was a consensus among experts regarding the interpretation of the evidence or there was no discussion of how the evidence could be interpreted differently.

The Accuracy of CSI

Case Characteristics. All of the respondents to the questionnaire disagreed with the way *CSI* portrays the frequency of cases in which there would be an in-depth investigation. While murder represents the overwhelming majority of cases presented on *CSI*, only two of the respondents ranked murder as the most frequent crime that would result in a crime scene investigation. Of the six respondents, four did not believe that the portrayal of the outcomes for murder cases was accurate. Two of

these four respondents believed that the typical outcome was a resolution due to another reason (such as a confession), while two believed that the most frequent outcome of a murder was an unsolved case (see Table 1). Four of the respondents did not agree with the depiction of the speed in which criminal cases are resolved on *CSI*, while two were unsure. Of the four respondents who did not agree with *CSI*, all indicated that the majority of cases take more than seven days to solve.

Table 1. Rankings by Forensic Experts of the Frequency of Outcomes for Murder Cases.

Outcome of Murder Case	Respondent			
	A	B	C	D
Case solved as a result of scientific evidence	2	2	3	3
Case solved as a result of another reason	1	1	2	2
Case is not solved	3	3	1	1

Note. A ranking of 1 represents the most frequent outcome and a ranking of 3 represents the least frequent outcome.

Forensic Evidence, Tools, and Procedures. The group of experts did not have a high level of agreement or disagreement with many of the statements (see Table 2). When examining the results within each type of evidence, there is a lack of clear patterns. However, when the survey responses are examined across types of evidence, some clear patterns emerge. Most notably, the responses that range from somewhat to strongly agree are typically related to the usage of various tools and procedures on *CSI* with particular types of evidence. For example, there was a high agreement with statements such as “luminol is used to detect blood” and “tape lifts are used to collect fingerprints”. The majority of the statements that the experts disagreed with relate to the timeframe to obtain the results of the various tools and procedures. The other areas of disagreement were in the preciseness and interpretation of evidence. Lastly, five of the six respondents disagreed with the portrayal on *CSI*

that conclusions made with various types of evidence do not depend on that evidence being of low or high quality.

Discussion

The purpose of this study was to determine the degree of accuracy of *CSI* in its presentation of the field of forensic science. It is evident that there is a lot of information presented on *CSI*. In some respects, the portrayals of forensic science are accurate. For instance, the types of tools and procedures used by real forensic experts correspond to those portrayed on *CSI*. In other regards, the depictions on *CSI* are inaccurate. For example, the timeframe to obtain results and the ability to make precise conclusions using these forensic techniques are distorted on *CSI*. This implies that the issue of the CSI Effect is not quite as simple as suggested. Within *CSI*, there are inconsistencies in the degree of accuracy in its portrayal of forensic science. Thus, we cannot simply ask: Is *CSI* accurate? We have to ask: In what ways is *CSI* accurate and inaccurate?

Limitations of the Study

The content analysis conducted in the current study was useful in determining how *CSI* portrays forensic science. However, the results may not be generalizable to other forensic programming or even other seasons of *CSI*. This is an important point because potential jurors who view *CSI* are also likely watching other forensic programs, such as *Law & Order* and *Criminal Minds*. These viewers may also watch factual forensic programs or read academic books, and it would be important to determine the accuracy of this material and examine whether these more accurate sources of information counteract *CSI*-type effects.

In relation to the content analysis, there are also potential reliability issues. For example, while inter-rater agreement was generally satisfactory, some coding domains posed problems. It was particularly difficult to code temporal variables, such as the time to obtain results from forensic techniques, because there was no objective indication of time in the *CSI* episodes.

Table 2. Descriptive Statistics of Survey Responses.

Type of Evidence Survey Question	N	Mean	Standard Deviation
Blood evidence			
Luminol is used to detect blood	6	1.67	1.21
The results of Luminol are obtained immediately	6	3.00	1.26
Luminol provides some information of the source of the blood	6	3.83	1.17
Phenolphthalein is used to detect blood	5	2.80	1.10
The results of phenolphthalein are obtained immediately	5	2.60	1.34
Phenolphthalein precisely indicates the source of the blood	5	4.20	1.30
Bloodstain pattern analysis produces immediate conclusions	6	4.33	1.21
Bloodstain pattern analysis provides some information regarding the source of blood	6	2.67	1.21
Blood evidence can also be used to estimate time of death	6	4.33	1.21
Estimating time of death with blood evidence can be done immediately	6	4.83	0.41
Estimating time of death with blood provides some information in terms of an approximate time period of death	6	3.33	1.51
There is consensus among experts regarding the interpretation of blood evidence	6	4.17	1.17
Bullet casing			
Ballistics analysis takes less than 48 hours with bullet casings	5	4.20	1.30
This procedure give some information regarding the source of a bullet casing	5	2.40	1.14
There is consensus among experts regarding the interpretation of bullet casings	5	3.00	1.58
Bullet			
Ballistics analysis takes less than 48 hours with bullets	5	4.20	1.30
Bullet evidence is precisely indicative of its source using ballistics analysis	5	3.60	1.34
There is consensus among experts regarding the interpretation of bullet evidence	4	3.75	1.26
Projectile trajectory analysis is used with bullets	4	2.25	0.96
The time frame to obtain results with projectile trajectory analysis is immediate	5	4.20	0.45
Bullet evidence is precisely indicative of its source using projectile trajectory analysis	5	3.60	1.34
Type of Evidence	N	Mean	Standard Deviation

Survey Question			
DNA			
DNA tests take less than 48 hours to complete	6	3.83	1.17
DNA evidence is precisely indicative of its source, either positively or negatively	6	3.67	1.51
There is consensus among experts regarding the interpretation of DNA evidence	5	3.20	1.30
Drug			
The presence of drugs is determined by toxicology	5	1.80	1.30
A toxicology report takes less than 48 hours	5	4.80	0.45
Toxicology provides precise conclusions regarding the particular drugs present	5	3.20	1.64
There is consensus among experts regarding the interpretation of drug evidence	5	3.00	1.00
Fibre			
Visual inspections and comparisons can be used to identify the source of fibres	5	2.20	0.48
Conclusions are made immediately with visual inspections and comparisons	5	4.00	0.71
Precise conclusions can be made in terms of the source of fibres	5	3.80	1.30
There is consensus among experts regarding the interpretation of fibres	5	4.00	1.00
Fingerprint			
The results of Lansberry's Ridge Builder are obtained immediately	3	3.67	1.52
Tape lifts are used to collect fingerprints	5	1.80	0.84
Tape lifts are completed immediately	5	3.20	1.10
Visual comparisons of fingerprint evidence allows for precise conclusions to be made as to its source	5	3.40	1.52
The AFIS produces results in less than 48 hours	5	3.20	1.64
The AFIS precisely indicates the source of fingerprints	5	4.00	1.22
Experts are in agreement on the interpretation of fingerprint evidence	5	3.80	1.10
Firearm			
Ballistics analysis takes less than 48 hours with firearms	5	3.80	1.10
Precise conclusions can be made about the source of firearms with ballistics analysis	5	3.20	1.30
There is consensus among experts regarding the interpretation of the same evidence	4	3.25	1.50
Footprint			

An ALS is used to detect footprints	4	2.75	0.96
A ALS produces immediate results	4	3.50	1.29
An ALS gives some information regarding the source of footprint evidence	4	3.25	0.96
Type of Evidence			
Survey Question	N	Mean	Standard Deviation
Footprint			
A tape lift is used to collect footprint evidence	4	2.50	1.29
A tape lift takes less than 48 hours to complete	4	2.50	1.73
Casting/molding is used to collect footprint evidence	4	1.75	0.50
Casting/molding takes less than 48 hours to complete	4	3.00	1.83
A visual comparison of footprint evidence gives some information regarding the source	4	1.75	0.50
There is consensus among experts regarding the interpretation of footprint evidence	4	3.50	1.29
Gunshot residue			
Gunshot residue can be detected immediately	5	3.20	1.64
Precise conclusions can be made as to the presence of gunshot residue	5	4.00	1.22
Hair			
The results of visual inspections or comparisons of hair are immediate	6	4.50	1.22
Hair evidence gives some information regarding its source	6	2.33	0.82
Experts are in agreement on the interpretation of hair evidence	6	3.33	1.21
Impression evidence			
Visual comparisons are used with impression evidence	6	2.17	0.98
The results of visual comparisons are concluded immediately	6	4.17	0.75
Visual comparison allow for precise conclusions to be made regarding the source of the evidence	6	3.83	1.17
Insect			
Forensic entomology takes less than 48 hours to reach conclusions	6	4.33	1.21
Insect evidence provides some information regarding the details of a crime	6	2.33	1.37
Rape kit/semen			
An ALS is used to detect semen	5	2.00	0.71
The length of time to obtain results with an ALS is immediate	5	2.60	1.34
An ALS precisely indicates the presence of semen	5	3.20	1.30

There is consensus among experts regarding the interpretation of semen using an ALS	5	3.20	1.30
Stomach contents			
Visual inspections are done on stomach contents	5	1.60	0.89
Precise conclusions regarding the source can be made with stomach contents	5	3.40	1.52
There is consensus among experts regarding the interpretation of stomach content evidence	5	3.00	1.58
Type of Evidence			
Survey Question	N	Mean	Standard Deviation
Teeth/bite marks			
Casting/molding is used to make impressions of teeth	3	1.33	0.58
The time frame of casting/molding and visual comparisons is immediate	4	4.50	0.58
Teeth impressions are able to allow for precise conclusions using casting/molding and visual comparisons	4	3.50	1.29
Tire tracks			
A computer program is used to identify the source of tire tracks	3	3.00	1.00
Using a computer program with tire tracks takes less than 48 hours to obtain results	3	4.00	1.00
The results of a computer program provide a precise indication of the source of tire tracks	3	3.67	1.53
Tire tracks can be visually compared or inspected	4	2.00	0
A visual comparison of tire tracks is an immediate technique	4	4.50	0.58
A visual comparison results in precise conclusions regarding the source of the tire tracks	4	4.00	1.41
A visual inspection of tire tracks provides some information regarding its source	4	2.50	1.00
Experts are in agreement on the interpretation of tire track evidence	4	4.00	0.82
Urine			
An ALS is used to detect urine	4	2.00	0.82
The length of time to obtain results with an ALS is immediate	4	3.00	1.41
An ALS gives some information regarding the source of urine evidence	4	3.25	0.96
Wound			
A wound is used to estimate time of death	6	3.83	1.47
Estimating time of death based on a wound is an immediate process	6	4.50	0.84

Some information can be obtained by estimating time of death with a wound	6	3.83	0.98
There is consensus among experts in the interpretation of estimating time of death with a wound	6	4.00	0.89
Visual inspections or comparison provide some information regarding the source of a wound	6	2.00	0.63
There is consensus among experts in the interpretation of a wound using visual inspections or comparisons	6	3.67	1.21

Note. Ratings were provided on a scale from 1 (strongly agree) to 5 (strongly disagree).

The survey presented to forensic experts provided the basis for further examination regarding the accuracy of forensic programming on *CSI*. However, while this approach appears to have some value, one of the main limitations to the findings in the current study was the extremely low sample size of experts that participated, and the subsequent problems this caused for interpreting responses. The small sample size could be a result of a number of problems with the survey, which should be remedied in future research. For example, the survey used in the current study dealt with a broad range of forensic issues and very few, if any; forensic scientists have expertise across forensic sub-disciplines. Therefore, many potential respondents likely felt unable to respond to the range of questions included on the survey. Due to the fact that the sample size was so low, the responses provided by the forensic scientists may not be representative opinions and should be interpreted with a degree of caution.

Future Research

Future research on the CSI Effect should first attempt to replicate this study, while remedying its limitations. Specifically, a content analysis of all types of forensic programming would allow for generalizability to all *CSI*-type programs. The findings of an updated content analysis could be used to revise the survey used in this study. In addition, instead of relying on an online questionnaire, conducting focus groups with forensic experts may prove more productive when trying to get at the accuracy of forensic programs. Each focus group could pertain to a different forensic sub-discipline, which would likely result in a larger sample of forensic experts and more meaningful results.

From this line of research, experimental studies with mock jurors could be pursued to examine the CSI Effect. First and foremost, experimental research must determine whether jurors have biases toward the field of forensic science that influence their decision-making. The precise conditions under which these biases influence behaviour should also be identified by examining such variables as the types of offence and evidence presented at trial. Second, in order to term this phenomenon the CSI Effect, it must be established that the biases jurors have are based on viewing *CSI* and other forensic programming. At this point, we should not rule out the possibility that *CSI*-type shows could be educating viewers. That is, instead of having a negative impact on jurors by encouraging false expectations, forensic programming could be educating jurors on the complex field of forensic science.

Conclusion

This research represents a preliminary examination of the source of biases toward forensic science that makes up the so-called CSI Effect. *CSI* is rich with information, both verbally and nonverbally conveyed, which may influence the attitudes of viewers. This study found that, although some aspects of *CSI* are inaccurate, others are accurate. This study provides the foundation to determine whether jurors have biases, whether they are a result of viewing *CSI*, and whether they are beneficial or detrimental to the decision making process of jurors.

References

Justis, G. G. (2006). *Images of legitimacy: Presentation of forensic programming in contemporary news*

publications. Unpublished master's thesis. Michigan State University, East Lansing.

Patry, M.W., Stinson, V., & Smith, S.M. (in press). The *CSI Effect*: Is popular television transforming Canadian society? In Greenberg, J., & Elliott, C (eds.). *Communications in Question: Canadian Perspectives on Controversial Issues in Communication Studies*. Scarborough, Ontario: Thompson-Nelson.

Blurring the Line between Fact and Fiction: Expert Opinions about Forensic Investigation Tools Represented on CSI

Marc W. Patry
St. Mary's University
E-mail: marc.patry@smu.ca

Veronica Stinson
St. Mary's University

Steven M. Smith
St. Mary's University

Tony McCulloch
Royal Canadian Mounted Police

People who follow the media are likely to have heard of the so-called *CSI effect*. *CSI: Crime Scene Investigation* and its related spin-offs are among the most popular television shows in North America (www.nielsenmedia.com). In recent years, newspapers and television news programs (e.g., *USA Today*, *the Toronto Star*, *CNN*) have reported on people's beliefs that *CSI* and other television crime dramas (e.g., *Law & Order*, *Bones*, *Crossing Jordan*) influence how people think about and behave relative to the legal system. As mentioned above, typically, the news media has referred to this influence as the *CSI effect*.

Recently, dozens of newspaper and magazine articles and TV news shows have reported on the *CSI Effect*. Examples of some of the headlines from articles about the *CSI effect* include "'CSI effect' has juries wanting more evidence" (Willing, 2006) and "'The CSI Effect': Does the TV crime drama influence how jurors think?" (CBS News, 2005). The results of one study in which news media reports about the *CSI Effect* were subjected to content analysis indicated that the news media reports frequently characterized the *CSI effect* as being negative, and assumed that juror expectations were being affected by the unrealistic portrayals of crime scene investigation on television crime dramas (see Patry et al., in press). Most references to the *CSI Effect* reflect the notion that it is an undesirable effect exhibited by jurors which results from their perceived expertise about forensic techniques and police investigations (see Tyler, 2006; Podlas, 2006). These reports suggest that watching *CSI* causes jurors to have unrealistic expectations about the quantity, and availability of scientific evidence. When the scientific evidence presented at trial fails to meet jurors' expectations, they are presumably more likely to acquit the defendant.

More specifically, our research shows that the *CSI effect* tended to be described in the popular media as having one of four impacts (Patry et al., in press). First, some articles described television crime dramas as having piqued student interest in topics such as biology, anthropology, psychology and other forensically relevant topics, and this has lead to increased enrolments in these courses and programs at Colleges and Universities. A second category the articles fell into claimed that these shows have actually educated criminals in how to engage in criminal activity without getting caught.

The third theme of stories in the media reflects how jurors are purportedly influenced by *CSI*. Often, specific cases were mentioned. For example, jurors in the Robert Blake case (the actor who was charged but acquitted of murdering his wife) cited a lack of gunshot residue as a main reason for acquitting the actor (Keating, 2005). In another case, a man was acquitted because a jury foreman, a fan of *CSI*, convinced the other jurors to acquit because of lack of fingerprint evidence (Lotstein, 2004). A final general category of articles focused on how lawyers and other legal professionals have responded to the perceived *CSI Effect*. Most of these stories have focused on how lawyers have changed their behavior in court to address the *CSI effect* and its purported influence on jury decision-making.

This preoccupation with a possible effect of television crime dramas on our legal system is not limited to the popular press but rather reflects a genuine concern about this issue emerging from the legal community. In 2005, the Maricopa County Attorney's Office released a study about the *CSI Effect* (Maricopa County, 2005). Their study of 102 prosecuting attorneys yielded overwhelming evidence that prosecutors believe these TV crime dramas affect jury verdicts. The report cites a number of cases after which jurors

cited “*CSI*-related reasons” for their verdicts (e.g., the prosecution should have provided more physical evidence such as DNA or fingerprint). There is also evidence that these programs are influencing prosecutors’ trial strategies and pre-trial negotiations. The Maricopa County study shows that “52% of the prosecutors have engaged in plea negotiations...because they anticipated problems [due to] the *CSI* Effect,” (p. 8). Furthermore, “83% [of prosecutors] agree that jury instructions should include directing jurors not to use outside standards like those used in forensic crime television shows when making judgments of guilt or innocence,” (p.10). Our preliminary research with Canadian law enforcement professionals and lawyers demonstrates that they are also concerned about the *CSI* Effect but they are unsure of its exact nature or impact (Stinson et al., 2007).

In general, there is not a great deal of empirical evidence to support the conceptualizations of the *CSI* effect put forward by the media. Given this general lack of evidence, some scholars have argued that there is no *CSI* effect, and if there is one, that it is equally probable that it favors the prosecution (e.g., Tyler, 2006). If researchers are to examine carefully the *CSI* effect, an important early step is understand the truthful and fictional aspects of the program.

In a prior study, we conducted a content analysis of the first seasons of *CSI* and *CSI: Miami* (see Patry et al., in press). We sought to document the types of forensic procedures portrayed, the frequency of errors, and the types of sentiments expressed by characters on the show, such as the theme that scientific evidence, when properly gathered, leads to the truth. We identified over 75 forensic evidence techniques portrayed on the shows. The two most common types of forensic evidence were DNA evidence, which was present in 18.9% of the main story lines, and fingerprint evidence, which appeared 12% of the time. There was a consistent theme that “science is the only truth”. Finally, the perpetrator of the crime was successfully identified in 98% of the storylines. This is related to the fact that on *CSI*, evidence which conclusively points to the guilt of one suspect is almost always available at the crime scene – a condition which is far less common in real life. Importantly, on the *CSI* episodes we

examined, forensic investigators conducted scientific tests 72% of the time, whereas in real life, specialized laboratory technicians conduct scientific tests of forensic evidence (T. McCullough, personal communication, 2005). Thus, there is a clear difference between actual forensic investigations and the way that these investigations are portrayed on *CSI* shows.

So, it seems clear that *CSI* programming makes unrealistic portrayals of forensic techniques. But are the techniques themselves realistic? The present study was designed to shed light on the ways in which the information presented in *CSI* programs does and does not reflect reality, in order to develop a better sense of whether these programs are likely to cause bias among the general public.

Method

Participants

This study was a survey of 15 forensic experts employed by the Royal Canadian Mounted Police (RCMP), including RCMP staff officers as well as independent civilian experts who regularly consult with the RCMP. The median age of sample participants was 48; the median number of years of experience was 13.5. Participants were selected for their expertise in specific areas of forensic analysis: domains of forensic expertise represented in these data included: identification services, major crimes unit, police dog service, blood stain analysis, audio and video analysis, facial identification artistry, firearms, biology, anthropology, odontology, traffic, entomology, and forward looking infrared.

Survey Instrument and Sampling Method

The objective of the study was to develop a comprehensive catalogue of expert opinions about the realism of various forensic techniques represented on *CSI*.

We developed a survey about numerous specific forensic techniques represented on the television crime drama program *CSI*. The survey represented 73 techniques that were portrayed on *CSI*, as identified in our prior content analysis of the program itself (see Patry et al., in press). In addition to its name, each technique in the master

survey included a brief description of the technique as portrayed on *CSI*.

The survey was broken up based on forensic content area and relevant portions of the overall survey distributed to experts in the respective content areas. Our objective was to obtain evaluations of techniques presented on the program by forensic experts in content areas closely aligned with the domain of expertise represented by the technique. In other words, experts commented only on those forensic techniques within their individual realms of expertise

Results

Consistent with our prior research on forensic investigators and police officers (Stinson et al., 2007), forensic experts in the current study reported fairly high changes in the public's expectation of their profession as a result of television crime dramas such as *CSI*, see Table 1. Experts in the current study reported somewhat less extreme ratings of crime drama-related changes in their interactions with the public, and in their professional roles including their behaviour in court (see Table 1) as compared to our prior finding with generally less specialized law enforcement professionals (Stinson et al., 2007). Consistent with our prior research, experts in the current study had low ratings of the accuracy of television crime dramas in terms of what happens in actual cases: $M = 2.5$ on a 7-point scale, see Table 1.

See Table 2 for the aggregated expert ratings of the 73 different forensic techniques represented in the survey (overall). The realism of the procedures was well above the median of the 7-point scale ($M = 4.6$), and the scientific research supporting the use of the techniques was very high ($M = 5.9$). The reliability/accuracy of the techniques, however, were rated quite low on the scale ($M = 1.9$).

Discussion

The current study utilized content-area experts in forensic evidence to help identify realistic and fictional aspects of the forensic techniques represented on *CSI*. Overall, forensic techniques portrayed on *CSI* have some basis in realism. This

Table 1. Expert Perceptions of CSI Effect

Item	Rating Mean (SD)
In your opinion, has the public's expectations of your profession changed due to these shows?	M = 5.0 (1.6)
In your opinion, have TV crime dramas changed the way in which you interact with the public?	M = 2.9 (1.8)
In your opinion, have TV crime dramas influenced your work or the way you do your job?	M = 2.9 (1.9)
In your opinion, have the public's expectations affected your behaviour in court?	M = 3.6 (1.0)
In your opinion, how accurate are TV crime dramas in portraying what happens in real court cases and investigations?	M = 2.5 (2.0)

Note. All measures were taken on seven-point bipolar scales ranging from 'not at all' to 'completely.'

Table 2. Expert Ratings of 73 Forensic Techniques Portrayed on CSI

Item	Rating Mean (SD)
Would you say the procedure described above is realistic?	M = 4.6 (2.0)
Is there scientific research supporting the use of such a technique?	M = 5.9 (1.6)
In your opinion, does this technique produce reliable or accurate results?	M = 1.9 (0.3)

Note. All measures were taken on seven-point bipolar scales ranging from 'not at all' to 'completely.'

is evidenced by the expert ratings of the realism of the procedures, and especially by their high ratings of the scientific research supporting the techniques. According to this survey of content-area experts, the techniques are generally technically feasible and are supported by research.

However, it is clear from these data that the reliability and accuracy of scientific techniques is distorted on these programs. This is consistent with our prior research with police officers and death investigators indicating that there are major distortions in terms of the ways in which the techniques are portrayed on *CSI* (e.g., timing, personnel, availability) (see Patry et al., 2007).

We conclude that the blurry line between fact and fiction lies just beyond the forensic tests themselves. While the tests themselves are technically feasible and their use is supported by research, according to this sample of content-area experts, the accuracy and reliability of the techniques are far more limited than they ways in which they are generally presented on *CSI*.

What implications does this have in terms of public awareness of police procedures? It is possible that viewers of these programs are 'learning' both accurate and inaccurate information about forensic evidence. Viewers of the program may be learning a good deal of valid information from the program, but they may also be inculcated with biases about the accuracy and reliability of these scientific tools. What is clear from the present study is that a good deal of the information presented on *CSI* has a firm basis in reality, but that this realism is countered by distortions of the reliability of the science as portrayed on the program. Further research is needed in order to pinpoint the specific ways in which this type of programming may impact public perceptions in general, and jury decision making more specifically.

References

Keating, J. (Aug. 26, 2005). *Lawyers debate a possible "CSI Effect" on Juries*. Inland Valley Daily Bulletin. MediaNews Group: Ontario, Canada.

Lotstien, B (2004). Fact and Fiction collide in the Jury Box. *Prosecutor Magazine*, 38.

Maricopa County (2005). *The CSI Effect and its Real-Life Impact on Justice: A Study by the Maricopa County*

Attorney's Office (June 30, 2005). Retrieved on September 20, 2005 from http://www.maricopacountyattorney.org/Press/PDF/CSI_Report.pdf

Patry, M. W., Stinson, V., & Smith, S. M. (in press). CSI Effect: Is popular television transforming Canadian society? In Greenberg, J., & Elliott, C. (eds.). *Communications in Question: Canadian Perspectives on Controversial Issues in Communication Studies*. Scarborough, Ontario: Thompson-Nelson.

Podlas, K. (2006). "The CSI Effect": Exposing the media myth. *Fordham Intellectual Property, Media and Entertainment Law Journal*, . 429-465.

Stinson, V., Patry, M. W., & Smith, S. M. (2007). The CSI Effect: Reflections from Police and Forensic Investigators. *Canadian Journal of Police and Security Services*, 5, 125-133.

Tyler, T. R. (2006). Viewing CSI and the threshold of guilt: Managing truth and justice in reality and fiction. *The Yale Law Journal*, 115, 1050-1085.

Willing, R. (2006, August 5). 'CSI effect' has juries wanting more evidence. *USA Today*, Retrieved August 10, 2006 from http://www.usatoday.com/news/nation/2004-08-05-csi-effect_x.htm.

Linking Serial Rapes: A Test of the Behavioural Frequency Hypothesis

Tamara Melnyk, Craig Bennell, & Natalie J. Jones
Carleton University
E-mail: tmelnyk@connect.carleton.ca

In the absence of physical evidence, investigators must primarily rely on offence behaviours when determining whether several crimes are linked to a common offender. For this to be possible, offenders must exhibit high levels of behavioural similarity across their respective crime series, while at the same time exhibiting high levels of behavioural distinctiveness (i.e., behaviours that are not exhibited by other offenders committing similar types of crimes) (Bennell, 2002; Bennell & Canter, 2002; Bennell & Jones, 2005; Canter et al., 1991; Woodhams & Toye, 2007). Researchers have argued that the level of behavioural similarity and distinctiveness that can be found, and therefore the degree of linking accuracy that can be achieved, will vary as a function of the behaviours being considered. Specifically, claims have been made that low frequency behaviours should be focused on when attempting to link crimes.

For example, Canter, Bennell, Alison, and Reddy (2003) suggested that the distinguishing features of a crime are most likely to be behaviours that are relatively rare, since higher frequency behaviours tend to be conceptually central to a given crime type and are likely evidenced by most offenders (e.g., certain sexual behaviours in the case of rapes). Specifically, Canter et al. (2003) stated that, “when attempting to link a potential series of unsolved crimes to the offender responsible, care should be taken to avoid spurious links that may result if too much weight is given to the sexual [i.e., high frequency] behaviours exhibited by the offender(s)” (p. 171). Similar arguments related to the frequency issue have been made by other researchers as well (e.g., Canter, Alison, Alison, & Wentink, 2004; Canter & Fritzon, 1998; Canter & Heritage, 1990; Salfati & Bateman, 2005; Salfati & Dupont, 2006; Woodhams, Hollin, & Bull, 2007).

While this argument may appear sensible on the surface, it is an assumption that has yet to be

tested empirically. The current study made an attempt to test this assumption by examining the level of linking accuracy that can be achieved in a sample of serial rapes using rape behaviours that vary with respect to their frequency of occurrence in the sample (i.e., low, moderate, and high frequencies). An attempt was therefore made to address whether it is possible to accurately link serial rapes by relying solely on the crime scene behaviours exhibited by offenders, as well as whether the ability to link serial rapes increases when using low frequency rape behaviours, as compared to moderate or high frequency behaviours.

Method

Sample

The current study analyzed 126 offences committed by 42 UK serial rapists (three crimes per offender). All of the data were extracted from victim statements prepared by police officers as part of criminal investigations. Twenty-seven variables relating directly to the behaviour of the offender at the scene of the crime were identified through content analysis of victim statements by trained researchers. For each crime, behaviours were either coded as 1 (present) or 0 (absent). The content categories were initially derived from the published literature on rape and from thorough analysis of the victim statements. The 27 dichotomous variables coded across the 126 offences were separated into three subsets representing “low”, “moderate”, and “high” frequency behaviours (see Table 1). These behaviours provided the basis upon which the subsequent analysis was conducted.

Analysis

The criterion variable in the analysis was whether a pair of crimes was committed by the same offender or different offenders. Jaccard's coefficient of similarity was calculated across

Table 1. Summary List of Variables

	Frequency		
	Low (<20%)	Moderate (20-50%)	High (>50%)
Demeans victim (18)		Forces victim to actively participate (39)	Uses surprise attack (92)
Tears victim's clothing (16)		Kisses the victim (37)	Vaginal penetration (83)
Uses disguise (16)		Steals unidentifiable items (35)	Makes sexual comments (52)
Gags victim (15)		Uses single violence (32)	Uses a weapon (52)
Anal penetration (14)		Threatens the victim not to report offence (29)	
Uses multiple violence (13)		Attempts to identify the victim (26)	
Compliments victim (10)		Blindfolds victim (25)	
Steals personal items (10)		Demands goods from victim (25)	
Implies knowing victim (10)		Forces victim to perform fellatio (25)	
Steals identifiable items (8)		Binds victim (21)	
Forces victim to make sexual comments (6)		Performs Cunnilingus (20)	

Note: Numbers in brackets refer to the percentage of cases where the behaviour was exhibited.

each crime pair using low, moderate, or high frequency behaviours. Receiver operating characteristic (ROC) graphs were then constructed to examine the degree to which the different sets of behaviours yielded different degrees of linking accuracy. The area under the ROC curve (i.e., the AUC) was calculated for each ROC curve and reflects the level of linking accuracy that was achieved using a specific subset of behaviours. The AUC can range from 0.50, indicating chance linking accuracy, to 1.00, indicating perfect linking accuracy.

Results

Prior to the linkage analysis, descriptive statistics

were calculated across all identified linked and unlinked crime pairs for each frequency set. Significance testing was conducted on this data to determine if linked crime pairs were associated with significantly higher levels of behavioural similarity than unlinked crime pairs. Results demonstrated that linked crime pairs were consistently associated with significantly higher similarity scores than unlinked crime pairs. However, cases did exist in which linked crimes were characterized by relatively low levels of across-crime similarity and unlinked crimes by relatively high levels of across-crime similarity. Thus, linking accuracy will not be perfect regardless of the behaviours considered.

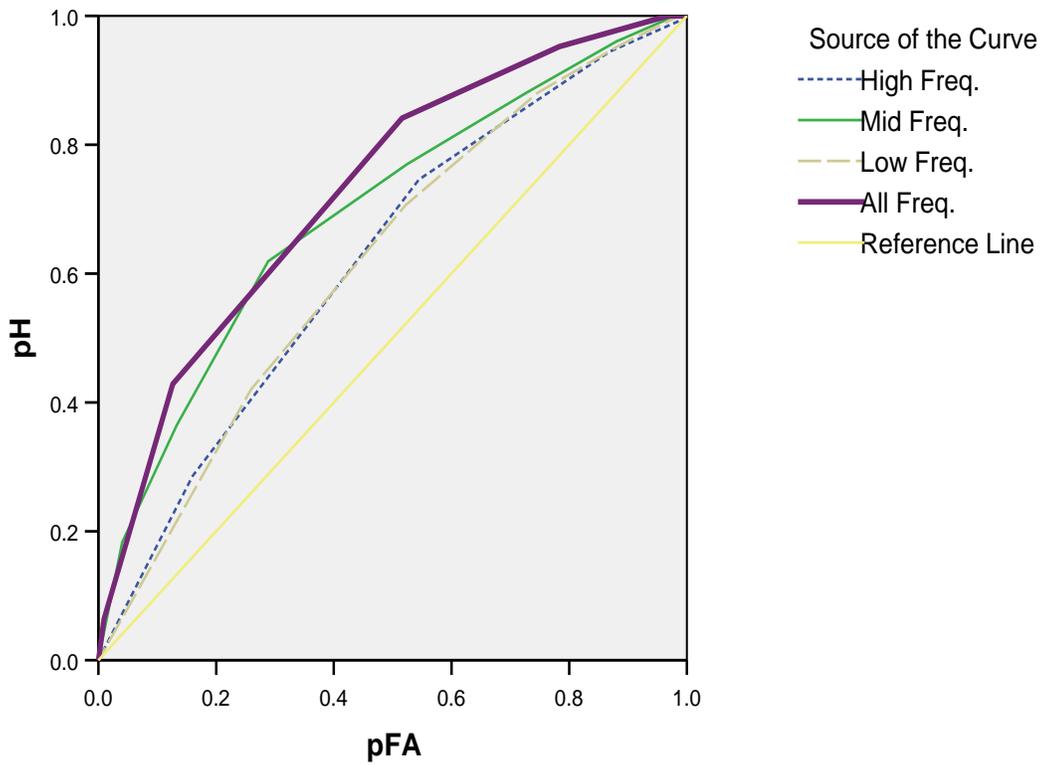


Figure 1. ROC curves representing the degree of linking accuracy associated with behaviours of varying frequencies.

Table 2. Summary Data from ROC Analysis

Frequency	AUC	Standard Error	95% Confidence Intervals
High	.62	.03	.56 to .67
Moderate	.70	.03	.65 to .75
Low	.62	.02	.58 to .67
All Frequencies	.75	.03	.70 to .80

ROC curves derived from the across-crime similarity scores are presented in Figure 1. The results confirm that it is possible to discriminate between linked and unlinked crimes significantly beyond chance level. This is the case with behaviours from all frequency sets. However, the highest levels of linking accuracy were associated with the use of all variables, followed by

moderate frequency variables, and high and low frequency variables. The summary data from the ROC analysis is presented in Table 2.

Discussion

Results from the current study indicate that the assumed importance of low frequency behaviours

for linkage analysis may be unfounded, as low frequency rape behaviours did not result in the highest levels of linking accuracy. This finding suggests that investigators should refrain from relying too heavily on low frequency crime scene behaviours when attempting to determine whether several crimes can be linked to a single offender and should instead use all available crime scene behaviours for this purpose. If behavioural frequency is going to be considered, the results from the current study suggest that moderate, rather than low, frequency behaviours should be relied on.

A number of potential avenues for future research emerge from this study. First, it would be important to replicate the results presented here using larger samples and different crime types. Second, researchers should explore the potential impact of using different similarity coefficients for the purpose of linking crimes (i.e., a coefficient other than Jaccard's) since different coefficients (e.g., the simple matching coefficient) will respond differently to behavioural frequencies. Finally, efforts should be made to identify other moderators of linking accuracy beyond behavioural frequencies, such as the degree to which crime scene behaviours are situation-dependent, since this will allow us to maximize success on the linking task.

References

- Bennell, C. (2002). *Behavioural consistency and discrimination in serial burglary*. Unpublished doctoral dissertation, University of Liverpool, Liverpool, UK.
- Bennell, C., & Canter, D. V. (2002). Linking commercial burglaries by modus operandi: Tests using regression and ROC analysis. *Science & Justice*, 42(3), 153-164.
- Bennell, C., & Jones, N. J. (2005). Between a ROC and a hard place: A method for linking serial burglaries by modus operandi. *Journal of Investigative Psychology and Offender Profiling*, 2(1), 23-41.
- Canter, D. V., Alison, L. J., Alison, E. K., & Wentink, N. (2004). The Organized/ Disorganized typology of serial murder: Myth or model? *Psychology, Public Policy, and Law*, 10(3), 293-320.
- Canter, D. V., Bennell, C., Alison, L. J., Reddy, S. (2003). Differentiating sex offences: A behaviorally based thematic classification of stranger rapes. *Behavioral Sciences and the Law*, 21, 157-174.
- Canter, D. V., & Fritzon, K. (1998). Differentiating arsonists: A model of firesetting actions and characteristics. *Legal and Criminological Psychology*, 3(Part 1), 73-96.
- Canter, D. V., & Heritage, R. (1990). A multivariate model of sexual offence behaviour: Developments in 'offender profiling'. *The Journal of Forensic Psychiatry*, 1, 185-212.
- Canter, D. V., Heritage, R., Wilson, M., Davies, A., Kirby, S., Holden, R., et al. (1991). *A facet approach to offender profiling*. London, UK: Home Office.
- Salfati, C. G., & Bateman, A. L. (2005). Serial homicide: An investigation of behavioural consistency. *Journal of Investigative Psychology and Offender Profiling*, 2(2), 121-144.
- Salfati, C. G., & Dupont, F. (2006). Canadian homicide: An investigation of crime-scene actions. *Homicide Studies*, 10(2), 118-139.
- Woodhams, J., Hollin, C. R., & Bull, R. (2007). The psychology of linking crimes: A review of the evidence. *Legal and Criminological Psychology*, 12(2), 233-249.
- Woodhams, J., & Toye, K. (2007). An empirical test of the assumptions of case linkage and offender profiling with serial commercial robberies. *Psychology, Public Policy, and Law*, 13, 59-85.

The influence of witness age, relation to crime, and eyewitness identification decision on jurors' perceptions and verdicts

Julie L. Dempsey & Joanna D. Pozzulo
Carleton University
E-mail: jdempsey@connect.carleton.ca

Jurors' perceptions of a witness' credibility and evaluation of the witness' testimony may be influenced by witness characteristics such as age of the witness (child vs. adult) and the witness' relation to the crime (victim vs. bystander). In addition, jurors' perceived accuracy of the identification decision made by the witness may influence how accurate other details in their testimony are perceived, and ultimately influence verdict.

Witness Age and Credibility

The prosecution of cases may rely heavily, if not exclusively, on the testimony of a child provided he/she is the only bystander/victim to the crime. Children, in comparison to adults, may be considered inferior witnesses based on the notion that children may be more likely to make false claims or are prone to suggestion and fantasy (e.g., Ceci & Bruck, 1993). However, child witnesses can be perceived as more or less credible as adult witnesses depending on whether their testimony confirms or disconfirms the stereotype of children as memory sources. Various studies have examined how the age of a witness impacts perceived witness credibility and final verdict. For example, Goodman and colleagues (1987) found that a child eyewitness was rated as less credible in comparison to the older eyewitnesses and credibility of the eyewitness increased as a function of age. Guilty verdicts, however, were *not* found to differ as function of the eyewitness' age. Conversely, Ross and colleagues (1990) found children to be viewed as more credible than adults by jurors. Given, the mixed results, witness' age alone does not appear to be predictive of perceived credibility and verdict. Rather, the inconsistent results of the two studies may be a function of other witness factors (Goodman et al., 1984) such as the witness' relation to the crime; bystander versus victim.

Witness Age and Relation to Crime

The testimony provided by an eyewitness may be the recollections of either a victim or a bystander (i.e., an individual who observed the event but who was not directly involved in the crime). Given the mixed results regarding child witness credibility it is possible that the eyewitness' relationship to the crime may influence a juror's perception of witness credibility. For instance, when a child is involved in a sexual assault case, the child is typically the victim of the alleged abuse. In cases such as this, jurors tend to view a child witness as more credible than either adolescent or adult witnesses (Ross, Jurden, Lindsay, & Keeney, 2003). However, child witnesses tend to be viewed more negatively and less credible in cases that do not involve sexual abuse, where typically the child is not the victim of the crime, but rather an observer of the event (Goodman et al., 1987). It is possible that jurors may consider a victim to be more accurate in their recollection of events because the victim plays such a central role in the event. Conversely, jurors may believe a bystander may be more accurate in their testimony because the bystander would not have experienced as much stress or arousal during the event and therefore would be more reliable. At present it is unclear how jurors interpret these different roles and how these roles impact upon witness credibility and verdict.

Type of Identification Decision

Lineup identification is used to gain proof of a culprit's identity beyond the verbal description provided by the witness (Wells, 1993). Three types of lineup identification decisions are possible; a positive identification (i.e., the suspect of the crime is identified by the witness), a non identification (i.e., the witness rejects the entire lineup and does not identify anyone), or a foil identification (i.e., the witness identifies an

individual known to be innocent). The only known error by a witness is a foil identification.

In reality, it is not possible to know the accuracy of either a positive or non identification. It is the responsibility of the jurors to determine the accuracy of the identification decision. Previous research investigating the influence of eyewitness identification on juror perceptions has primarily focused on cases where the eyewitness made a positive identification. However, all types of identification decisions may end up in court.

Eyewitness identification evidence has been found to be influential on jury verdicts (Skolnick & Shaw, 2001). Jurors tend to attribute a high level of credibility to eyewitness identification evidence which may be problematic given the fact that such identification decisions have repeatedly been shown to be vulnerable to various errors including perception and memory errors (Skolnick & Shaw, 2001). It is also possible that adults may be perceived as more accurate in lineup identification decisions in comparison to children (Parker & Ryan, 1993). However, numerous studies have indicated that children are able to correctly identify the criminal, shown a target-present lineup (i.e., the criminal's picture is among the photos presented to the witness), at a rate comparable to adults (Pozzulo & Lindsay, 1998). Given that these misconceptions of children's identification abilities exist, it is possible that a juror may consider an identification decision provided by a child witness to be less accurate compared to an adult witness, regardless of what the identification decision may be.

Method

Participants

Adults ($N = 363$, 129 males and 234 females) were recruited from the first year psychology participant pool and from upper year courses at an Eastern Ontario University, Canada (age range from 18 to 56 years, $M = 21.51$ years, $SD = 4.82$).

Design

This study used a 2 (age of witness; child vs. adult) x 2 (crime relation; bystander vs. victim) x

3 (identification type; positive vs. non vs. foil identification) between-subjects factorial design.

Materials

Transcripts. The mock-trial transcripts consisted of an excerpt of a trial in which a drug deal was interrupted, subsequently killing an ice-rink arena attendant. A second-degree murder charge was applied to the defendant thus requiring a jury trial. The portion of the trial provided to mock-jurors dealt exclusively with the eyewitness testimony concerning the drug deal. The trial transcript included the witness first being questioned by the prosecutor, followed by a cross-examination by the defense attorney. Witness age (10 vs. 40 years old), relation to the crime (victim vs. bystander), and identification decision (positive vs. non vs. foil identification) were crossed to produce 12 transcripts. In addition to the identification testimony, bystanders/victims provided testimony that included a description of the culprit, description of the culprit's jacket, time of the drug deal, and description of the item passed between the culprits. Testimony on these details was held constant across the transcripts.

Reliability of crime details. Using a 6-point scale (1 = not at all reliable; 6 = very reliable), jurors rated how accurate they perceived the crime details to be that were provided by the witness; description of the culprit, description of the culprit's jacket, time of the drug deal, identification decision, description of the item passed between the culprits.

Witness credibility. Using a 6-point scale (1 = not at all credible; 6 = very credible), jurors were asked to rate the overall credibility of the witness.

Verdict. Jurors were asked to render a guilty or not guilty verdict for the defendant (i.e., the person accused of committing the murder at the ice-arena).

Results

Verdict

A Logit loglinear analysis was conducted to examine the influence of age (i.e., child vs. adult), relation to crime (i.e., victim vs. bystander), and type of identification decision (i.e., positive vs.

non vs. foil) on verdict. A main effect of type of identification decision on verdict was significant, $Z = -3.33, p < .01$. Mock-jurors rendered more guilty verdicts when a witness made a positive identification decision than a foil identification decision, $\chi^2(1, n = 249) = 36.89, p < .01$, and a non identification decision, $\chi^2(1, n = 250) = 22.36, p < .01$. Guilty verdicts were made at a comparable rate when the witness made a foil identification vs. a non identification.

Perceived Accuracy of Identification Decision

Type of identification decision was found to influence the perceived accuracy of the identification decision made by the witness, $F(2, 375) = 66.00, p < .05$. Mock-jurors perceived positive identification decisions ($M = 4.27, SE = .11$) as more reliable than non identification decisions ($M = 3.47, SE = .11$; mean difference = $-.81, p < .01$) and more reliable than foil identification decisions ($M = 2.53, SE = .11$; mean difference = $-1.75, p < .01$). Foil identifications were viewed as less accurate than both non identifications and positive identifications.

Perceived Accuracy of the Criminal's Description

Type of identification decision was also found to influence the perceived accuracy of the criminal's description provided by the witness, $F(2, 375) = 66.00, p < .05$. Mock-jurors perceived the criminal's description as more accurate when the witness made a positive identification ($M = 4.35, SE = .09$) than a foil identification ($M = 3.77, SE = .90$; mean difference = $-.57$). Also, mock-jurors perceived the criminal's description as more accurate when the witness made a non identification ($M = 4.12, SE = .90$) than a foil identification (mean difference = $-.36$). No significant difference was found between positive and non identifications (mean difference = $-.22, ns$).

Perceived Credibility of the Witness

The child victim was viewed as equally credible to the adult victim, $t(1,189) = 1.00, ns$. However, when the child was a bystander, the child was viewed as less credible than the adult who was a bystander, $t(1, 182) = 4.54, p < .01$. The adult

victim and adult bystander did not receive different credibility ratings. However, the child victim was viewed as more credible than the child bystander.

Discussion

The present study found that the child bystander was viewed by mock jurors as less credible than the child victim. The same difference was not found for adults. If crimes involve a child, the child may be viewed as more credible than if the child is merely a bystander. Jurors may hold negative stereotypes of a child's ability to accurately report witnessed events, these negative stereotypes in turn may impact perceptions of children testifying (Eaton, Ball, & O'Callaghan, 2001). Social psychology research has found that stereotypes tend to have a strong influence on social judgment (Ross et al., 1990). Thus, jurors may hold preconceived notions and doubts in regards to the credibility of child eyewitnesses testifying (Goodman et al., 1984). However, age is not considered in isolation of other key factors.

The present study found that age interacted with the witness' relation to the crime to influence a witness' perceived credibility. Adult and child victims were viewed as equally credible. Jurors may view a victim's testimony as more reliable as a result of the victim's salient role in the crime. Jurors may believe that a victim would have a more accurate recollection of the criminal and details pertaining to the crime because the victim was directly involved in the crime and this may hold true for child and adult victims. Further research is required to examine why relation to crime may effect a juror's perceptions.

In addition, type of eyewitness identification decision was found to significantly affect a jurors' perceived accuracy of the lineup decision and other testimony about the culprit provided by the witness, specifically, the culprit's description. In reality, the accuracy of a positive and non identification cannot be determined, in and of itself. A foil identification however, is a known identification error. Mock-jurors appeared somewhat cognizant of these distinctions. Overall, mock-jurors seemed to place the accuracy of the

various types of identification decisions on a continuum with positive identification as most reliable and foil identifications as least reliable. Non identifications fell between the other two types of identification.

More intriguing was to examine whether jurors make a link between recognition performance and recall. Previous research illustrates that these two memory tasks are primarily independent, with little information to be gleaned about one's recall performance based on his/her recognition performance (Wells, 1984). In the current study, type of identification decision (recognition) made by a witness influenced how reliable testimony for other crime details (recall) were perceived. Mock jurors perceived a witness' description of the culprit as more accurate when a positive or non identification was made than a foil identification. Thus, when a known identification error is made by a witness, his/her other testimony about the culprit may be discounted by jurors. A non identification was viewed similarly to a positive identification by mock-jurors when evaluating the culprit's description.

Of the factors examined, only type of identification decision influenced verdict. Not surprisingly, more guilty verdicts were rendered when a positive identification was made than a foil or non identification. Not guilty verdicts were comparable between foil and non identifications.

Although each of the factors examined in the present study had an individual effect on jurors' perceptions and verdict, it is crucial to consider the complex interaction between each of these factors that were found. When evaluating evidence, jurors consider these factors in conjunction with each other. It is problematic to infer that any one of these individual factors contribute to or have a greater impact on a juror's evaluation of witness testimony. Rather, it is more important to recognize the impact these factors have in combination.

References

Ceci, S. J., & Bruck, M. (1993). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin*, 113, 403-437.

- Eaton, T. J., Ball, P.J., & O'Callaghan, M. G. (2001). Child-witness and defendant credibility: child evidence presentation mode and judicial instructions. *Journal of Applied Social Psychology*, 31, 1845-1858.
- Goodman, G. S., Golding, J. M., & Haith, M. M. (1984). Juror's reactions to child witnesses. *Journal of Social Issues*, 40, 139-156.
- Goodman, G. S., Golding, J. M., Helgson, V. S., Haith, M. M., & Michelli, J. (1987). When a child takes the stand: Juror's perceptions of Children's Eyewitness Testimony. *Law and Human Behavior*, 11, 27-40.
- Parker, J. F. & Ryan, V. (1993). An attempt to reduce guessing behavior in children's and adults' eyewitness identification. *Law and Human Behavior*, 17, 11-26.
- Pozzulo, J. D. & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, 22, 549-570.
- Ross, D. F., Dunning, D., Toglia, M. P., & Ceci, S. J. (1990). The child in the eyes of the jury: assessing mock jurors' perceptions of the child witness. *Law and Human Behavior*, 14, 5-23.
- Ross, D. F., Jurden, F. H., Lindsay, R. C. L., & Keeney, J. M. (2003). Replications and limitations of a two-factor model of child witness credibility. *Journal of Applied Social Psychology*, 33, 418-431.
- Skolnick, P. & Shaw, J. I. (2001). A comparison of eyewitness and physical evidence on mock-juror decision making. *Criminal Justice and Behavior*, 28, 614-630.
- Wells, G.L., (1984). The psychology of lineup identifications. *Journal of Applied Social Psychology*, 14, 89-103.
- Wells, G.L. (1993). What do we know about eyewitness identification? *American Psychologist*, 48, 553-571.

Preschoolers' Person Description and Identification Accuracy: A Comparison of the Simultaneous and Elimination Lineup Procedures

Charmagne Crescini & Joanna D. Pozzulo
Carleton University
E-mail: jpozzulo@connect.carleton.ca

Over the past few decades, much research has been devoted to children's capacities as eyewitnesses (for a review, see Ceci & Bruck, 1995). Past research has indicated that even very young children are capable of accurately recalling events and persons previously encountered (e.g., Marin, Holmes, Guth, & Kovac, 1979). Further, previous studies have shown that children are not only able to accurately recall and describe such persons, but they are also able to recognize the person out of a lineup (e.g., Parker & Ryan, 1993). Children over the age of five have been found to be as accurate as adults when presented with a target-present simultaneous lineup (Pozzulo & Lindsay, 1998). The problem for child witnesses arises when they are faced with a target-absent lineup. Whereas adults are capable of correctly rejecting the lineup, children (aged 5 and up) are likely to make an identification, resulting in a false positive identification (i.e., identifying an innocent person).

In an attempt to improve children's identification accuracy, Pozzulo and Lindsay (1999) developed an elimination lineup procedure specifically for child witnesses. The elimination lineup procedure requires two separate judgments. First, the witness is asked to make a relative judgment by selecting the person most similar to the culprit relative to the other lineup members. Second, the witness is asked to make an absolute judgment by comparing the chosen lineup member to his/her memory of the culprit and decide if it is in fact the culprit. It was suggested that children often fail to make an absolute judgment and thereby produce greater false-positives than adult witnesses (Pozzulo & Lindsay, 1999). An absolute judgment is especially critical when faced with a target-absent lineup; identification accuracy (i.e., correct rejection) depends on whether the witness realizes that the most similar looking lineup member is not the true culprit. Thus, the elimination lineup that explicitly requests an absolute judgment should

effectively increase children's identification accuracy, particularly by reducing false positive rates when faced with target-absent lineups.

In the original study, Pozzulo and Lindsay (1999) proposed several variations of the elimination lineup and compared it with the simultaneous lineup. They found that children's (10- to 14-year-olds) correct identification rates when presented with any of the elimination lineup procedures did not differ considerably from those observed when children were presented with a simultaneous lineup. More importantly however, all the variations of the elimination lineup procedure significantly reduced children's false positive rates compared to the simultaneous lineup procedure (.15 - .27 vs. .46, respectively; Pozzulo and Lindsay, 1999). Recently, Pozzulo and Balfour (2006) found the elimination procedure to be more effective than the simultaneous procedure at increasing correct rejection rates among adults and children (aged 8-13 years) when the lineup members matched the culprit's appearance at the time of the crime. These studies demonstrate the effectiveness of the elimination lineup in reducing false positives among older children. However, no study has yet investigated the elimination procedure among younger children and it is unclear if similar low false positive rates would be achieved with this age group.

The present study examined the identification accuracy of 4- to 6-year-olds with the elimination lineup procedure compared to the simultaneous lineup. It was predicted that both the elimination and simultaneous lineups would produce comparable correct identification rates when the target is present in the lineup. For target-absent lineups, it was predicted that the elimination lineup would produce a higher correct rejection rate compared to the simultaneous lineup. The present study also investigated the nature, quantity, and accuracy of person descriptions provided by the preschoolers. Consistent with

previous research (Marin et al., 1979), preschoolers were expected to report few descriptors (1-2 on average). Despite the few descriptors however, it was predicted that these would be accurate. Lastly, gender differences in terms of the accuracy of descriptions provided were also explored.

Method

Participants

One hundred preschoolers ($M = 59.10$ months, range = 39 - 78 months, $SD = 8.07$ months) were recruited from 19 preschools/daycare centres in Eastern Ontario, Canada. Five additional children who originally participated were excluded in the final analyses due to incomplete data. Approximately equal numbers of boys and girls participated in the study (59 and 41, respectively) and the sample was predominantly Caucasian (77%).

Design

A 2 (presence of target; present vs. absent) \times 2 (lineup procedure; elimination vs. simultaneous) between-subjects factorial design was used.

Procedure

Children in small groups of up to five were given a mask-making demonstration (approximately 20 minutes) by a female confederate who acted as the mask-making teacher. After the demonstration, participants joined other children in the regular routine. After a 20-minute delay, the preschool teacher announced that two of the mask-making teacher's helpers (i.e., research assistants) were there to ask about the mask-making session. The research assistants then interviewed the children individually, asking them to describe the mask-making teacher using an open-ended format. Specifically, children were asked, "Remember the mask-making teacher, what did she look like?" Once the child responded, s/he was further prompted, "Do you remember anything else about the mask-making teacher?" Finally, the children were randomly assigned to view one of the four lineup conditions and their responses were recorded accordingly.

Lineup Construction

Photos of volunteers who resembled the female confederate were taken to construct a six person lineup. The photos were frontal, head and shoulder shots, printed in color (4"x6"). The same photos were used in both lineups. Target-present lineups included a photo of the confederate while in target-absent lineups, her photo was replaced with a similar one. All the foils and the confederate wore different coloured clothing in the photos from that worn by the confederate during the mask-making sessions in order to avoid a potential clothing bias. The order of presentation was held constant, with the confederate or her replacement always presented in position 4.

Results

Identification Accuracy

Correct identification rates were higher, but not significantly different with the elimination procedure (0.68) compared to the simultaneous procedure (0.44), $\chi^2(1, n = 25) = 2.92, ns$.

A significantly higher correct rejection rate was obtained with the elimination procedure (0.80) compared to the simultaneous lineup (0.52), $\chi^2(1, n = 25) = 4.37, p < .05$.

Nature, Quantity, and Accuracy of Descriptors

Hair colour was the most frequently reported descriptor, mentioned by 58% of the participants. This was followed by clothing colour (47%), hair length (20%), and clothing type (17%). Other descriptors reported were height (7%), complexion (6%), eyes (1%), body type (1%), and accessories (1%).

In terms of the quantity of descriptors, the children reported an average of 1.57 descriptors. The number of descriptors did not vary by gender, $t(98) = .236, ns$, with males reporting an average of 1.6 descriptors and females an average of 1.5 descriptors.

A multivariate analysis of variance (MANOVA) was performed on these three dependent variables (DVs) with gender as the independent variable (IV). Using Hotelling's T^2 , there was a non-significant association between the combined DVs and the IV, $F(3, 96) = 1.28, ns$. Girls and boys did not significantly differ in the number of correct

descriptors ($M = .85$ for girls; $M = 1.14$ for boys), incorrect descriptors ($M = .10$ for girls; $M = .07$ for boys), or intrusions ($M = .68$ for girls; $M = .46$ for boys) reported.

Discussion

Consistent with previous findings (Pozzulo & Lindsay, 1999; Pozzulo & Balfour, 2006), correct identification rates of preschoolers presented with target-present lineups did not significantly differ as a function of lineup procedure. Although the elimination lineup was not specifically designed to increase correct identification rates, it is notable that it produced a higher, albeit not significantly different correct identification rate than the simultaneous lineup. Future studies could investigate how the elimination lineup may increase correct identification rates among child witnesses.

Across target-absent conditions, the elimination lineup produced a significantly higher correct rejection rate than the simultaneous lineup for preschoolers. As Pozzulo and Lindsay (1999) speculated, it is possible that children may not be aware of implicit lineup instructions such as having the option to completely reject the lineup, or at the very least, express uncertainty with their decision. Thus, the dual-judgment process of the elimination procedure that explicitly asked witnesses for a second, absolute judgment (i.e., whether the lineup member selected is in fact the target) may have facilitated children's lineup decision-making, and consequently, reduced false-positive responding.

In terms of the quantity of descriptions provided, the preschoolers reported on average, 1.57 descriptors, consistent with past findings that younger children report very few details (Marin et al., 1979). Despite the few descriptors mentioned, the preschoolers were accurate in the descriptors that they provided. Consistent with Pozzulo and Warren (2003), the preschoolers were most likely to report hair items, with hair colour being reported by more than half of the participants. Further, clothing color and type, as well as hair length were the items most frequently reported. Pozzulo and Warren argued that the tendency to report such items could reflect a developmental trend wherein exterior features (e.g., hair, clothing, etc.) may be more salient and hence,

focused on. With increasing age, interior features (e.g., eyes, nose) may be observed and incorporated into the descriptions. Thus, considering the young age of the participants in the present study, it is understandable that almost all of the children disregarded interior facial features.

Lastly, the present study found no effects of gender on the quantity or accuracy of the preschoolers' descriptions of the teacher. Girls and boys reported similar rates of correct descriptors, incorrect descriptors, and intrusions.

Implications

Overall, the present study provides evidence that the elimination lineup procedure may be effectively implemented for preschool-aged children in order to reduce false-positive responding. The elimination lineup could potentially eliminate concerns that young children's identification decisions are less accurate than identification decisions of older children and adult witnesses. Further, the present study indicates that even very young children are capable of providing accurate descriptions and should therefore be considered as credible witnesses.

References

- Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom*. Washington, DC: American Psychological Association.
- Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law and Human Behavior*, 3, 295-305.
- Pozzulo, J. D., & Balfour, J. (2006). Children's and adults' eyewitness identification accuracy when a culprit changes his appearance: Comparing simultaneous and elimination lineup procedures. *Legal and Criminological Psychology*, 11, 25-34.
- Pozzulo, J. D., & Lindsay, R. C. L. (1998). Identification accuracy of children versus adults: A meta-analysis. *Law and Human Behavior*, 22, 549-570.
- Pozzulo, J. D., & Lindsay, R. C. L. (1999). Elimination lineups: An improved identification procedure for child witnesses. *Journal of Applied Psychology*, 84, 167-176.
- Pozzulo, J. D., & Warren, K. L. (2003). Descriptions and identifications of strangers by youth and adult witnesses. *Journal of Applied Psychology*, 88, 315-323.

Section E

Professional & Ethical Issues

Ethical Dilemmas in Practicing Correctional Psychology

James McGuire
University of Liverpool
E-mail: merc@liv.ac.uk

During a typical career spent in forensic or correctional psychology, whether in criminal justice agencies or secure mental health services, it will be virtually impossible to avoid confronting ethical problems at some stage. Indeed such problems arise with remarkable frequency. The place where individuals who in all probability have acted against the interests of others, meet with organisations allotted the complex responsibility of simultaneously controlling and rehabilitating them, is one that is likely to generate many intricate questions of this kind.

This paper has three objectives. First, the opening section will highlight the nature of the ethical dilemmas likely to arise in this and adjoining areas of professional practise. Second, we will briefly examine the philosophical bases of ethical action in correctional and clinical psychology. Several formalised frameworks and approaches have been proposed for supporting practitioners in their efforts to understand and address issues of this kind: to bridge the gap, as it were, between relatively abstract moral precepts and concrete instances of everyday ethical problems. The third objective then is to present and briefly review some approaches and procedures that have been proposed for bridging that gap in day-to-day practice.

Surveys of ethical problems

Self-report studies of ethical dilemmas described by practicing psychologists and by psychology graduate students yield a reasonably consistent picture regarding the principal kinds of difficulties that individuals experience. The largest such study was conducted under the auspices of the American Psychological Association and entailed distribution of a questionnaire to 1319 of its members (Pope & Vetter 1992). The response rate was 51%, and the principal dilemmas reported were in the areas of confidentiality, dual or conflicting relationships, aspects of fee payment and related issues, and five per cent of the sample

reported specific dilemmas in relation to forensic work. A parallel study was subsequently undertaken with members of the British Psychological Society (Lindsay & Colley 1995), and despite a lower response rate (28%) a broadly similar pattern of concerns emerged, though a surprising number of respondents stated they experienced no ethical dilemmas. Working on behalf of the Belgian Psychological Association and with a different population, Nederlandt (1996) surveyed 100 graduate students pursuing research degrees in psychology. Again confidentiality and allied issues emerged as important, though there were also some context-specific issues, a function of respondents' status at the stage of their careers. Another more recent survey has been reported of the extent to which psychologists are provided with or participate in ethics education during their professional studies. Davidson, Garton and Joyce (2003) surveyed 35 accredited psychology programs in Australian universities and found that the primary focus of such ethics teaching as was provided focused on research ethics, and in the majority of cases, this aspect of psychology was not assessed at any stage. The authors called for a revision of the way ethics education was conducted and forwarded reasons for sustaining and expanding it; arguments that were amplified by Miner (2005).

Table 1 displays the findings of another very small-scale and informal survey carried out by the present author, of a cohort (n=25) of graduate psychologists pursuing a programme that leads to a qualification as a clinical psychologist working in the UK's National Health Service. Whilst only a proportion of their work brings them into contact with forensic services (and some go on to specialise in this area), it is evident that even the direct provision of psychological services via a large organisation can generate ethical unease in numerous ways.

Moral principles and codes of conduct

Table 1. Ethical problems experienced by trainee clinical psychologists ($n = 25$)

-
- Receiving/being offered gifts by service users
 - Personal boundaries (where clients act inappropriately)
 - Flirtatious clients
 - Help being withdrawn from someone who needs it
 - Challenging one's "superiors"
 - Prioritizing clients where resources are limited
 - Friendship / professional boundaries
 - Status and competence of colleagues
 - Lack of clarity concerning uses of information (purpose, source)
 - Confidentiality versus action in risk situations
 - How to be aware of the limits of own competence
 - Breach / power
 - Audio- and video-taping (client consent, comfort)
 - Disagreements with employer's policies (e.g. exploring risk issues)
 - Inappropriate behaviour – competence and confidence to address it
 - Unprofessional conduct: criticism being given out of context
 - Inaction by agencies in cases of need
 - Actions that are discriminatory against and individual or group
-

The sometimes acute discomfort we experience when encountering an ethical problem is not something that needs to be explicitly trained: on the contrary it appears to arise naturally in most individuals when they encounter difficulties of these kinds. There are long-standing divisions and deeply-held beliefs concerning the potential origins of such reactions within us. Traditionally the domain of moral philosophy, these debates have expanded and are now the subject matter of numerous disciplines. At the risk of gross over-

simplification, the dominant views could be characterised as falling into three broad categories. One widespread - and what might even be termed conventional - view is that these types of thoughts and beliefs and our emotional connection with them are endowed from an external source; a superior being which perhaps is also imbued with powers of having created the universe, and who may or may not have an interest in intervening in human affairs. This is the core assumptions of numerous forms of theism (Dawkins 2006). To many psychologists, whether or not they hold religious beliefs of their own, an individual's stance on moral issues is thought to be a function of cultural and social learning, likely to emerge alongside, or as a function of, other cognitive operations (Tomasello 1999). A more recently emerging viewpoint is that our moral reactions are the product of an evolutionary process, and several authors have forwarded theories concerning how both these and other profound systems of beliefs (for example with respect to religion, or the freedom of the will) are composed (Boyer 2002; Dennett 2004; Hauser 2006).

However, most discussion of key concepts concerning how moral beliefs are constituted still occurs within the framework of philosophical ethics. Some key distinctions within this domain include that between *normative* ethics and *practical* ethics (Singer 1993). The former refers to an inquiry concerning the basic values or norms that govern morality from which elementary ethical principles can be derived. The latter entails consideration of the implications of such principles for individual conduct in specific instances. Another core distinction to some extent reflects disputes mentioned above over the origin of moral response, and corresponds to two major "schools of thought" or traditions in the philosophical study of ethics. One approach is grounded in *utilitarianism*, the view that moral questions can be decided on the basis of their consequences or outcomes; there are several different bases on which these can be evaluated when making a moral decision. This is attributed to a school of English philosophy whose chief representative is usually taken to be Jeremy Bentham (1742-1832). The other is *deontology*, "a theory that some features of actions other than or

in addition to consequences make actions right or wrong” (Beauchamp & Childress 2001, p.56). This is generally attributed to the thought of the 18th century German philosopher Immanuel Kant (1724-1804).

The ethics codes that have been adopted by most professional associations of psychology worldwide have their roots in the deontological approach to establishing ethical standards. Several essential concepts inform the content of most ethics codes. They include the concepts of *beneficence*: the acceptance of a responsibility to do good; and conversely *non-maleficence*, acceptance of a need to avoid doing harm. The principle of *autonomy* insists on respect for every person’s freedom of thought and action. *Justice* requires professionals to base their actions on principles of fairness between individuals; and *fidelity* requires observance of trustworthiness, reliability and faithfulness to commitments.

Seeking to translate these ideas meaningfully into practise, and in response to a recognition that the application of psychology to human and social problems has important ethical dimensions, most psychological associations have developed ethics codes of varying levels of elaboration. The first professional Code of Ethics for psychologists was prepared by the American Psychological Association in 1953; it has since been revised several times (most recently 2002). The Canadian Psychological Association / Société canadienne de psychologie published the third edition of its Code of Ethics in 2000. The British Psychological Society produced its code in 1985 (subsequently revised in 1995 and 2006) and established an Ethics Committee in 2004. All psychological associations in Europe have their own ethics codes, and their collective “umbrella” organisation, the European Association of Psychologists’ Associations (EFPA) has developed a “meta-code” of ethics for all member societies.

Dimensions of ethical dilemmas in professional psychology

Some of the core dilemmas arising in the field of applied psychology have been discussed by McGuire (1997). In correctional settings, at their core these dilemmas turn upon a question of what

is the basic purpose of a psychologist’s work, and who is the client. Is the psychologist there primarily to address the needs of the individual who has broken (or might break) the law, and to provide counselling, therapy or other services that will promote and sustain psychological well-being? Alternatively, is the pre-eminent objective to protect society, to serve an employer (possibly the state), to assess and manage risk? One way to think of this is as a conflict between “individualist” and “collectivist” orientations towards the application of ethical principles. Clearly, in many circumstances the practitioner will to some extent be seeking to achieve both, as far as that may be possible. It is within our efforts to reconcile these inherent conflicts that moral unease and ethical dilemmas often originate.

Driven by these competing demands, other aspects of the dilemma are also activated. When addressing individual needs, the client or service user is likely to be self-motivated and to seek help on a voluntary basis. By contrast, if legally detained he or she is in a situation that is coerced, and may not be motivated to comply with a psychologist’s requests – and perhaps actively resistant. This has further implications for the boundaries of confidentiality. Under what circumstances can or should information be passed to others, and of what kinds, to whom, and with what justification? The collection and accumulation of information places the correctional psychologist in an influential position, and ethical issues arise from the degree of power he or she then possesses. To what extent can someone who has acted in a therapeutic role with an individual then contribute to decisions made about him or her, and when does this cross the boundary into the role of the expert witness? This conflict may be at its most acute in the area of risk assessment, where both the collection of data at the individual level and the process whereby it is converted into predictions pose substantial ethical challenges (McGuire, 2004).

Not surprisingly, there is a marked divergence of opinion concerning whether the fundamental conflict underpinning these dilemmas, and which pervades many facets of correctional psychology, can be resolved. Is it possible to work in a dual role, for example as psychological therapist or program tutor and also as an evaluator of

someone's progress and potential expert witness? Most commentators insist that many aspects of these and other roles are intrinsically incompatible, and argue against "wearing two hats" (Strasburger, Gutheil & Brodsky 1997). Unquestionably, these are significant difficulties to overcome.

Drawing on a set of recommendations initially made by Shapiro (1979) concerning the accurate recording of information elicited during clinical assessment interviews, McGuire (1997) suggested that if it is possible to obtain accurate description of client problems, an informed awareness of individualist and collectivist positions, and clear definitions of boundaries concerning expectations and responsibilities, the two sets of roles can be combined. Doing so however, and coping with numerous other dilemmas that may manifest themselves in correctional or clinical work, also means having resort to robust systems for analysing ethical questions. But equally, it is possible to argue that questions of this kind should be addressed at a more systemic level, engendering changes in the criminal law that allow it to absorb more person-centered practices and principles, for example by applying the concepts of therapeutic jurisprudence (McGuire 2000).

These disputes notwithstanding, there is a consensus amongst psychologists who have contributed to this area on several general points. First, prior to engaging in professional practise it is important to give thought to ethical issues, rather than being suddenly confronted with them and so feeling totally unprepared, and possibly even experiencing significant distress. There is a parallel argument that it is better to avoid the learning of ethical concepts "by osmosis" from the working environment (Handelsman 1986). Second, many commentators also suggest that addressing ethical concerns cannot be properly accomplished using lectures or other didactic methods (though they may play a small part in it). It is more appropriate to use discussion, perhaps centered on specially devised learning tasks, in conjunction with a problem-solving or problem-oriented framework.

Resolving ethical conflicts

An aspect of self-awareness that is seldom discussed in psychology is the feelings practitioners experience when they sense that something intrinsically wrong has occurred. This is an important, and can sometimes be a distressing reaction. Some psychologists refer to it as the "ordinary moral sense" which can be activated by certain events or circumstances and progress into a condition of "ethical unease". When an individual becomes aware of this sensation, it is appropriate - and it may be vital - to reflect on what is causing it, and review what options there may be for addressing it. It may be important to detach this somewhat from a purely personal reaction, and bring it into focus within a more formal ethical framework. As Smith, McGuire, Abbott and Blau (1991) showed, at least some time ago that was something psychologists often failed to do.

With reference to professional training, given a general consensus that development in this area requires what may be broadly termed a non-didactic approach, several proposals have been forwarded. For example Francis (1999) makes suggestions on stimulating self-awareness of the ethical ramifications of a problem. He suggests asking students to imagine they have such a difficulty and have made an initial decision concerning what to do about it. Next they are invited to imagine the following scenario: that the knowledge of the circumstances becomes public in a court case; the student psychologist is called to the witness box and has to defend his/her decision. The exercise then entails preparing a possible defence of the position adopted. In a variation, participants imagine that for some reason the circumstances have become known to their families, and have to explain their decisions to them. If, having considered an ethical problem, the individual arrived at any specific proposal about how to solve it, subjecting it to the above tests or "thought experiments" (in a safe learning environment) can be a useful developmental and learning process.

When faced with a moral or ethical dilemma in the course of professional practise, what steps can be followed? Are there any agreed procedures that can guide a professional's actions in such a situation? There is no formal consensus regarding this, but a number of proposals have been made

concerning it and the closing section of this paper will briefly present several of them in turn. Authors in this area have developed a number of approaches to practical problem-solving on practical/ethical questions.

First, Canter, Bennett, Jones and Nagy (1994) and Pryzwansky and Wendt (1999) describe a series of preventive steps that can be taken to address ethical problems arising in everyday practice. They describe this as an “ethical practice model” and a summary of its main ingredients is shown in Table 2. This is a general orientation within which practitioners are alerted to the ethical dimensions of their work and it could be characterized as offering a “protective” framework. When ethical problems arise, more specific procedures are then activated.

Table 2. Ethical practice model (Canter et al, 1996; Pryzwansky & Wendt, 1999)

-
- Know the ethics code applicable to the situation in question.
 - Know the legal framework or statutes that have a potential bearing on this issue.
 - Know the rules and regulations where you work (agency policies, etc.).
 - Be involved in continuing education in ethics.
 - Identify ethical problems when they arise: this involves both personal awareness and awareness of the context. Prompt yourself to ask ethical questions in work settings.
 - Learn or develop a method for analysing ethical obligations or the dimensions of ethical aspects of a situation.
 - Consult with colleagues and other professionals knowledgeable about ethical matters.
-

For example, Keith-Spiegel and Koocher (1998) have developed a problem-solving approach and the sequence of steps suggested by them is depicted in Table 3. The approach of Tymchuk (1986), closely allied to that recommended by Sinclair, Poizner, Gilmour-Barrett and Randall (1987), is shown in Table 4, and entails a form of Socratic questioning rather than a series of recommendations. Packard (1997, cited in Pryzwansky & Wendt, 1999) has specified the

Table 3. A problem-solving approach (Keith-Spiegel & Koocher, 1998)

-
- Describe the parameters of the situation; determine that the matter is an ethical one.
 - Consult the guidelines, if any, that are already available and might apply to the resolution of each issue; identify possible mechanisms for resolution.
 - Consider, as best possible, all the sources that might influence the kind of decision you will make.
 - Locate a trusted colleague with whom you can consult.
 - Evaluate the rights, responsibilities, welfare and vulnerabilities of all affected parties.
 - Generate the alternative decisions possible for each issue.
 - Enumerate the consequences of making each decision.
 - Make the decision.
-

Table 4. The approach of Tymchuk (1986) and Sinclair, Poizner, Gilmour-Barrett & Randall (1987)

-
- Indicate the individuals and/or groups that need to be considered in arriving at a solution to the dilemma posed by the situation.
 - Take each of the individuals or groups that you feel should be considered, and explain in detail what consideration each is owed and why, particularly in terms of the rights and responsibilities involved.
 - What would be your choice of action, and why would you choose it?
 - What alternative choices of action would you consider, and why would you not choose them?
 - What is the minimal change of circumstances in the situation that you can conceive of that could lead you to a different course of action? What would that action be, and why would you choose to do it?
 - Do you have any further thoughts or comments about this or similar situations? If so, explain.
-

Table 5. Components of an ethical decision-making model (adapted from Packard, 1997, cited in Pryzwansky & Wendt 1999)

1. Identify the known facts of the situation.
2. Assess your personal response ("ordinary moral sense").
3. Review the relevant Ethics Code carefully.
4. Check for relevant legal requirements.
5. Review Agency/Organisational policies.
6. Apply general moral principles as concretely as possible.
7. Ask the deontological and teleological questions.
8. Consider how a "virtuous" person might respond to the situation.
9. Identify and analyse possible conflicts between codes, laws, policies, moral principles, and ethical theories.
10. Identify and analyse possible conflicts between codes, laws, policies, moral principles, and ethical theories
11. Consult regularly with trusted and respected colleagues.
12. Be aware of and take into account relevant cultural and diversity issues.
13. Construct alternative courses of action.
14. Act by implementing the alternative that according to your thoughtful and reflective analysis seems most credible.
15. Review the consequences of your actions and be prepared to re-involve yourself in the decision-making process if the issue is not resolved.
16. Keep in mind that the decision-making process is interactive, reciprocal, and systemic and not a set of linear steps as the previous listing implies.

components of an ethical decision-making model in perhaps greater detail than in the other approaches represented here and this is illustrated in Table 5. Finally, Haas and Malouf (1995) devised a decision-making flowchart that indicates possible pathways through the ethical jungle and crucial choice-points along each route. Their format for this is shown in Figure 1.

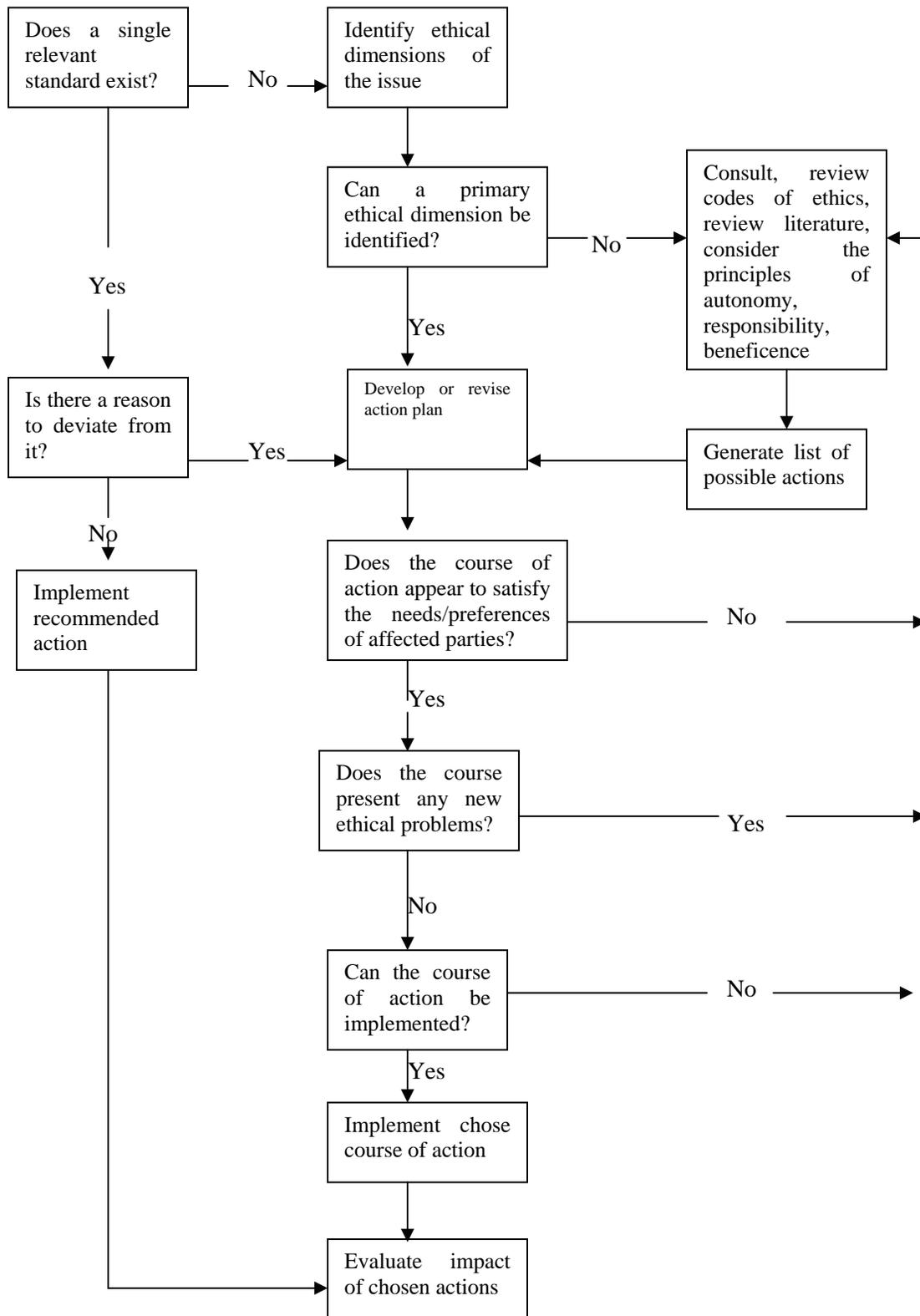
The approaches summarised in Tables 3-5 and in Figure 1 have certain features in common. They make a distinction between ethical problems that (relatively speaking) require straightforward courses of action, either because the dimensions of the situation are sufficiently clear, or because the relevant ethical code or agency policy specifies what should be done, on the one hand; and ethical problems that are not covered by existing guidelines, or are more complex, on the other. They all emphasize careful reflection, and the consideration of alternatives at every stage.

They also uniformly recommend not trying to solve problems of this kind alone: it is strongly advisable always to consult and seek support from others. Ethical dilemmas are not only intellectually but also emotionally demanding. The frameworks collected here offer significant advantages in attempting to reduce confusion and unease and where possible to find constructive solutions to professional ethical problems.

Ethical aspects of applied psychology have been increasingly widely recognised in recent years. Several excellent volumes now exist that provide general introductions to the area (Pope & Vasquez 1998; Pryzwansky & Wendt 1999). Some include in-depth analysis of specific issues (Koocher & Keith-Spiegel 1998), or incorporate readings on key themes (Bersoff 2007). Similar issues of course arise in psychiatry, as illustrated in another recent volume (Green & Bloch 2006).

Dilemmas in Correctional Psychology

Figure 1. A decision-making flowchart (Haas & Malouf, 1995)



References

- Beachamp, T. L. & Childress, J. F. (1994). *Principles of Biomedical Ethics*. 4th edition. New York, NY: Oxford University Press.
- Bersoff, D. N. (ed) (2007). *Ethical Conflicts in Psychology*. 3rd edition. Washington, DC: American Psychological Association.
- Boyer, P. (2002). *Religion Explained: The Human Instincts that Fashion Gods, Spirits and Ancestors*. New York, NY: Vintage.
- Canter, M. B., Bennett, B. E., Jones, S. E. & Nagy, T. F. (1994). *Ethics for Psychologists: A Commentary on the APA Ethics Code*. Washington, DC: American Psychological Association.
- Davidson, G., Garton, A. F. & Joyce, M. (2003). Survey of ethics education in Australian university schools and departments of psychology. *Australian Psychologist*, 38, 216-222.
- Dawkins, R. (2006). *The God Delusion*. New York, NY: Bantam Press.
- Dennett, D. C. (2004). *Freedom Evolves*. London: Penguin Books.
- Francis, R. (1999). *Ethics for Psychologists: A Handbook*. Leicester: BPS Books.
- Green, S. A. & Bloch, S. (eds) (2006). *An Anthology of Psychiatric Ethics*. Oxford: Oxford University Press
- Haas, L. J. & Malouf, J. L. (1995). *Keeping Up the Good Work: A Practitioner's Guide to Mental Health Ethics*. Sarasota, FL: Professional Resource Exchange.
- Handelsman, M. M. (1986). Problems with ethics training by "osmosis". *Professional Psychology: Research and Practice*, 17, 371-372.
- Hausser, M. D. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York, NY: HarperCollins.
- Koocher, G. P. & Keith-Spiegel, P. (1998). *Ethics in Psychology: Professional Standards and Cases*. New York, NY: Oxford University Press.
- Lindsay, G. & Colley, A. (1995). Ethical dilemmas of members of the British Psychological Society. *The Psychologist: Bulletin of the British Psychological Society*, 8, 448-453.
- McGuire, J. (1997). Ethical dilemmas in forensic clinical psychology. *Legal and Criminological Psychology*, 2, 177-192.
- McGuire, J. (2000). Can the criminal law ever be therapeutic? *Behavioral Sciences and the Law*, 18, 413-426.
- McGuire, J. (2004). Minimising harm in violence risk assessments: Practical solutions to ethical problems? *Health, Risk and Society*, 6, 327-345.
- Miner, M. H. (2005). Ethics education: further reasons why a grounding in ethical theory is essential. Comment on Davidson, Garton, and Joyce (2003). *Australian Psychologist*, 40, 54-56.
- Nagy, T. F. (2000). *Ethics in Plain English: An Illustrative Casebook for Psychologists*. Washington, DC: American Psychological Association.
- Nederlandt, P. (1996). *Comment l'étudiant en psychologie découvre-t-il la déontologie?* Paper presented at the First European Meeting on Psychology and Ethics, Calouste Gulbenkian Foundation, Lisbon, Portugal.
- Pope, K. S. & Vasquez, M. J. T. (1998). *Ethics in Psychotherapy and Counseling: A Practical Guide*. San Francisco, CA: Jossey-Bass.
- Pope, K. S. & Vetter, V. A. (1992). Ethical dilemmas encountered by members of the American Psychological Association. *American Psychologist*, 47, 397-411.
- Pryzwansky, W. B. & Wendt, R. N. (1999). *Professional and Ethical Issues in Psychology: Foundations of Practice*. New York and London: W. W. Norton.
- Sinclair, C., Poizner, S., Gilmour-Barrett, K. & Randall, D. (1987). The development of a code of ethics for Canadian psychologists. *Canadian Psychology*, 28, 1-8.
- Singer, P. (ed) (1993) *A Companion to Ethics*. Oxford: Blackwell.
- Smith, T. S., McGuire, J. M., Abbott, D. W. and Blau, B. I. (1991). Clinical ethical decision making: an investigation of the rationales used to justify doing less than one believes one should. *Professional Psychology: Research and Practice*, 22, 235-239.
- Strasburger, L. H., Gutheil, T. G. & Brodsky, A. (1997). On wearing two hats: role conflict in serving as both psychotherapist and expert witness. *American Journal of Psychiatry*, 154, 448-456.
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Tymchuk, A. J. (1986). Guidelines for ethical decision making. *Canadian Psychology*, 27, 36-43.

Ethical Dilemmas, Forensic Psychology and Therapeutic Jurisprudence

Ida Dickie

E-mail: IDickie@spalding.edu

Some have argued that irreconcilable conflict exists between therapeutic and forensic psychological roles (Greenberg & Shuman, 1997). This position contributes to the myth that effective psychological assessment and intervention can not occur within the legal system. However, it can be argued that the examination of the ethical dilemmas that result when the legal and psychological worlds intersect has been too narrow. The discussion should center on the following question; how can we as professional forensic psychologists function as psychologists who can contribute simultaneously to the advancement and well being of the offender, the community and society. This article explores how the Therapeutic Jurisprudence (TJ) model may allow for reconciliation between traditional clinical and forensic psychological roles. However, before discussing how the TJ Model can assist with such a reconciliation two basic assumptions have to be accepted.

First and foremost, the assumption that forensic psychologists and the law have a common goal; the prevention and management of criminal behaviour has to be embraced. Many of the ethical dilemmas discussed as existing in forensic settings for psychologists stem from the belief that psychology and the law do not have a common goal (McGuire, 1997). Furthermore, in order for forensic psychologists and the law to truly work together towards the shared goal of prevention and management of criminal behaviour, the legal system has to produce an environment that is conducive to therapeutic efforts. The legal theory of Therapeutic Jurisprudence may provide such an environment. Therapeutic Jurisprudence (TJ) is defined as the study of the legal system's role as a "therapeutic agent" (Wexler & Winick, 1991). This does not change the focus of TJ which is the law, but it allows for movement away from the traditional adversarial nature of the legal system and also provides a framework that allows the law to take into account human needs and emotional wellbeing of offenders (Schneider, 1999).

Slobogin (1995) defines TJ as "the use of social science to study the extent to which a legal rule or practice promotes the psychological and physical well being of the people it affects". Therefore, the second assumption that has to be accepted is that the law and the legal system can indeed function as a therapeutic agent of change. Critical reviews of the TJ model have suggested that the legal system is not capable of operating as a therapeutic agent or within a therapeutic framework because it is impossible to overlook the punitive ideology, power imbalance and violence embedded in the law (Arrigo, 2004). However, this article embraces the assumption that in spite of the legal system being predicated upon a punitive model, the law can be applied in a therapeutic way and in a fashion that minimizes the inevitable anti-therapeutic impacts of the law. Therapeutic is defined as promoting the advancement and well being of the offender, the community and society.

Principles of Ethically Sound Psychological Conduct (Meta-Code) have been articulated in the literature and include: 1) beneficence: the acceptance of responsibility to do good; 2) non-Maleficence: the acceptance of a need to do no harm; 3) autonomy: the respect for others freedom of thought and action; 4) justice: basing actions on fairness between individuals; 5) fidelity: trustworthiness to commitments and 6) respect for a person rights and dignity, competence, responsibility and integrity (Welfel & Kitchner, 1992; European Meta-code). It has been discussed that these ethical principles are difficult to apply within the legal system as it currently operates (McGuire, 1997). However, it may be that a broader interpretation of these principles needs to occur, with particular attention given to effective treatment principles within forensic settings, most notably the responsibility principle. (Andrews & Bonta, 2004).

For example, the aforementioned ethical principles may need to be conceptualized within a collectivistic model within legal settings in which the individual, the victim and the community are

considered inseparable as clients. Why does it have to be unethical to consider the needs of the victim as well as the offender when treating the offender? It is possible that offenders require a different therapeutic approach, one where the needs of the victim are considered. Similarly, a commitment to beneficence and non-maleficence within the forensic setting may require a collectivist approach because without considering the offender, the victim, and the community together then harm is more likely. Why does helping to offender manage his risk level to harm others have to be considered harmful? Could it not be re-stated that by not assisting the offender, sometimes with incarceration, to manage his risk level that he or she is being harmed. Ultimately, the question that needs to be asked is: "does ethical principles look different within legal settings when practically effectively as psychologists". This question can only be asked if the legal system is committed to upholding the ethical psychological principles.

The TJ model calls for an increased psychological sensitivity in the attorney/client relationship and awareness of the basic principles and techniques of psychology (Winick & Wexler, 2006). This means that lawyers will utilize motivational interviewing skills and considering all aspects of the offenders' lives when devising a defence or prosecution strategies such that the rights of the offenders and public safety can be balanced. For more information on how clinical concepts are being introduced into legal training refer to Winick & Wexler, 2006. The underlying principles of TJ are as follows. 1) The way law is implemented can increase or decrease/have a neutral impact on well being. 2) The law should capitalize on the entry into the legal system to promote a pro-social lifestyle. 3) The law can be a multi-disciplinary endeavour in which psychology and law co-operate to enhance well-being. 4) The law balances community protection/justice principles against individual needs/therapeutic principles (Birgden 2002; Winick & Wexler, 2006). These underlying principles promote a collectivistic approach to ethical dilemmas within the legal system.

The Ethical Code suggested by TJ Model suggests that 1) the offender must be treated with procedural fairness; 2) amount and type of

treatment is governed equally by the seriousness of offence and the need for treatment 3) infringements on an offenders legal rights minimized- least restrictive principle and 4) non-maleficence (Glasser, 2003). The TJ model allows for a more therapeutic approach to managing criminal behaviour in which all invested parties, psychologist, lawyers, and correctional officers can balance the rights of the offenders and public safety. TJ model promotes an environment where collaboration between the law, psychiatry, community mental health and psychology can assist offenders in asking "How can I live my life differently". The TJ ethical code can be subsumed under the larger ethical code in which psychologists practice under and make it easier to deal with some of the ethical dilemmas that psychologists face while working within the legal setting. The TJ model offers psychologists a way to respect offenders self determination and autonomy even when it has to be reduced at times, and suggest that such reductions can have important therapeutic value.

In conclusion, the TJ model allows Forensic psychologists to practice within a legal system that respects an interdisciplinary and psychologically oriented approach to management of criminal behaviour. This paradigm concerns itself with client needs and emotional well being as well as offenders rights which ultimately contributes to adherence to ethical standards and resolution of forensic and clinical roles.

References

- Andrews, D. & Bonta, J. (2006). *The psychology of criminal conduct (4th ed.)*. Albany, New York: Matthew Bender & Company, Inc.
- Arrigo, B. (2004) The ethics of therapeutic jurisprudence: A critical and theoretical enquiry of law, psychology, and crime. *Psychiatry, Psychology, and Law*, 11 (1), 23-43
- Birgden, A. (2002) Therapeutic jurisprudence and the "good lives.": A rehabilitative framework for corrections *Australian Psychologist*, 37,180-186.
- Glaser, B. (2003). Therapeutic jurisprudence: an ethical paradigm for therapists in sex offender treatment programs. *Western Criminology Review*, 4 (2), 143-154.
- Greenberg, S.A. and Shuman, D.W. (1997). Irreconcilable Conflict Between Therapeutic and Forensic Roles. *Professional Psychology: Research and Practice*, 28, 50-157.

- McGuire, J. (1997). Ethical dilemmas in forensic clinical psychology. *Legal and Criminological*, (2), 177-192.
- Schneider, A. (1999).. The Intersection of Therapeutic Jurisprudence, Preventive Law, and Alternative Dispute Resolution. *Psychology, Public Policy and Law*, 1084-1102
- Slobogin, C. (1995). Therapeutic Jurisprudence: Five dilemmas to ponder. *Psychology, Public Policy and Law*, 193-196.
- Welfel, E. R., & Kitchener, K. S. (1992). Introduction to the special section: ethics education – an agenda for the '90s. *Professional Psychology: Research and Practice*, 23, 179-181.
- Wexler, D. & Winick (1991). Therapeutic Jurisprudence as a New Approach to Mental Health Law and Policy Analysis and Research. *University of Miami Law Review*, 975.
- Winick, B. & Wexler, D. (2006). The use of therapeutic jurisprudence in law school clinical education: Transforming the criminal law clinic. *Therapeutic Jurisprudence and Clinical Legal Education and Skills Training*.

Training Needs for Graduate Students in Forensic Psychology

Brandi Reynolds, Amanda Reinhardt, & Ida Dickie
E-mail: IDickie@spalding.edu

APA guidelines and Forensic Psychology Speciality Guidelines outline principles that mandate psychologists to practice only in their area of competence which is established by their training and educational experiences. Therefore, forensic psychologists need specific training to practice in an ethical and competent manner.

Duties of a Forensic Psychologist

Forensic and correctional psychologists may find themselves having many roles. They may do individual therapy, group therapy, crisis intervention, provide employee assistance, participate in personnel selection, or serve as consultants for institutions, lawyers, courts to name a few (Brandt, 2005). They may also work in a variety of settings such as hospitals, community mental health centers, prisons, or jails.

In order to perform so many different duties, psychologists in the field of forensic and correctional psychology must have adequate training. The first general principle set forth by the APA (1992) guidelines suggests that psychologists maintain awareness of the boundaries of their own competence and limit professional activities to those in which their competence is established by education, training, and experience. In addition, Specialty Guidelines for Forensic Psychologists state that psychologists have the responsibility to limit their forensic work to those areas in which they possess specific knowledge, skill, experience, and education (CEGFP, 1991).

Therefore, it is essential that psychologists trained in clinical psychology obtain training in forensic psychology before practicing in the field. Likewise, it is mandatory that forensic psychologists obtain training in the field of clinical psychology. These two fields of training must go hand in hand for a psychologist to be effective at the variety of duties they are called upon to perform.

Education on Law, Psychology, and Ethics

According to Melton, Huss, and Tompkins (1999), "all that is needed to integrate psychology and law is to cure a problem in communication (702)," however, this is unrealistic. Therefore, basic training in law, including knowledge of the adversary system, is mandatory for ethical practice for forensic psychologists in training.

Psychologists in training need to be taught how the legal and psychological systems differ in mind-set (Melton, et al., 1999). For example, psychologists desire to learn "the truth" about human behavior, whereas lawyers are concerned with human behavior but are seeking justice (Ogloff, 1990). By learning about the differences, psychologists can best communicate with lawyers. It is suggested that attorneys be included in this training to teach students the appropriate relevant laws and definitive differences between the fields from the legal system's perspective, as well as learn to prevent oneself from invading client's rights in the adversary system (Melton, et al., 1999).

This sufficient knowledge in law is also necessary to avoid ethical violations. According to APA guidelines and in congruence with CPA guidelines, "it is mandatory that psychologists "avoid any action that will violate or diminish the legal or civil rights of clients or others who will be affected by their actions" (American Psychological Association [APA], 1981, p. 634; Canadian Psychological Association [CPA], 1986, as cited in Ogloff, 1990).

Therapeutic Jurisprudence

It is suggested that the Therapeutic Jurisprudence model would be helpful in reducing the amount of ethical violations because both the psychological and legal systems keep the best interest of the client as their primary goal (Arrigo, 2004). Further, therapeutic jurisprudence is an extremely important model and provides a therapeutic framework for new therapists. It is mandatory for new therapists to learn how to balance treatment

goals and the punitive correctional setting in order to assist patients to make treatment gains and also to maintain safety for the community.

A student pursuing psychology needs to be taught how to do no harm to their client, yet still protect others from harm (Stone, 1996). Because forensic psychologists have responsibility to the state, the fine line between an offender's autonomy and society's protection needs to be distinguished before working with the population.

In order to learn this distinction, students need to be taught who "the client" is in specific cases (Krock & Zibbell, 1998). The client may be the individual or even a family if they come to the clinician at their own will. The client may be the court if the individual or family is referred by the court. An attorney may be the client if the attorney refers the client for forensic assessment. Each of these clients require different methodology and legal considerations. By teaching students to recognize the client, the student will be able to avoid ethical dilemmas by respecting autonomy through consent and by learning to notify the individuals of their rights, what will be happening in therapy and/or testing, and why they are referred (Stone, 1996; Krock & Zibbell, 1998).

One reason that students favor the therapeutic jurisprudence framework is because "the client" will always include the individual. The individual's human needs and emotional well-being are always considered by lawyers and attorneys, as well as the psychologist. Therefore, this model can aid forensic students by allowing them to create a therapeutic relationship, while following APA and FP guidelines.

Forensic students who are educated about the therapeutic jurisprudence model can bring their knowledge to lawyers and attorneys within their jurisdiction and make sure that they are following best practice when working with their clients (Winick & Wexler, 2005). This allows for a collaborative effort between the legal and psychological systems.

Clinical Training

According to Otto, Heilbrun, & Grisso (1990) students need exposure to multiple forensic and

correctional settings through practicum experiences. Training programs should strive to make community connections with hospitals, courts, community mental health centers, prisons, and jails that offer forensic and/or correctional training to students. Additionally, programs should also strive to involve forensic professionals in applied settings and those from other disciplines in research in the forensic/correctional field in order to ensure that research is clinically and legally relevant (Otto et al., 1990).

These practicum experiences will be more beneficial when they are coordinated with student coursework. If students are able to tie in what they are learning in the classroom to what they are doing at their practicum placement they will gain a richer experience. Applying learning in the classroom will also help students to understand and work through challenging forensic and correctional issues.

Forensic Assessment Training

APA guidelines and Guidelines for Forensic Psychologists outline that professionals must practice within the boundaries of their own competence. Included in this ethical duty is forensic assessment. Students must ensure that their programs provide adequate training in the area of forensic assessment so that they are competent to provide ethical assessment services in this area. According to Butcher and Pope (1993) students must be trained to utilize the appropriate standardized assessment instruments in order to ensure they are administered appropriately. Students must also be exposed to a variety of assessment instruments so that they feel comfortable using different assessment tools when needed (Butcher and Pope, 1993). In addition, it is important that students are able to choose the right measures for each client in order to answer the referral question.

It is essential that training programs educate students on how to write assessment reports effectively in order communicate the referral question for the client. Forensic assessments are powerful tools in the field and students must feel comfortable in this area. Finally, communicating the results of the assessment can be a complicated task. Students must be trained to effectively

communicate assessment results to many different groups including the individual client, the court, lawyers, parole boards, etc. Without this training students run the risk of incorrectly communicating results of the assessment, which may affect the individual client and the community at large.

Empirically Supported Treatment Strategies

How do we measure what is effective treatment in the field of forensic psychology? Is it increased insight on the part of the offender or is it a reduction in the recidivism rate? (Stone, 1997). Students must continually assess whether or not the treatment that they are providing is effective with the population with which they are working. This can be a difficult task; however, it is important that students are encouraged to remain familiar with relevant empirical research in order to practice ethically in the field. This also requires that students examine themselves in the treatment room.

Training for students must focus on helping them to assess risk factors and criminogenic needs of clients with which they work. According to Andrews and Bonta (2006), research has indicated that to provide clinically relevant and psychologically informed human service, clinicians must focus service delivery on the high risk population and focus on the dynamic or criminogenic needs of the client. Once students learn about the Risk-Need principles they must be able to put them into practice in their clinical training.

The Responsivity principle requires that students utilize effective treatment strategies that will influence the clients and cause behavior changes (Andrews and Bonta, 2006). According to Andrews and Bonta (2006) behavioral, social learning, and/or cognitive behavior strategies have been proven to be most effective. Training in these treatment methods, both in the classroom and in clinical internships, is important for students to learn. The responsivity principle also requires that the clinician examine themselves in terms of effective treatment delivery. It is imperative that students learn to examine themselves and the treatment that they are providing in order to ensure that they are

providing the most effective treatment for their population of interest.

According to Stone (1997) students must also “be aware of their limitations and lack of experience and be particularly aware of counter-transference (257).” It can often be hard for students and psychologist alike to admit that they are having a strong reaction to their clients. If this issue is ignored treatment may be greatly impacted. Students must learn to examine themselves and their biases in regards to their clients. Often clients’ situations in the forensic/correctional field may trigger thoughts and emotions within us that are hard to admit and discuss. However, it is the acknowledgement and the discussion of these issues that helps the clinician work through these thoughts and emotions in order to provide effective treatment.

Using Supervision

Clinical supervision is an important factor in the training of forensic/correctional psychologists. It should be considered an integral part of any training program. According to Mothersole (2000) “Good supervision by its very nature involves a challenging of attitudes and actions, and may well lead to a change in old established practices. A good supervisor will help the student question themselves and what they are doing in treatment in a supportive manner. The supervision relationship is a chance for students to learn about themselves in terms of their clinical work and their interactions with other professionals in the field. Therefore, this relationship should be used in a way that promotes professional growth for the student (47).”

If a student experiences poor supervision they must be trained on how to address this problem. If a student feels they are not getting what they need, then they must talk with their supervisor to communicate their expectations and offer suggestions to work out the problem. It is also helpful for students to have a supportive advisor or faculty member to consult with that is separate from their practicum placement.

Education on Credentialing

Credentialing allows courts, lawyers, and other private parties to identify mental health

professionals that have a certain level of training and experience in forensic practice (Otto et al, 1990). This makes credential programs a form of forensic training because it involves continuing education that keeps mental health professionals up to date about current topics in the field. According to Otto, Heilbrun, and Grisso (1990) credentialing currently occurs in two forms at both the state and national level.

Further research must be done regarding the credentialing by state boards in order to organize and monitor professionals in the field that are practicing forensic psychology. National credentialing is more regulated and monitored but the value of this certification has yet to be empirically tested. Students must be encouraged to research specialty training in the area of forensic psychology and consider credentialing by the state or national boards in order to remain abreast of the important advances in the field of forensic psychology.

In conclusion, in order to practice ethically and competently in the forensic setting, specialized training is required to assist psychologists with the ethical dilemmas unique to the setting. Students interested in entering the field of forensic psychology must be given opportunities in their academic and clinical work to be exposed to these unique aspects of the field and to obtain the specialty training required in order to practice as ethical and competent professionals.

References

- Andrews, D. & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Albany, New York: Matthew Bender & Company, Inc.
- Arrigo, B. (2004) The ethics of therapeutic jurisprudence: A critical and theoretical enquiry of law, psychology, and crime. *Psychiatry, Psychology, and Law*, 11 (1), 23-43.
- Birgden, A. (2002) Therapeutic jurisprudence and the "good lives." *Australian Psychologist*, 37, 180-186.
- Brandt, S. (2005). Correctional psychology: Looking beyond the bars. In R. Morgan, T.L. Kuher, & C.J. Habben (Eds.). *Life after graduate school in psychology: Insider's advice from new psychologists* (pp.151-165). New York: Psychology Press.
- Butcher, J. & Pope, K. (1993). Seven issues in conducting forensic assessments: Ethical responsibilities in light of new standards and new tests. *Ethics and Behavior*, 3 (3&4), 267-288.
- Glaser, B. (2003). Therapeutic jurisprudence: an ethical paradigm for therapists in sex offender treatment programs. *Western Criminology Review*, 4 (2), 143-154.
- McMurrin, M. (2004) Motivating offenders to change in therapy: An organizing framework. *Legal and Criminological Psychiatry*, 9, 295-311.
- Melton, G. B., Huss, M. T., & Tomkins, A. J. (1999). Training in forensic psychology and the law. In A. K. Hess & I. B. Weiner (Eds.), *Handbook of forensic psychology* (2nd ed., pp. 700-720). New York: Wiley.
- Mothersole, G. (2000). Clinical supervision and forensic work. *The Journal of SexualAggression*, 3 (1), 45-58.
- Ogloff, J. (1990). Law and psychology in Canada: The need for training and research. *Canadian Psychology*, 31 (1), 61-73.
- Ogloff, J. (2002). Two steps forward and one step backward. In J.R. Ogloff (Ed.). *Taking psychology and law into the twenty-first century* (pp. 1-31). New York: Plenum Publishers.
- Otto, R.; Heilbrun, K.; & Grisso, T. (1990). Training and credentialing in forensic psychology. *Behavioral Sciences and the Law*, 8, 217-231.
- Stone, J. (1996). Medico-legal ethics in forensic psychotherapy. In E.V. Welldon & C.V. Velson (Eds.). *A practical guide to forensic psychotherapy* (pp.253-258). Bristol, PA: J. Kingsley Publishers.
- Winick, B. & Wexler, D. (2005). The use of therapeutic jurisprudence in law school clinical education: Transforming the criminal law clinic. *Therapeutic Jurisprudence and Clinical Legal Education and Skills Training*.

A Distance Learning Course on Mental Health Issues in Jails and Prisons

Michael L. Perlin
New York Law School
E-mail: mperlin@nyls.edu

Through the technology of the Internet-based education, the author (a professor at New York Law School) (NYLS) has created a program of on-line mental disability law courses for attorneys, activists, advocates, important stakeholder groups (consisting of consumers and users of psychiatric services, sometimes referred to as “survivor groups”), mental health professionals and governmental officials in an effort to both teach participants the bases of mental disability law and to encourage and support the creation and expansion of grass-roots advocacy movements that may optimally lead to lasting, progressive change in this area. This is especially timely in light of recent research demonstrating how the Internet has already become an important provider of advocacy services and advocacy information to many persons with disabilities (*see* Blanck et al, 2003; Ritchie & Blanck, 2003), and how inaccessible most current websites are to many persons with disabilities.

First, this paper will briefly discuss the use of distance learning in a law school environment, and will next consider the special implications of distance learning for persons with disabilities. Then, it will explain the structure and rationale of this program, and it will discuss the most recent course added to the array offered at NYLS: a course entitled *Mental Health Issues in Jails and Prisons*.

I. Distance Learning in Law Schools

Distance learning is generally defined as “communication which connects instructors and students who are separated by geography and, often, by time,” or as “the electronic connection of multiple classrooms” (Leskovac, 1998). Distance learning courses enable students to share different perspectives, and provide a new environment for teaching law students to collaborate with other types of professionals (Berg, 2003), a

characteristic “increasingly essential to the effective practice of law” (Berg, 2003, p. 34). Distance learning – the use of computers, telecommunications, and digital networking to permit learning outside the boundaries of the classroom – “holds the potential to expand the availability of cross-listed courses by reducing these barriers ... [and] can provide professors of cross-listed courses with pedagogical tools for enhancing interdisciplinary communication and collaboration, and circumventing some of the problems inherent in teaching students from different disciplines” (Berg, 2003, p. 35).

Self-evidently, distance learning has great implications for international legal education as well as for domestic legal education. A report in the Fletcher Forum of World Affairs concluded: “[T]here is no doubt that ICTs [Information and Communication Technologies], if properly adopted and implemented, can bring economic and cultural opportunities to developing countries. Education facilities may be greatly improved through distance learning and Internet access” (Cukor & McKnight, 2001, p. 47).

II. The Special Implications of Distance Learning Education for Persons with Disabilities

One of the specific challenges in creating a distance learning pedagogy in mental disability law is the need to provide a program that can also be meaningfully accessed by persons with disabilities. By way of example, a recent study by the UK-based Disability Rights Commission showed that 81% of British websites are inaccessible to persons with disabilities.¹ Scholars have begun to explore how the Internet can

¹See <http://www.drc-gb.org/newsroom/newsdetails.asp?id=805§ion=4>. See also, Axel Schmetzke, *Online Distance Education - "Anytime, Anywhere" but Not for Everyone*, manuscript at 12, available online at <http://www.rit.edu/~easi/itd/itdv07n2/axel.htm>

provide individuals with disabilities the tools to enable them to live independently and “to gain greater independence and social integration”(Rich, 2002), and have thus begun to call for a coordinated program of study to examine the extent to which Internet sites are accessible to persons with disabilities (Blanck & Sandler, 2000). A study of 200 websites affiliated with Centers for Independent Living concluded:

(1) Accessible technology for persons with disabilities has the potential to enhance independence in life. Its future development holds promise for a wide range of persons with disability...

(2) The commitment to digital equality as a civil right must be founded in policy that incorporates accessibility and universal design in public and private programs providing technological access to all.

(Ritchie & Blanck, 2003, p. 24)

III. Internet Courses

NYLS created the first Internet-based mental disability law courses in an attempt to disseminate the core universal principles of mental disability law to the full range of activists, advocates, professionals and stakeholders described above. The courses were developed at the suggestion and urging of Dean Richard A. Matasar, an early visionary in computer-based legal education. (Matasar, 1998; Matasar, 2002-03; Matasar & Shiels, 1995). There are currently six courses being offered: (1) *Survey of Mental Disability Law*, (2) *The Americans with Disabilities Act: Law, Policy and Practice*, (3) *International Human Rights Law and Mental Disability Law*, (4) *Lawyering Skills for the Representation of Persons with Mental Disabilities*, (5) *Mental Health Issues in Jails and Prisons*, and (6) *Sex Offenders*. Three more courses are scheduled to be offered in the 2008-09 school year: *Forensic Ethics and Forensic Evidence*; *Competency and the Civil Law*, and *Mental Illness, Dangerousness, the Police Power and Risk Assessment*. Following that, NYLS plans on launching an online Masters degree program in mental disability law studies (open to lawyers and mental health professionals) starting in the Fall 2008 term (Perlin, 2006; Perlin, 2007).

IV. The Jails and Prisons Course

In the seven years that we have been offering courses in this modality, it has become clear to us that one of the “missing pieces” has been a course that is focused directly and solely on issues that involve the provision of mental health services to persons in jails and prisons. It has become a well-worn cliché that the largest providers of mental health services in the US are the county jails in Chicago, New York City and Los Angeles (Lurigio & Swartz, 2006)

Even if this observation is slightly apocryphal, we know the figures are appalling: Nearly two million new jail admissions are of people with mental illnesses—35,000 individuals a week. At the end of 2000, nearly one million individuals with mental illnesses were in the criminal justice system, and, at the end of 1998, over 280,000 of such persons were in prisons and jails (Perlin & Dlugacz, 2007). One in six U.S. prisoners is mentally ill, and many of that cohort suffer from serious illnesses such as schizophrenia, bipolar disorder, and major depression. There are three times as many men and women with mental illness in U.S. prisons and jails as in mental health hospitals (Lurigio & Swartz, 2006). Moreover, the rate of mental illness in the prison population is three times higher than in the general population (Perlin & Dlugacz, 2007).

Prisons are dangerous and damaging places for mentally ill people. Other prisoners victimize and exploit them. Prison staff often punish mentally ill offenders for symptoms of their illness – such as being noisy or refusing orders, or even self-mutilation and attempted suicide. Mentally ill prisoners are more likely than others to end up housed in especially harsh conditions, such as isolation, that can push them over the edge into acute psychosis (Smith, 2006). Woefully deficient mental health services in some prisons leave prisoners under treated – or not treated at all. Across the country, prisoners cannot get appropriate care because of a shortage of qualified staff, lack of facilities, and prison rules that interfere with treatment (Lopez, 1999, p. 761; Vosberg, 1988, pp. 766-67). And, there are fewer and fewer lawyers available to provide legal

services to this population.²

The concomitance of all these factors – the unprecedented explosion of the number of persons with mental illness and mental retardation in jails and prisons, the statutory limitations on law suits by prisoners, the diminution of legal services available to prisoners, the increase in litigation on behalf of persons institutionalized because of mental illness in civil and forensic settings, the blurring of the borderlines between criminal and civil mental disability law – have led us to conclude that (1) the time was right for a course dedicated solely to the topic of mental health issues in jails and prisons. We thus added this course to our online program in an effort to make accessible to potential students – in the US and around the world – the basic elements of this area of the law.

In addition to this course, there is also a casebook (titled, *Mental Health Issues in Jails and Prisons: Cases and Materials* (Perlin & Dlugacz, 2007)). Some of the major points of focus of the casebook include “Persons with mental illness and criminal justice system: Why should we care?”, “Conditions in jails and prisons: deliberate indifference to a serious medical need”, “Statutory issues”, “International law issues”, and “Legal counseling of persons with mental disabilities in jails and prisons: on sanism and pretextuality”.

Conclusion

The Internet has the capacity to transform and invigorate legal education through the use of distance learning methodologies. New York Law School, which created the first Internet-based mental disability law course domestically in 2000, has now expanded that course so as to offer six different courses, with an online Masters soon to come. Also, faculty are actively seeking to expand the course to law schools and universities in the US and abroad (see, for a discussion of past international programs, Perlin, 2007). Our hopes and expectations are that, through this course as

part of this program, we can focus in on the law and policy issues that are at the heart of the underlying problems.

References

- Berg, P. (2003). Using distance learning to enhance cross-listed interdisciplinary law school courses. *Rutgers University Computer and Technology Law Journal*, 29, 33-52.
- Blanck, P. & Sandler, L. (2000). ADA title III and the internet: Technology and civil rights. *Mental and Physical Disability Law Reports*, 24, 855-858.
- Blanck, P. et al. (2003). Technology for independence: A community-based resource center. *Behavioral Science and Law* 21(1), 51-62.
- Blanck, P. & Ritchie, H. (2003). The promise of the internet for disability: A study of on-line services and web-site accessibility at centers for independent living. *Behavioral Science and Law*, 21(1), 5-26.
- Cukor, P & McKnight, L. (2001). Knowledge, networks, the internet, and development. *Fletcher Forum of World Affairs*, 25, 43-57.
- Leskovic, H. (1998). Distance learning in legal education: Implications of frame relay videoconferencing. *Albany Law Journal of Science and Technology*, 8, 305-334.
- Lopez v. Lemaster, 172 F.3d 756 (10th Cir. 1999).
- Lurigio, A. & Swartz, J. (2006). Mental illness in correctional populations: The use of standardized screening tools for further evaluation or treatment. *Federal Probation*, 70, 29.
- Matasar, R. A. & Shiels, R. (1995). Electronic law students: Repercussions on legal education. *Valparaiso Law Review* 29, 909-933.
- Matasar, R. A. (1998). Private publics, public privates: An essay on convergence in higher education. *University of Florida Journal of Law & Public Policy*, 10, 5-26.
- Matasar, R. A. (2002-2003). Legal education skills and values education: Debate about the continuum continues. *New York Law School Law Review*, 46, 395-428.
- Perlin, M. (2006). “Ain’t no goin’ back”: Teaching mental disability law courses on line. *New York Law School Law Review*, 51, 991-1001.
- Perlin, M. (2007). An internet-based mental disability law program: Implications for social change in nations with developing economies. *Fordham International Law Journal*, 40, 435-461.
- Perlin, M. & Dlugacz. (2007). *Mental health issues in jails and prisons: Cases and materials*. Carolina Academic Press: Durham, NC.

²On the impact on prisoners in New York of the 2006 decision of Governor Pataki to veto \$12 million in financing for legal aid programs, see Abby Goodnough, *The Governor's Budget Vetoes Imperil Legal Aid Programs: Lawyers for Poor Warn of Unequal Justice*, accessible at <http://www.pili.org/2005r/content/view/152/26> (2006).

Rich, R. et al. (2002). Critical legal and policy issues for people with disabilities. *DePaul Journal of Health Care Law*, 6, 1-53.

Smith, P.S. (2006). The effects of solitary confinement on prison inmates: A brief history and review of the literature. *Crime and Justice*, 34, 441-528.

Vosberg v. Solem, 845 F.2d 763 (8th Cir. 1988).

Prison, Stress and the Aging Process

Donovick, P. J.

State University of New York at Binghamton
E-mail: donovick@binghamton.edu

Holtzer, R.

Yeshiva University

Portocarrero, J. S.

State University of New York at Binghamton

Yee, R. S.

State University of New York at Binghamton

With approximately seven percent of adults incarcerated, the United States has more individuals in jail and prison per capita than any country in the world. Associated with the growth in American prisons has been an increase in prisoners over 55 years of age. This “graying” of prisoners can be attributed to changes in policy such as three strike laws, longer sentences for some crimes (e.g., drugs), increased use of mandatory minimum sentences, and an increase in the age of first-time offenders (Mara 2002). In the relatively brief period of time of 1995 to 2003 this older population has nearly doubled (Harrison & Beck 2004). Paralleling the *civilian world*, older prisoners need more medical and social services than younger prisoners (Jones, Connelly, & Wagner 2001) resulting in increased medical costs. The present article seeks to address some of the causes and ramifications of this shift in the demographics of the prison population.

Although the degree of stress undoubtedly varies between facilities (e.g., maximum vs. minimum security prisons) it is an understatement to say that prison life is stressful. The crowding, lack of privacy, noise and interpersonal tension of prison life make it inherently stressful as reflected in such symptoms as fatigue, sleeplessness, headaches, and weight fluctuations (c.f., Lindquist & Lindquist, 1997).

Prison housing provides ideal conditions for the spread of communicable diseases (c.f., Lindquist & Lindquist, 1999). Individuals are subject to interpersonal violence as represented, for instance, in forced sexual contact (e.g., Beck & Harrison 2006), and prisoner on prisoner homicide (Cox et al., 1984; Tartaro, 2002). Controllability and predictability of aversive stimuli determine, in part, the nature and magnitude of the stress response (e.g., Lindquist & Lindquist, 1997). Prisoners have minimal control over their lives. While some aspects of their daily routine

are predictable, such as meal time and work placement, frightening interpersonal interactions are not (cf., Lindquist & Lindquist, 1997). Ongoing social support is typically lacking in prisons. Conflict, fear, and despair are common factors in prison suicides (Way, Miraglia, Sawyer, Beer & Eddy, 2005). Separation from family and friends, which is often compounded by the great distance between the prison and home, exacerbates loneliness, isolation, and the ability to deal with family problems.

Due to the dynamics of prison life, the aging prisoner may be faced with different, and sometimes more intense, stressors than the younger prisoner. Given that physical prowess and perceived *dangerousness* are valued attributes in prison, and are often a critical component of keeping an individual “safe,” the natural aging process is a risk factor for predation that the individual prisoner may face. The gradual change in prison status may be inevitable but, there are significant individual differences in susceptibility to the negative effects of stress (cf., Coriell & Adler, 2001, Donovanick & Burright, 1984, Sapolsky, 2001, Turner, Hartman, & Bishop, 2007).

Exacerbating stress, state and federal prisons are functioning at “*over-capacity*” by as much as 15-20% (Harrison & Beck, 2004). Overcrowding in prisons exacerbates stressful conditions, affecting the behavior of inmates and may play a role in maladaptive *self-injurious* acts (Jeglic, Vanderhoff, & Donovanick, 2005). Cox and co-workers (Cox et al., 1984) conducted an in-depth archival analysis of the impact of crowding on more than 175,000 inmates. They found that increased crowding in prisons was associated with an increase in the number of deaths, suicides, disciplinary infractions, psychiatric crises and inmate on inmate assaults. As correctional facilities are forced to function with

overcrowding, prisoners have less access to physical, medical, and mental health resources.

Socio-economic-status is directly related to mortality and morbidity associated with a number of chronic diseases (Coriell & Adler, 2001). The prison is populated most heavily with individuals from disadvantaged communities (c.f., Turner, et al., 2007) where marginal employment is common. Many prisoners have limited education and minimal history of legal employment. Incarcerated individuals are much more likely to have a history of physical and sexual abuse than is found in the non-prisoner population (c.f., Andover, Crossman, Vanderhoff & Donovan, in preparation; Jeglic, Vanderhoff, & Donovan, 2005). Such histories contribute to the vulnerability of the individual (c.f., Donovan & Burright, 1984). Many come from communities where health insurance is rare and often inadequate. It is of no surprise that the frequency of many chronic disorders is higher in prisoners than it is in the general population (Lindquist & Lindquist, 1999). Limited resources in prisons, including the high case loads of mental health workers may produce frustration and competition among prisoners and may lead to disruptive behavior (Johnston, 1991). As might be expected, older inmates rate their health status more poorly, and account for higher utilization and costs of health care, than younger inmates (Lindquist & Lindquist, 1999). A life history of less than optimal healthcare, risky behaviors and a highly stressful environment are associated with an increase in inmate health problems and increases the probability of developing dementia.

Dementia, a global and progressive loss of cognitive abilities, typically includes deficits in the ability to learn new material. Cognitive changes are often associated with changes in personality (Lishman, 1988). There are a multitude of causes for dementia, with the most common being Alzheimer's disease, followed by arteriosclerotic neurovascular disorders, and several frontal-temporal dementias including Lewy Body and Pick's disease (Lishman, 1988). Disorders may be distinguished, in part, by the age of onset, and the rate of decline. A distinction is sometimes made between the cortical dementias such as Alzheimer's disease and the several frontal-temporal dementias, and subcortical

dementias such as Huntington's disease and AIDS dementia. In cortical dementias, symptoms typically include deficits in memory and executive functioning (Holtzer, Burright, & Donovan, 2004). Initially in subcortical dementia impairments in motor functioning are seen; later in the course of the disorder, cognitive impairments are observed. Prevalence studies in the general population suggest a steady increase in dementia from about 2% at age 65 to around 30% by age 90 (Riggs 2001).

It is reasonable to assume that the stressful life associated with incarceration increases the probability of dementia (Porter, Herman & Lansfield, 2001), particularly given that many prisoners have a history of perinatal and neuropsychological problems (c.f., Turner et al., 2007). A basic component of the stress system is glucocorticoid activity. Glucocorticoid receptors are plentiful in the hippocampus, a brain structure critical for laying down of new memories, including encoding of the environmental conditions surrounding an event (cf. Brown & Kulik, 1977). However, long-term exposure of the hippocampus to glucocorticoids impairs cellular functioning and thereby interferes with the laying down of new memories (Sapolsky, 2003). The impact is compounded as impaired hippocampal functioning subsequently reduces negative feedback on the adrenal-pituitary axis. The result is enhanced responsiveness to stressors. To compound the problem, a decline in hippocampal functioning is typically associated with aging. Stress compounds aging changes in the hippocampus increasing the probability of dementia (Porter et al., 2001).

Executive control processes decline at an accelerating rate in old age (Holtzer et al., 2004, 2005). This decline is related to functional impairment in activities of daily living (ADL) including the most basic self-care activities such as bathing and dressing. Behavioural disregulation including but not limited to judgment and impulse control is a frequent consequence of aging (Stuss et al., 2000; 2002). This age-related decline in functioning is particularly relevant to prison inmates whose executive processes, on average, may be compromised compared to the general population (Williams, Wagner, Hallquist, & Donovan, in preparation). Such a decline may

translate into increased behavioral problems in older inmates. Emotional and behavioral disturbances are common in dementia (e.g., Holtzer et al., 2003, 2005) and likely cause disciplinary problems for the individual living in the *rule-bound* prison.

Prisons have their own social ranking system and the health and age of an incarcerated person is related to the individual's rank. Cognitive changes associated with dementia will put the individual at odds with the rules and regulations that govern prisons and how a prisoner is viewed by both other prisoners and correctional officers. For instance, prisoners are supposed to be in designated places at specific times, but persons with dementia, tend to wander (c.f., Edgerly & Donovan, 1998; Holtzer, et al., 2003). Most incarcerated individuals are assigned to various *programs*, that are designed to help the prisoner re-integrate into the community when they are released. The older prisoner, including individuals with degenerative disorders such as dementia may become less able to carry out the required tasks and thus become a risk to him/herself and/or others.

In a community sample, older African-Americans showed a greater cognitive decline over a three-year period than did community dwelling older Caucasians (Sachs-Ericsson & Blazer 2005) and were also over-represented in the dementia population (Husaini, Sherkat, Moonis, Levine, Holzer, & Cain, 2003). In the non-incarcerated world, compared to Caucasians, African-Americans with dementia have better community social support. However there is no indication that the prison system, where African-Americans are over represented, allows for the development of social networks that would facilitate the care for an individual as he or she loses abilities due to dementia.

Conclusion

The design of American Prisons has been driven by the need to maintain secure and safe facilities. However, increasingly we ask the Department of Corrections to provide services for a growing heterogeneous population that includes the mentally ill, drug addicted, the chronically ill, and

now the elderly - many of whom will become demented.

There has been a steady increase in the number of individuals who are incarcerated and the typical prisoner is serving longer prison terms than was true in the past. This fact is due to mandatory sentencing laws which decrease the ability of courts to take into account mitigating circumstances, thus resulting in longer sentences. Further, *three-strike* laws have increased the number of individuals serving sentences of life without parole. As a consequence of such policy changes, the average prisoner is older than in previous decades.

Of course, prisoners are prone to the normal afflictions of aging, including dementia. As discussed above, incarcerated persons are more prone to such disorders than individuals living on the outside. With an increasingly older population, prisons will need to cope with inmates who are visually and auditorially impaired and motorically limited. Many will have cognitive limitations above and beyond what they had previously (c.f., Williams et al., in preparation). Stress of prison life magnifies the ravages of age increasing the morbidity of disease, including illnesses of dementia.

For correctional facilities to adequately provide required services for the increasingly diverse populations that they house, changes in the structure of prisons and how they provide services are needed. With longer sentences more prisoners will die in prison, necessitating the need for hospice care (Craig & Craig, 1999). Hospice care is challenging for the prisons system for financial and logistic reasons, and is absent in most prison systems (Craig & Craig, 1999). Many prisoners who desperately need such care are left to "fend for themselves."

Changes in the characteristics of prisoners must in the long run change the structure of prisons themselves and the services that they need to perform. Older, *limited* prisoners will need to be protected from predation from younger, more able bodied, prisoners. Programs for the aging, and often ill, prisoner will need to be implemented. Services to help individuals with their ADLs will need to be designed. And the physical structure of, at least some portions of the prison, may need

to be redesigned. Given the rapidly increasing number of older prisons, steps to implement these changes should begin now.

References

- Andover, M.S., Crossman, D.M., Vanderhoff, H.A., & Donovan, P.J. (In preparation) Neuropsychological functioning in self-mutilating and suicidal prisoners
- Brown, R. & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5, 73-99.
- Coriell, M. & Adler, N.E. (2001). Social ordering and health. In B.S. McEwen & H.M. Goodman (Eds.), *Handbook of physiology, Vol. IV: Coping with the environment: Neural and Endocrine Mechanisms* (pp. 533-546). New York: Oxford University Press.
- Cox, V.C., Paulus, P.B., & McCain, G. (1984). Prison crowding research: The relevance of prison housing standards and a general approach regarding crowding phenomena. *American Psychologist*, 39, 1148-1160.
- Craig, E.L. & Craig, R.E. (1999). Prison hospice: An unlikely success. *American Journal of Hospice and Palliative Medicine*, 16, 725-729.
- Donovick, P.J. & Burreight, R.G. (1984). Roots to the future: Gene-environment coaction and individual vulnerability to neural insult. In C.R. Almlie and S. Finger (Eds.), *The behavioral biology of early brain damage*. New York: Academic Press.
- Edgerly, E. S. & Donovan, P.J. (1998). Neuropsychological correlates of wandering in persons with dementia. *American J. Alzheimer's Disease*, 13, 317-329.
- Fazel, S., Hope, T., O'Donnell, I., Piper, M., & Jacoby, R. (2001). Health of elderly male prisoners: worse than the general population, worse than younger prisoners. *Age & Aging*, 30, 403-407.
- Fuller, J.L. (1979). Genetic analysis of deviant behavior. In J. D. Keehn (Ed.), *Psychopathology in Animals: research and clinical implications* (pp. 61-79). New York: Academic Press.
- Harrison, P.M. & Beck, A.J. (2004). Prisoners in 2003. *Bureau of Justice Statistics*, NCJ, 2053355.
- Harrison, P.M. & Beck, A.J. (2006). Prison and jail inmates at midyear 2005. *Bureau of Justice Statistics*, NCJ 213133.
- Holtzer, R., Burreight, R.G., & Donovan, P.J. (2004). The sensitivity of dual-task performance to cognitive status in aging. *Journal of Clinical and Experimental Neuropsychology*, 10, 230-238.
- Holtzer, R., Stern, Y., & Raktin, B.C. (2004). Age-related differences in executive control of working memory. *Memory & Cognition*, 24, 1333-1345.
- Holtzer, R., Stern, Y., & Raktin, B.C. (2005). Predicting age related dual-task effects with individual differences on neuropsychological tests. *Neuropsychology*, 19, 18-27.
- Holtzer, R., Tang, M.X., Devanand, D.P., Albert, S.M., Wegesin, D.J., Marder, K., et al. (2003). Psychopathological features in Alzheimer's Disease: Course and relationship with cognitive status. *Journal of American Geriatric Society*, 51, 953-960.
- Husaini, B. A., Sherkat, D.E., Moonis, M., Levine, R., Holzer, C., & Cain, V. A. (2003) Racial differences in the diagnosis of dementia and its effects on the use and costs of health care services. *Psychiatric Services*, 54, 92-96.
- Jeglic, E.L., Vanderhoff, H.A., & Donovan, P.J., (2005). The function of self-harm behavior in a forensic population. *International Journal of Offender Therapy & Comparative Criminology*, 49, 131-142.
- Jones, G., Connelly, M & Wagner, K. (2001). *Aging offenders and the Criminal Justice System*. <http://www.msccsp.org/publications/aging.html>
- Lindquist, C.H. & Lindquist, C.A. (1997). Gender differences in distress: mental health consequences of environmental stress among jail inmates. *Behavioral Sciences and the Law*, 15, 503-523.
- Lindquist, C.H. & Lindquist, C.A. (1999). Health behind bars: utilization and evaluation of medical care among jail inmates. *Journal of Community Health*, 24, 285-303
- Lishman, W.A. (1988). *Organic psychiatry: The psychological consequences of cerebral disorder* (3rd ed.). Oxford: Blackwell Science Ltd.
- Mara, C.M. (2002). Expansion of long-term care in the prison system: An aging inmate population poses policy and programmatic questions. *Journal of Aging & Social Policy*, 14, 43-61.
- Sachs-Ericsson, N. & Blazer, D.G. (2005). Racial differences in cognitive decline in a sample of community-dwelling older adults: The mediating role of education and literacy. *American J. of Geriatric Psychiatry*, 13, 968-975.
- Sapolsky, R.M. (2001). Physiological and pathophysiological implications of social stress in mammals. In B.S. McEwen & H.M. Goodman (Eds.), *Handbook of physiology, Vol. IV: Coping with the environment: Neural and Endocrine Mechanisms* (pp.517-532). New York: Oxford University Press.
- Stuss, D.T. & Alexander, M.P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research*, 63, 289-298.
- Turner, M.G., Hartman, J.L. & Bishop, D.M. (2007) The effects of prenatal problems, family functioning, and neighborhood disadvantage in predicting life-course-persistent offending. *Criminal Justice and Behavior*, 34, 1241-1261.
- Way, B.B., Miraglia, R., Sawyer, D. A., Beer, R. & Eddy, J. (2005). Factors related to suicide in New York state prisons. *International Journal of Law & Psychiatry*, 28, 207-221.

Williams, B.R., Wagner, L.A., Hallquist, M.N., & Donovick,
P.J. (In preparation) Exploring *g*: Prediction of general
intellectual ability by motor speed and complex process.

The Importance of Context: Probation Officers Working With Adolescent Sex Offenders

Teresa D. Griep
Loma Linda University, Lacombe, AB
E-mail: teresagriep@hotmail.com

When you are dealing with adolescent sex offenders, just how much do stereotypes and preconceived notions factor into the way these predators are treated by probation officers? It was found in the literature that attitudes of probation officers working with sex offenders were significantly more punitive when compared to working with non-sex offenders (Craig, 2005).

My interest in this topic developed three years ago when I began working for a Young Offenders Center, co-facilitating a sex offender treatment program. I wondered what would happen to them when they were released into the community. How would their probation officers view them? What role if any would context play in how the sex offenders were viewed by the probation officers? Would the probation officer view a given sex offender as a youth with a sex offense or just as a sex offender? How would their probation officers treat the sex offenders? What kind of relationship would a probation officer build with the adolescent sex offender?

Few studies have examined the experiences of probation officers working with sex offenders. Studies to date have explored areas including the characteristics of adolescent sex offenders, how to treat sex offenders, the impacts of workshops on attitudes towards sex offenders, and the responsibilities and risks involved for probation officers working with sex offenders. None have explored the personal opinions, assumptions, beliefs or attitudes of probation officers who work with sex offenders (Blanchette, 1996; City of Calgary, 2005; Chaffin, Bonner, & Pierce, 2002; Collier, 1998; Craig, 2005; Harris & Hanson, 2003; National Clearing House on Family Violence, 1997).

As a result of both the previously mentioned questions, and the deficiencies in the literature, I decided to do a little research into the perceptions, attitudes, beliefs and assumptions probation officers hold towards adolescent sex offenders.

For this limited qualitative research project I contacted three probation officers whose names I had been given by a professional in the criminal justice field. The three individuals interviewed, worked or had worked in youth probation and had been responsible for adolescent sex offenders. For this study, grounded theory methodology was used to compare individual experiences with the goal of identifying common themes in the data (Charmaz, 2006). Grounded theory research is a systemic, qualitative procedure used to create a general explanation that explains a process, action or interaction among people, in which a theory emerges from the data (Creswell, 2005). Each of the individuals was interviewed for one hour using intensive interviewing. Although the sample size was small, and any conclusions from the mentioned themes are speculative, the research did uncover some interesting thoughts on the experiences of probation officers who work with adolescent sex offenders.

The findings were organized around one salient theme that had emerged from the data, that is, the importance of context in working with adolescent sex offenders. This overarching theme contained three sub-themes: a) the importance of forming a relationship with the offenders, b) the importance of seeing the crime in the context of the person and not the person as their crime, and c) the importance of probation officers seeing themselves in the role of an educator.

Although three sub-themes have been identified, they do not exist as separate entities under the salient theme. Many overlapping ideas and connecting pieces exist in the data to form a story of the importance of context in working with adolescent sex offenders.

The Importance of Forming a Relationship with the Offenders

The probation officers interviewed expressed a strong belief in building rapport with the sex

offender as a key factor to the re-integration of the youth back into the community. Although developing a good relationship was essential, it had to be realistic. On the one hand, the probation officers worked to engage in a positive, strength-based, rapport building approach and, on the other, they had to keep in mind the precautionary measures and concerns needed to ensure public safety. If a realistic relationship was not developed then the safety of the community and the probation officer's job were jeopardized. Forming a relationship with the adolescent sex offender involved finding a balance between helping the youth and protecting the community.

The Importance of Seeing the Crime in the Context of the Person and not the Person as Their Crime

This sub-theme is a further description of the importance of forming a relationship with a sex offender. Having to take a history of an offender and using "collateral contacts" as part of their policy, probation officers gain a better understanding of where the offender came from and what influences have been in his or her life. Forming a realistic relationship is taking the crime into consideration but not making it the focus of the relationship. As stated by one interviewee, "the more you know about a person, the better decisions you can make about the person in terms of risk levels, conditions and restrictions." The crime cannot be forgotten, but can be put into the context of the youth's social and personal circumstances.

Another important aspect of the role of the probation officer and public safety is to be an advocate for the community. This poses with some difficulty because the community as well as others involved with the sex offender sees sex offenders for their crimes rather than the crime in the context of the person. It is the stereotypical attitudes existing in the community that the probation officers have to deal with when at the same time being partially responsible for the very stereotype being labeled.

The Importance of Probation Officers Seeing Themselves in the Role of an Educator

In order to form a relationship and view an adolescent sex offenders' crime in the context of

the person, a probation officer approaches the youth as an educator. In seeing themselves as an educator, or in a teacher role, a probation officer shares the responsibility for change with the sex offender. This in turn empowers the sex offender and makes the probation officer less vulnerable to burnout. With an educator role, probation officers are there to help sex offenders re-integrate back into the community; they are there to share the work with the sex offender, rather than doing all the work on their own. In this sense, success is measured as the result of team work involving the adolescent sex offender, the probation officer and others who support the youth from a positive approach.

Conclusions

When probation officers first start out in their job they enter the job with preconceived notions that have developed from such influences as their family's beliefs, religion, their personal morals and values, their culture and the media. The more time they spend in the job, the more they learn about seeing the offender in the context of their entire life story. The attitude of all three interviewees is one of being helpful. In order to be helpful they need to be an educator, and in order to take on the role of an educator, probation officers need to build a relationship with the adolescent, a realistic relationship. This realistic relationship entails probation officers to look beyond the face value – beyond the crime, to discover how the crime became a part of the offenders' history.

The findings of this study not only demonstrate the complexities involved for probation officers with sex offenders on their case load, but also the importance of context when working with sex offenders.

Implications that can be gathered from this research include the need for all professional services and community members involved with sex offenders to establish a diligent approach to thoroughly understand and see the offense in the context of the person. It may also be useful for therapists, counselors, psychologists and other mental health professionals to be aware of the complexities probation officers face. It can better equip the clinicians by giving them a better

understanding of the issues and experiences of the probation officers they may work with.

Implications for future research include the area of educating the community on sex offenders as an effort to begin breaking down the stereotypes and preconceived thinking that exist.

References

- Blanchette, K. (1996, August). *Sex offender assessment, treatment and recidivism: A literature review*. Retrieved September 29, 2006, from Correctional Service Canada Web site: http://www.csc-scc.gc.ca/text/rsrch/reports/r48/r48e_e.shtml
- The City of Calgary. (2005, November). What youth probation does. Retrieved September 22, 2006, from The City of Calgary Web site: http://www.calgary.ca/portal/server.pt/gateway/PTARG_S_0_2_104_0_0_35/http:/content.calgary.ca/CCA/Cityliving/PeopleResource/Youth/HelpforYouth/YouthProbation/WhatYouthProbationDoes/WhatYouthProbationDoes.htm?Prn=1content.calgary.ca/CCA/City%20Living/People%20Resource/Youth/Help%20for%20Youth/Youth%20Probation/What%20Youth%20Probation%20Does/What%20Youth%20Probation%20Does.htm
- Chaffin, M., Bonner, B., & Pierce, K. (2002). *What research shows about adolescent sex offenders*. Retrieved September 22, 2006, from National
- District Attorneys Association Web site: http://www.ndaa.org/publications/newsletters/in_re_vol_ume_5_number_2_2002.html
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Collier, A. (1998). The probation services work with sex offenders: Findings and implications of thematic inspection. *Child Abuse Review*, 7, 391-401.
- Craig, L. (2005). The impact of training on attitudes towards sex offenders. *Journal of Sexual Aggression*, 11(2), 197-207.
- Creswell, J.W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd. Ed.). Upper Saddle River, NJ: Pearson Education.
- Harris, A., & Hanson, R. (2003). The dynamic supervision project: Improving the community supervision of sex offenders. *Corrections Today*, 65(5), 60-63.
- National Clearinghouse on Family Violence. (1997). Adolescent sex offenders. Retrieved September 22, 2006, from Public Health Agency of Canada Web site: http://www.phac-aspc.gc.ca/ncfv-cnivf/familyviolence/html/nfntxadolinfractions_e.html

Juvenile Sex Offending Within an Attachment Theory Framework.

Alicia L. Szielasko
Acadia University
E-mail: aszielas@dal.ca

Douglas K. Symons
Acadia University

Joan Boutilier
Annapolis Valley Health

Attachment theory has influenced approaches to the assessment and treatment of juvenile sex offence for over 15 years (Marshall, Hudson, & Hodgkinson, 1993). As overviewed in the work of various researchers, there are several attachment-based criminogenic factors (e.g., Marshall 1994, 1996; Ward & Hudson, 1998), which include poor emotional regulation, emotional loneliness, limited or inappropriate intimacy skills, dysfunctional schemas of relationships, and poor social support. These predictors of risk are all relationship-orientated stable dynamic predictors factors that are typically targeted in sex offender treatment programs. Although attachment researchers have successfully explained links between interpersonal deficits and sexually aggressive behaviour (e.g., Davis, 2006; Fonagy, Target, Steele, & Steele, 1997; Hudson & Ward, 1997), they have yet to demonstrate a model that can delineate juvenile sex offenders from non-sexual juvenile delinquents, and classify juvenile sex offenders based on attachment-related characteristics. These goals remain important in order to improve an understanding of etiology, risk prediction, and treatment outcomes for juvenile sex offenders. This discussion paper will focus on etiological and assessment issues of sex offending from an attachment-informed framework, as well as introducing a new measure of attachment that is specific to sexual behaviour.

Attachment can be described as a capacity to form affectionate bonds and relationships with others (Bowlby, 1973). Although Bowlby's original work on attachment theory (1969/1982) was initially applied to infant-caregiver relationships, research of the past two decades has affirmed the continuity and predictive nature of attachment styles across various developmental stages, including adolescence (e.g., Greenberg, 1999; Sroufe, Egeland, Carleson, & Collins, 2005). This is consistent with Bowlby's initial suggestion that attachment theory operates "from cradle to grave" across

the lifespan, and is best seen as a model of personality development.

In infancy, attachment begins as a coordinated relationship between a caregiver and her or his child that develops over time to protect the infant, facilitate exploration, resolve distress during threat, and increase the infant's chances of survival to reproductive age (Bowlby, 1969/1982, 1973). During early development, there is a critical window in which an emotional-cognitive goal for personality is to establish trust and confidence in the responsiveness of attachment figures. Based on these early experiences, a child eventually develops expectations of how attachment figures will respond in certain situations, that is, whether figures meet needs and do so in a consistent fashion. These expectations are the groundwork for working models or inner representations of the individuals' experience of the self and others within the context of relationships. Attachment theory proposes that these expectations influence other close relationships later on in life (Hazan & Shaver, 1994). For instance, when attachment figures are poorly attuned or inconsistently responsive, a child is likely to learn that his or her needs are not important and unlikely to be met, others are untrustworthy, and the world is generally not a "safe place". These experiences are thought to form the basis of insecure attachments, which are primarily classified as either avoidant or anxious in childhood. Research has shown that sex offenders tend to have either anxious or avoidant types of insecure attachment (Ward, Hudson, & Marshall, 1996), and although both styles are characterized by different patterns of experience, each is associated with detrimental implications for emotional development and metacognition (Crittenden, 1992).

Attachment-related working models form the basis of metacognition, or the "ability to reflect on the minds of self and others" (Rich, 2006, p. 165), and associated deficits result in a lack of

insight, failure to anticipate or appreciate the consequences of antisocial behaviour, poor behavioural regulations manifested as poor coping skills and impulsivity. Behaviour can be described as *reactive* rather than *reflective* (Fonagy, 2004). It has been argued that such deficits are associated with alexithymia, that is, a poor ability to identify emotions (Donovan & Leavitt, 1985). Alexithymic individuals have a poor sense of self awareness, marked inability to express emotions, and limited empathic skills (Parker, Taylor, & Bagby, 1993; Taylor, 2000). In addition, insecurely-attached children often grow up to have difficulty with intimacy and relationships, and often become socially isolated from their peers (Ainsworth, Blehar, Waters & Wall, 1978; Feeney & Noller, 1990; Sroufe et al., 2005). It has been well-established through research and clinical experience that social isolation and empathy deficits are two key risk factors for sex offending (Marshall et al., 1993).

Empathy provides a basis for the kind of intimacy required for secure relationships. Cassidy (2001) describes how a secure attachment style is associated with a capacity to participate in intimate relationships based on someone's ability to seek care, give care to others, feel comfortable with an autonomous self, and negotiate how much closeness they prefer. She reviews evidence that links this intimacy within the attachment system to the sexual behaviour system, particularly in adolescence when sexuality is so important. Going further, insecure attachment style serves as a risk factor for specific kinds of problematic sexual behaviours. For example, Cassidy (2001) suggests that secure attachment style is related to mutually initiated sexual activity within committed relationships. By contrast, anxious attachment is associated with higher probability of exhibitionism, voyeurism, and dominance/bondage within sexual relationships, and avoidant individuals are more likely to take coercive approaches toward sexual relations that are devoid of intimacy (Kunce & Shaver, 1994). Although these kinds of sexual behaviours certainly present in normative populations, they are seen more commonly among sexual offenders.

There are a number of limitations associated with assessment of attachment style in adolescence specifically related to sexual behaviour and sex offenders. One limitation is that much of the research with sex offenders has

been based on retrospective data from adult sex offenders (e.g., Worling, 1995), as there has been little longitudinal research with juvenile offenders that has spanned through to adulthood. This type of research is difficult to conduct due to feasibility, time constraints, and low base-rate behaviour, but it is the only way of reliably assessing the various developmental outcomes associated with the risk of re-offence. Another limitation is that attachment style in adolescence is typically assessed through reports of either early parent-child relationships or romantic attachment styles (see review by O'Connor & Byrne, 2007). The problem with using the latter type of measure is that assessing attachment with early caregivers might not reflect important qualitative changes that occur during adolescence, as this is a time when attachment relationships shift abruptly from parents to peers as well as potential romantic partners (see Allan & Land, 1999). The obvious problem with assessing romantic attachment in adolescents is that many of these youth have not established romantic relationships, particularly sexually aggressive youth whose intimacy deficits often preclude any type of close relationship (Hudson & Ward, 1997). Alternatively, sexual behaviour in adolescence often occurs outside of romantic relationships characterised by the communication and commitment of many relationships in adulthood. A final issue is the approach taken to the assessment of attachment style, which often takes the form of discourse analyses of standard questions about critical and historical interactions with caregivers. Such methods are too time-consuming to learn to administer and use easily in the field. Self-report measures are easier to use as long as they are not too transparent to respondents.

Many of these limitations can be addressed through domain-specific measures of attachment that capture behaviours and interpersonal styles relevant to sexual behaviour in adolescence. In the case of juvenile sex offenders, it would be meaningful to assess attachment styles within the context of sexual behaviour. Indeed, recent developmental research has shown that attachment style is associated with predictable patterns of sexual behaviour in adults and adolescents (e.g., Davis, 2006; Davis, Shaver, & Vernon, 2004). For instance, Davis et al. (2004) argue that sexual behaviour is a form of proximity-seeking and can be triggered by the same factors that activate the attachment system,

Table 1: Positive and negatively worded sample items from the Sexual Attachment Measure

Secure Positive	Ambivalent Positive	Avoidant Positive
5. I prefer sex that is planned and discussed with my partner	1. I sometimes want to have sex just for reassurance that my partner still wants me.	15. I am comfortable having sex with someone that I barely know.
23. Friendship is an important part of my sexual relationships.	13. I tend to get preoccupied with sex when I am in a relationship.	18. I tend to "cool-off" sexually when my partner wants to get too close.
39. Trust is an important part of my sexual relationship(s).	25. I have had sex when I didn't want to just to keep my partner happy.	27. I have tricked someone into having sex with me.
48. It is important that both my sexual partner and I respect each other equally.	38. I tend to attract sexual partners who want to keep me at a distance.	46. For me, physical pleasure is the most important reason to have sex.
Secure Negative	Ambivalent Negative	Avoidant Negative
2. I feel comfortable with casual sexual relationships (e.g., "one-night stands").	4. My partner's sexual pleasure is not the most important part of sex.	3. It is important for me to be emotionally close to a sexual partner.
8. If I wanted sex and my partner didn't I would use pressure to get what I wanted.	16. I don't get upset when my partner refuses my sexual advances.	9. Discussing sex with my partner is important.
20. I need to have sex to express affection to my partner.	28. I rarely worry about my own sexual performance.	11. My sexual experiences have mostly occurred within committed relationships.
21. I don't talk about sex with my partner.	47. I feel comfortable talking about my own needs to my sexual partner.	30. I don't like playing around with lots of sexual partners.

such as threat to an attachment relationship, internal threats such as emotional distress, and external threats and dangers. Insecure attachments are associated with sex behaviour that is activated by negative cognitive and affective states reflective of early attachment experiences.

Although anxious and avoidant attachment styles are activated by different patterns of sexual motivation, each predicts a specific attachment-related pathway to sexual aggression (Davis, 2006). Anxiously-attached individuals have overactive attachment needs that involve excessive proximity-seeking and poor distress-regulation. As Davis et al. (2004) argue, attachment-related distress is often linked to sexual motivation, and Davis (2006) has conceptualized a link between anxious attachment and sexually coercive behaviour in male youth. For instance, attachment-related distress, such as perceived threat to or dissolution of the relationship, may trigger sexual motivation in these individuals, who then

turn to the intended sexual partner for distress regulation. Partner refusal may be construed as rejection, which maintains distress and further activates the attachment system. This process leads to increased negative emotions and sexual motivation. All of these dynamics are coupled with the fact that ambivalently-attached males are more likely to suffer pre-existing deficits in interpersonal skills and confidence. Although they desire intimacy as a form of attachment-related proximity, they have considerable difficulty establishing intimate relationships and may try and satisfy their intimacy needs through sexual behaviour. However, poor communication and negotiation of sexual needs may result in rejection of their sexual advances (Davis, 2006). For these reasons, these youth may be more prone to demonstrating reactive sexual aggression within some form of attachment relationship (e.g., sibling, friend or relative).

On the other hand, avoidant attachment is associated with a habitual deactivation of

attachment needs such as closeness and intimacy, and sex is more associated with physical gratification and other instrumental goals (e.g., enhancement of peer status). These youth are inclined to take an emotionally detached and exploitative approach to relationships. For one, they are apt to engage in spontaneous sexual encounters with multiple sexual partners to whom they are not emotionally connected. As Davis (2006) argues, avoidant males are also more likely to demonstrate instrumental sexual aggression with people they are not emotionally connected to (e.g., strangers or acquaintances) for reasons of self-enhancement or gratification.

Given the link between attachment styles and sexual motivation, and no known assessment tools to capture these patterns, we developed a measure of sexual attachment in our current research called the *Sexual Attachment Measure* (SAM; see Szielasko, Symons, & Price, 2007). Such a measure could address limitations of attachment style assessment noted earlier by addressing patterns of attachment that are specific to sexual behaviour. This is a 48-item self-report measure of sexual attachment style that provides dimensional scores on the subscales of security, anxious-ambivalence, and avoidance. Behaviourally-based items have been written to reflect these styles within sexual relationships. Sample items which indicate the presence and absence of behaviours on each subscale are contained in Figure 1. Each subscale has shown good internal consistency, appropriate intercorrelations with other subscales, and concurrent validation with measures of romantic attachment and other sexual style measures. Most importantly, even though this study examined normally developing 18 and 19 year-olds, there were still positive relations between avoidance scores and various forms of sexual coercion.

For the purposes of this paper, it would be important to reflect on how juvenile sex offenders would respond to these kinds of items, as a number of related questions arise. For instance, is sexually abusive behaviour among youth associated with different relational contexts for those with anxious and avoidant attachment styles as Davis (2006) would argue? Specifically, would anxious sexuality be linked to sexual abuse within established relationships, and triggered by conditions associated with relationship insecurity and attachment-related

distress (e.g., separation from parent or a favoured social worker). Alternatively, would avoidant sexuality be linked to sexually abusive behaviour within unestablished relationships to gain access to an initial target, as opposed to being motivated by threatened needs within established relationships? Answers to these kinds of questions would elucidate how specific attachment styles may reflect certain kinds of problematic sexual behaviour, and promises to advance research and clinical practice within an attachment-informed model of sex offending.

References

- Ainsworth, M. D. S., Blehar, M. C., Waters, E. & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.
- Allen, J. P., & Land, D. (1999). Attachment in adolescence. In J. Cassidy, P. R. Shaver (Eds.). *Handbook of attachment: Theory, research, and clinical applications* (pp. 355-377). Guilford: New York.
- Bowlby, J. (1969/1982). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.
- Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation*. New York: Basic.
- Cassidy, J. (2001). Truth, lies, and intimacy: An attachment perspective. *Attachment and Human Development*, 3, 121-155.
- Crittenden, P. M. (1992). Treatment of anxious attachment in infancy and early childhood. *Development and Psychopathology*, 4, 575-602.
- Davis, D. (2006). Attachment-related pathways to sexual coercion. In M. Mikulincer & G. S. Goodman (Eds.). *Dynamics of romantic love: Attachment, caregiving, and sex* (pp. 293-336). New York: Guilford Press.
- Davis, D., Shaver, P. R., & Vernon, M. L. (2004). Attachment style and subjective motivations for sex. *Personality and Social Psychology Bulletin*, 30, 1076-1090.
- Donovan, W. L., & Leavitt, L. A. (1985). Physiologic assessment of mother-infant attachment. *Journal of the American Academy of Child Psychology*, 24, 65-70.
- Feeney, J. A., & Noller, P. (1990). Attachment style as a predictor of adult romance relationships. *Journal of Personality and Social Psychology*, 58, 281-291.
- Fonagy, P. (2004). The developmental roots of violence in the failure of mentalization. In F. Pfafflin & G. Adshhead (Eds), *A matter of security: The application of attachment theory to forensic psychiatry and psychotherapy* (pp. 13-56). London: Jessica Kingsley Publishers.
- Fonagy, P., Target, M., Steele, M. M., & Steele, H. (1997). The development of violence and crime as it relates to security of attachment. In J. D. Osofsky (Ed.),

- Children in a violent society* (pp. 150-177). New York: Guilford.
- Greenberg, M. T. (1999). Attachment and psychopathology in childhood. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of Attachment: Theory, Research, and clinical applications* (pp. 469-496). New York: The Guilford Press.
- Hazan, C., & Shaver, P.R. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry*, 5, 1-22.
- Hudson, S. M., & Ward, T. (1997). Intimacy, loneliness and attachment styles in sexual offenders. *Journal of Interpersonal Violence*, 12(3), 323-339.
- Marshall, W. L. (1994). Poverty of bonding and deficiencies in intimate relationships in the Sexually aggressive. *Criminologie*, 27, 55-69.
- Marshall, W. L. (1996). Assessment, treatment and theorizing about sex offenders: Developments during the past twenty years and future directions. *Criminal Justice And Behaviour*, 23, 162-199.
- Marshall, W. L., & Marshall, L. E. (2000). The origins of sex offending. *Trauma, Violence and Abuse*, 1, 250-263.
- Marshall, W. L., Hudson, S. M., & Hodgkinson, S. (1993). The importance of attachment bonds in the development of juvenile sex offending. In H. E. Barbaree, W. L. Marshall & S. M. Hudson (Eds.), *The juvenile sex offender*(pp. 164-181). New York: Guilford Press.
- O'Connor, T.G., & Byrne, G. (2007). Attachment measures for research and practice. *Child and Adolescent Mental Health*, 12, 187-192.
- Parker, J. D., Taylor, G. J., & Bagby, R. M. (1993). Alexithymia and the processing of emotional stimuli: An experimental study. *New Trends in Experimental and Clinical Psychiatry*, 9, 9-14.
- Rich, P. (2006). *Attachment and Sexual Offending: Understanding and Applying Attachment Theory to the Treatment of Juvenile Sexual Offending*. West Sussex, England: John Wiley & Sons, Ltd.
- Sroufe, L.A., Egeland, B., Carleson, E. A., & Collins, W. A. (2005). *The development of person: The Minnesota study of risk and adaptation from birth to adulthood*. New York: Guilford.
- Szielasko, A. L., Symons, D. K., & Price, E.L. (2007). The role of attachment in adolescent sexual risk-taking. Manuscript under review.
- Taylor, G. J. (2000). Recent developments in alexithymia: Theory and research. *Canadian Journal of Psychiatry*, 45, 134-142.
- Ward, T., Hudson, S. M., & Marshall, W. L. (1996). Attachment style in sex offenders: A preliminary study. *The Journal of Sex Research*, 33, 17-26.
- Ward, T. & Hudson, S. M. (1998). The construction and development theory in the sexual offending area: A metatheoretical framework. *Sexual Abuse*, 9, 47-64.
- Ward, T., & Keenan, T. (1999). Child molesters' implicit theories. *Journal of Interpersonal Violence*, 14, 821-838.
- Worling, J. (1995). Sexual abuse histories of adolescent male sex offenders: Differences on the basis of the age and gender of their victims. *Journal of Abnormal Psychology*, 104, 610-616.

Section F

Special Topics

Criminal Trajectories from Adolescence to Adulthood in an Ontario Sample of Offenders

David M. Day
Ryerson University
E-mail: dday@psych.ryerson.ca

Thierry Duchesne
Laval University

Ye Sun
York University

Irene Bevc
Hincks-Dellcrest Children's Centre

Jeffrey S. Rosenthal
University of Toronto

Frances Theodor
Hincks-Dellcrest Children's Centre

Since Blumstein et al. (1986) published their seminal work on criminal careers, the criminal career paradigm (CCP) has dominated the criminology literature. With an emphasis on longitudinal methodology and sophisticated mathematical approaches to model criminal trajectories, the CCP has made significant contributions to elucidating the nature and pattern of crime over time.

A criminal career is defined as “the longitudinal sequence of offending committed by an individual offender” (Blumstein et al., 1986, p. 12) that is characterized during a lifetime by three components: an initiation or onset, a termination or end, and a duration or career length (Blumstein et al., 1988). During their career, offenders may display changes and continuities in criminal activity on a variety of dimensions, including rate, type, timing, versatility, and severity. It is the pattern of transition and stability on these sorts of variables across different developmental periods as well as the underlying reasons for the observed patterns that is of interest to researchers, theoreticians, practitioners, and policy makers.

The CCP also implies that offenders vary on the dimensions of rate, type, severity, versatility, etc. The task for the CCP is to model the criminal offending data to take into account this inherent heterogeneity. Semi-parametric group-based trajectory analyses (Nagin, 2005) provide an ideal way to address this issue by estimating distinct unobserved (latent) trajectory groups based

on longitudinal rates of offending. Drawing upon the notion of a criminal career, the two aims of the present study were to (a) describe the criminal trajectories of the Toronto sample on the dimensions of rate, type, versatility, and severity and (b) estimate distinct, latent criminal trajectories using current techniques of group-based trajectory analysis.

Method

Sample

The sample comprised 378 males who had been sentenced between 1986-1996 to one of two open custody young offender facilities in Toronto. Their mean age at the time of admission into the youth homes was 17.6 years ($SD = .85$). The mean age at first conviction was 15.5 years ($SD = 1.8$). The sample was, on average, 27.6 years ($SD = 2.6$) as of March 17, 2001, the time of the most recent follow-up.

Data

The official criminal data were derived from four sources: (a) the (Ontario) Ministry of Community and Social Services (MCSS) for Phase I young offender records; (b) the (Ontario) Ministry of Correctional Services (MCS) for Phase II young offender and adult records; (c) the Canadian Police Information Centre (CPIC) for youth and adult records; and (d) for additional young offender offences, the Predisposition Reports (PDR) maintained in the clinical files of the

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 143-150. Public Safety Canada.

This research was supported by grants from the Ministry of Children and Youth Services, Ryerson University, The Hincks-Dellcrest Children's Centre, the Samuel Rogers Memorial Trust, and the Natural Sciences and Engineering Research Council of Canada.

children’s mental health centre that operated the young offender facilities.

The criminal trajectories were tracked for an average of 12.1 years (*SD* = 3.0), from early adolescence into adulthood, with 73% of the sample being followed for 10 years or more. The criminal activity variable of interest was a count of *all* their unique court contacts, that is, all court contacts arising from a new set of charges. The task was to arrange the unique court contacts into a temporally sequenced order of criminal activity.

Coding

Criminal records were coded for a range of variables for each unique court contact, including disposition date, disposition received, sentence length in days, including both time given and time served, and offence type, based on the seriousness rating of the most serious offence (MSO) for each court contact. A number of variables were created to measure various dimensions of the criminal career concept, including frequency, rate, versatility, and severity.

Frequency was a simple count of the total number of unique court contacts amassed by each individual.

Rate was the frequency of court contacts committed in a given time period (e.g., a year) corrected by two variables, *time-at-risk* and the *age at offence-age at court contact time lag*. The statistical techniques we used for these adjustments are described in Day et al. (2007).

Offence Type denoted the types of offences that were committed, grouped into five offence type categories: property, violent, drug, sex, and “other” (e.g., administration of justice or “breach” offences, obstructing justice, traffic offences).

Versatility was measured with the Diversity of Offending Index score (D) (Sullivan et al., 2006) calculated as follows:

$$D = 1 - \sum_{m=1}^M p_m^2$$

“in which p equals the proportion of offences in crime category ‘m’” (p. 207). The D score was

based on a count of the number of *different* offences committed at each conviction, across six broad offence types (property, violent, drug, sex, other, and breach), not just the MSO, even if the offences were included in the same broad offence type.

Severity was a measure of the seriousness of the offences taken from the MCS Statistical Reporting System User Manual (Ministry of the Solicitor General and Correctional Services, 1995). The severity ratings were ranked from 1 (murder/attempt) = most serious to 26 (unknown) = least serious.

Results

Over the duration of the tracking period, the sample amassed a total of 4,964 unique court contacts. This amounted to an average of 13.1 court contacts for each offender (*SD* = 9.6). The average criminal career length was 8.4 years (*SD* = 4.5), from ages 15.5 to 23.9 years.

At What Age Does the Court Contact Rate Peak?

The age-crime curve that results from averaging the court contact histories, corrected for the offence-court contact time lag and time-at-risk, is plotted in Figure 1. The curve is unimodal and skewed and the offence rate peaks at age 17, gradually declining thereafter.

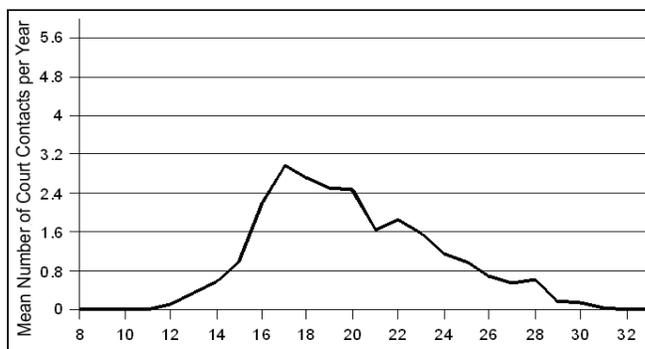


Figure 1. Average Age-Crime Curve Based on Individual Careers

Do Offenders Become More Versatile in their Offending Over Time?

This question was examined by calculating D scores for each individual by age. The results (see

Figure 2) indicate that the diversity of offending increased sharply from ages 12 to 16 years, reached a plateau until age 20, and decreased until age 29. The second peak at age 30 reflects a high rate of diversity among a small subgroup of this sample.



Figure 2. Diversity (D) Scores by Age

Are Some Offences More Common at Different Developmental Periods?

To address this question, a Relative Offence Type Involvement (ROTI) score was calculated for each individual at each age using the following formula:

$$ROTI_{ij} = \frac{r_{ij}}{\sum r_i} \times 100,$$

where r = age-specific court contact frequency, i = age category, and j = offence type. The ROTI score is based on *all* the different offences committed and is calculated as the total number of charges for each of five offence types (breaches

were included in the “other” category) committed at a given age divided by the total number of charges incurred at that age. At each age, the scores across offence types sum to 100%.

The pattern of scores in Figure 3 yields four interesting results. First, the relative involvement in property offences was much higher in early adolescence than at any other period. Second, as the D scores reached their peak at ages 16 to 20 years, the relative involvement in property offences decreased and involvement in violent and “other” types of offences increased. Third, sexual offences were primarily committed during adolescence and, fourth, drug offences were primarily an adult pursuit.

Do Offences Become More Severe over Time?

A quadratic equation was fit to the severity by age data to generate the curve presented in Figure 4 (actual and predicted values are shown). As indicated by the curve, the severity level peaks at about age 24, where a lower score denotes greater severity.

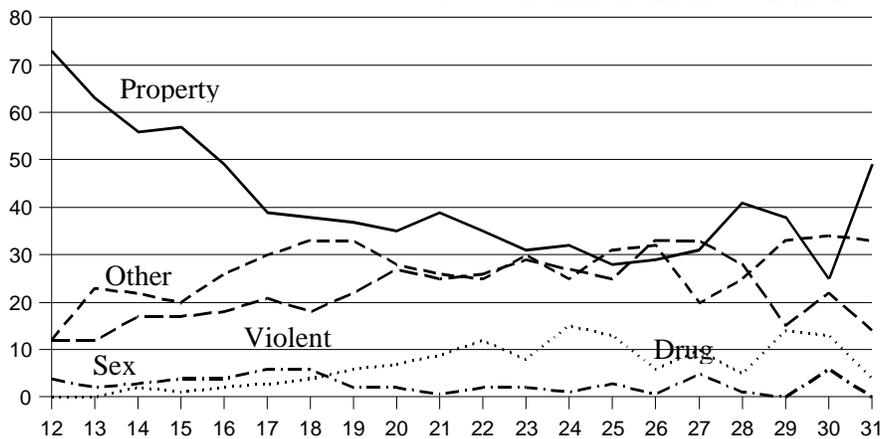
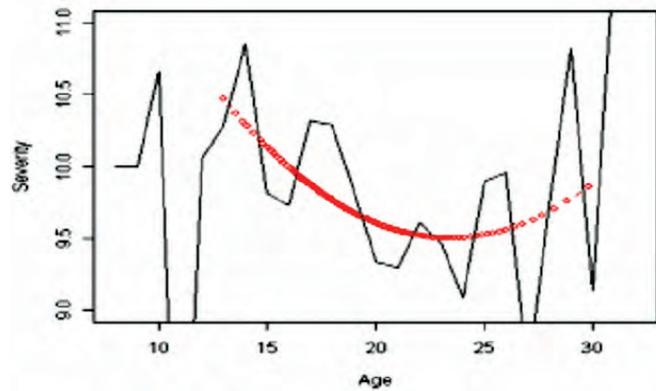


Figure 3. Relative Offence Type Involvement (ROTI) Scores by Age for Five Offence Types

Semi-Parametric Group-Based Trajectories: Making Sense of Heterogeneity

Criminal trajectories were estimated using a semi-parametric group-based trajectory model, a special application of the finite mixture modeling framework. This framework assumes that the sample comprises a finite number of unobserved (latent) groups. Because the criminal activity variable takes the form of an event count, we modeled the data as variations on the Poisson process. We first tested the homogeneity of the court contact variable and found that the data were such that it would be better to use a mixture model. Our Poisson model will be:

$$\log(\lambda_{it}^k) = \beta_0^k + \beta_1^k \text{Age}_{it} + \beta_2^k \text{Age}_{it}^2 + \beta_3^k \text{Age}_{it}^3$$

where the parameter λ_{it}^k is the predicted rate of court contacts for individual i at age t given membership in group k . Following the method described by Blokland et al. (2005), the β parameters were estimated by the method of maximum likelihood under the assumption that, within trajectory groups, the number of court contacts followed a Poisson process with rate parameter λ_{it}^k . The model was applied using PROC TRAJ, a SAS-based procedure described by Jones et al. (2001).

The Bayesian Information Criterion (BIC) was

used to determine that the optimal number of latent groups was three (BIC = -9956.54). Using the maximum likelihood estimate to obtain coefficients (β), where $k = 3$, we derived the following results:

$$\log(\lambda_{it}^1) = -41.62 + 5.76\text{Age}^1 + -0.25\text{Age}^2 + 0.00\text{Age}^3$$

$$\log(\lambda_{it}^2) = -30.12 + 4.27\text{Age}^1 + -0.19\text{Age}^2 + 0.00\text{Age}^3$$

$$\log(\lambda_{it}^3) = -25.66 + 3.43\text{Age}^1 + -0.14\text{Age}^2 + 0.00\text{Age}^3$$

Assignment of group membership was based on the *posterior probability* of individual i 's membership in group k (see Nagin, 2005 for a discussion of the posterior probability). The actual and predicted group trajectories for the three-group model are presented in Figure 5.

Group 1 (Low Rate) comprised 59.5% of the sample. For individuals following the LR trajectory, the average number of (corrected) court contacts was 8.8 ($SD = 5.5$). Their average criminal career length lasted for 6.7 years, from ages 16.0 to 22.7 years. This trajectory group also spent the least amount of time in secure custody, having been sentenced, on average, to a total of 299.90 days ($SD = 353.91$).

Criminal Trajectories for Three-Group Model

Zero Inflated Poisson Model

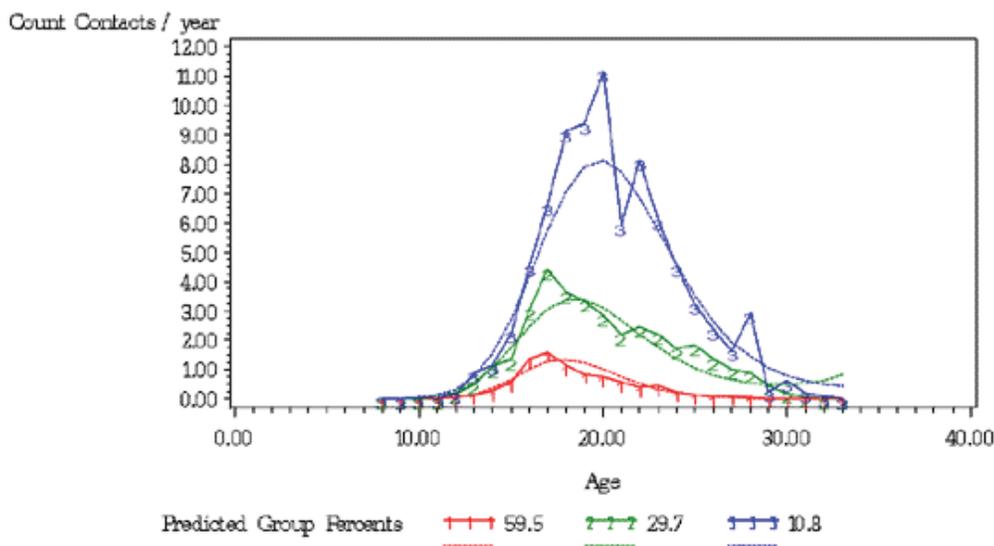


Figure 5. Criminal Trajectories for Three-Group Model

Group 2 (Moderate Rate) comprised 29.6% of the sample. The average individual in the MR trajectory incurred 35.0 (corrected) court contacts ($SD = 15.3$). Their average criminal career lasted for 11.3 years, beginning at age 14.8 and ending at age 26.1 years. This trajectory group was sentenced, on average, to a total of 1,284.3 days ($SD = 935.3$) in secure custody.

Group 3 (High Rate) comprised 10.8% of the sample. For individuals following the HR trajectory, the average number of (corrected) court contacts was 80.9 ($SD = 38.3$). Their average criminal career lasted for 10.1 years, beginning at age 14.6 and ending at age 24.7 years. This trajectory group also spent the most amount of time in secure custody, having been sentenced, on average, to a total of 3,026.3 days ($SD = 1,957.9$).

Discussion

The aims of this study were to describe the nature and pattern of criminal offending over time on the dimensions of rate, type, diversity, and severity and to estimate latent criminal trajectory groups using current techniques of group-based trajectory analysis. The results for the various criminal career dimensions are generally congruent with findings from other longitudinal studies. First, the aggregated age-crime curve generated for the Toronto sample resembles the classic age-crime curve reported in many studies (Blumstein et al., 1988). Second, while the rate of offending decreased into adulthood, the diversity and severity of offending increased, followed by a moderate decline. This increase in diversity and severity is an interesting finding that may reflect either a “normative” delinquent trend, like the age-crime curve, or an atypical pattern that is in need of further investigation.

In another way, the observed trajectory of type, versatility, and severity of offending may reflect a marked developmental shift as these young people negotiate the transition from adolescence to adulthood, a time when life paths become more sharply focused (Johnson et al., 2004). As the Toronto sample reached late adolescence, they became more physically mature, perhaps more menacing, threatening, and deeply entrenched in a criminal lifestyle, possibly due to an involvement in street gangs (Thornberry, 2005). This trend

may reflect a narrowing of options for them in terms of engagement in legitimate employment and academic opportunities.

According to developmental theory, involvement in serious antisocial behaviour during adolescence, particularly if it begins at an early age, is protracted, and involves contact with the justice system, may lead to a disruption in normative developmental processes, bringing about a premature transition from adolescence into adulthood and a concomitant redefinition of roles and contexts (e.g., being processed as a “criminal,” making court appearances, and spending a great deal of time with police, correctional, and probation officers) (Johnson et al., 2004). It also impedes the young person's ability to accomplish the developmental tasks of adolescence, such as completing school, developing positive peer relations, and forming a healthy and integrated sense of self, referred to as “adaptational failure” (Masten & Coatsworth, 1998). The cumulative impact is a continued disruption in normative developmental functioning that can interfere with the person's ability to develop the requisite skills to assume the socially accepted roles and expectations of adulthood. This process can result in an increased likelihood of maintaining criminal activity into adulthood, as opportunities for completing high school and entering the labour force diminish. This entrenchment in a criminal lifestyle may be particularly acute for members of the HR group who showed the deepest involvement in criminal activity.

However, caution must be exercised in describing these outcomes as developmental trajectories are meant to be understood as probabilistic not deterministic (Dumas & Nilson, 2003). Considerable plasticity in adaptation and adjustment allows for both continuity and discontinuity in developmental outcomes. This opens up the possibility for rehabilitative efforts to provide missed opportunities for youth in contact with the justice system to facilitate their positive growth and development. Ideally, such intervention strategies are informed by a thorough understanding of developmental trajectories of offending behaviour. As well, this characterization may only apply to a relatively small number of cases. For example, based on the

results of our group-based trajectory analyses, the majority of the offenders in the Toronto sample followed a LR course whose involvement in criminal activity was significantly less than members of either the MR or HR groups. This issue could be a focus for further investigation.

References

- Blokland, A. A. J., Nagin, D.S., & Nieuwbeerta, P. (2005). Life span offending trajectories of a Dutch conviction cohort. *Criminology*, *43*, 919-954.
- Blumstein, A., Cohen, J., & Farrington, D. P. (1988). Criminal career research: Its value for criminology. *Criminology*, *26*, 1-35.
- Blumstein, A., Cohen, J., Roth, J., & Visher, C. (1986). *Criminal careers and "career criminals"* (Vols. 1 & 2). Washington, DC: National Academy Press.
- Day, D. M., Bevc, I., Duchesne, T., Rosenthal, J., Rossman, L., & Theodor, F. (2007). Comparison of adult offense prediction methods based on juvenile offense trajectories using cross-validation. *Advances and Applications in Statistics*, *7*, 1-46.
- Dumas, J. E. & Nilson, W. J. (2003). *Abnormal child and adolescent psychology*. NY: Allyn and Bacon.
- Johnson, L. M., Simons, R. L., & Conger, D. D. (2004). Criminal justice system involvement and continuity of youth crime. *Youth & Society*, *36*, 3-29.
- Jones, B. L., Nagin D. S., & Roeder, K. (2001). A SAS Procedure based on mixed models for estimating developmental trajectories. *Sociological Methods & Research*, *29*, 374-393.
- Masten A. S. & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist*, *53*, 205-220.
- Ministry of the Solicitor General and Correctional Services. (1995). *Statistical Reporting System User Manual*. Statistical Services, Correctional Services Division, North Bay, Ontario.
- Nagin, D. S. (2005). *Group-based modeling of development*. Cambridge, MA: Harvard Press.
- Sullivan, C. J., McGloin, J. M., Pratt, T., Piquero, A. R. (2006). Rethinking the "norm" of offender generality: Investigating specialization in the short-term. *Criminology*, *44*, 199-233.
- Thornberry, T. P. (2005). Explaining multiple patterns of offending across the life course across generations. *Annals of the American Academy of Political and Social Science*, *602*, 156-195.

Criminal Attitudes in Young Female Offenders: A Psychometric Evaluation

Leigh Greiner
Carleton University
E-mail: lgreiner@connect.carleton.ca

Shelley Brown
Carleton University

Vicki Jennings
Eastern Ontario Youth Justice Agency

As the need for research on female offenders is being recognized, there continues to be a significant gap between what is known about female offenders in comparison to their male counterparts. Historically, girls have received little attention within criminology (Hubbard & Pratt, 2002; Simourd & Andrews, 1994); this is logical as males account for significantly more crime than females. However, due to the dearth of research on youthful female offenders, there is little empirical evidence to help inform theory and practice.

This knowledge gap is most evident in the field of risk assessment. The level of risk an offender poses ascertains not only the probability of future criminal behaviour, but also assists in determining the level of intervention and the intensity of services provided to manage that risk (Hoge, 2002).

Several studies have demonstrated that both females and males have similar risk factors. Andrews and Bonta (2003) identified four primary risk factors equally important for both genders including criminal attitudes, antisocial peers, antisocial personality, and criminal history. All of these have been shown to be significant factors in the prediction of criminal re-offending for both female and male youthful offenders (see e.g., Simourd & Andrews, 1994; Hubbard & Pratt, 2002). However this study was only concerned with the construct of criminal attitudes.

Traditionally, models of delinquency often assumed that criminal attitudes were the result of having some other risk factor, such as antisocial peers (Hoge, Andrews, & Leschied, 1994). This has led to some misleading results from studies that have combined criminal attitude variables with other types of factors (see, e.g., Simourd & Andrews, 1994). Despite this limitation in the research, a number of studies have demonstrated

the importance of assessing criminal attitudes in determining the level of risk a young offender poses to re-offend (Engels, Luijpers, Landsheer, & Meeus, 2004; Hoge, Andrews, & Leschied, 1994;). Although this notion has been labelled with several different terminologies, including antisocial attitudes (Andrews & Bonta, 2003), the general concept remains the same. The construct of criminal attitudes, encompassing criminally orientated attitudes, rationalizations and beliefs (Simourd & Olver, 2002) has been empirically demonstrated to have high validity in the field of risk assessment.

Due to the similarities in risk factors, existing assessment approaches that were originally developed and validated on males have been arbitrarily applied to females. Although, both genders have been shown to have similar risk factors, little research has examined whether or not instruments developed on samples of males are equally valid for females. Notably, research examining the reliability and validity of criminal attitude measures used with female offender samples is scarce.

Accordingly, the primary objective of this study was to assess the reliability, and the convergent and predictive validity of four gender-neutral criminal attitude measures in a sample of youthful female offenders ($n = 134$) and male offenders ($n=133$). The measures included the: 1) Criminal Sentiments Scale (CSS; Andrews & Wormith, 1984), 2) Pride in Delinquency Scale (PID; Shields & Whitehall, 1991), 3) Neutralization Scale (Shields & Whitehall, 1994), and the 4) criminal attitude component of the Young Offender Level of Service Inventory (YO-LSI; Shields & Simourd, 1991). The 3-item criminal attitude component of the YO-LSI was comprised of item 69 (supportive of delinquency), item 70 (poor attitude towards sentence) and item 73

Criminal Attitudes in Female Offenders

(intends to continue crime). Although not an independent measure of criminal attitudes, the YO-LSI total score was also evaluated. Each of these instruments measures one of the variables (i.e., criminal sentiments, neutralization) that together form attitudes, values, and beliefs that are supportive of crime. The Young Children's Social Desirability Scale Revised (YCSDS; Ford & Rubin, 1970) was also included to determine if socially desirable responding needed to be factored into the analyses.

Method

Participants

Participants were youthful offenders who took part in the programming offered at the Eastern Ontario Youth Justice Agency (EOYJA) in Ottawa, Ontario between 1989 and 2005. The original sample was comprised of 134 females and 133 males however the sample size fluctuated for each of the instruments as a result of missing data. The females ranged in age from 13 to 18 ($M = 15.11$, $SD = .94$) and had an average of 9.28 years of education. The males ranged in age from 12 to 19 ($M = 14.95$, $SD = 1.12$) and had an average of 9.10 years of education. A t-test yielded no significant differences in age between genders. All data were retrieved from archival files, electronic or hard copy.

Measures

As dependent measures, two program outcome variables were selected including program success (i.e., completion) and program termination due to re-offending. Re-offending was defined as any new charge and/or conviction which occurred while participating in the program warranting termination. Offences ranged from minor offences, such as a breach of probation, to more serious crimes, such as assault. The length of follow-up (i.e., time in program) ranged considerably, from less than one week to 195 weeks for males, and from one week to 146 weeks for females. Although not statistically significant, there were modest gender differences in the length of time spent in the program with males spending on average 45.0 weeks ($n = 91$, $SD = 34.5$) and females spending on average 37.4 weeks ($n = 95$, $SD = 25.3$). It is important to note that both of

these dependent measures were based upon program outcomes as recorded by the EOYJA and that official criminal records (with a longer follow-up) were not yet available at the time of this study.

Results

The psychometric properties of the aforementioned instruments were assessed including their reliability, and convergent and predictive validity. To establish reliability, a series of Cronbach's alpha were calculated. All of the criminal attitude measures demonstrated high reliability (see Table 1) when used on both female and male young offenders.

Table 1. Reliability: Cronbach's alpha (α) based on initial assessment

	Female $\alpha / (n)$	Male $\alpha / (n)$
YOLSI ^a	.85 / (134)	.89 / (133)
YO-LSI: Att ^b	.53 / (134)	.52 / (133)
CSS ^c	.93 / (91)	.95 / (85)
PID ^d	.78 / (104)	.82 / (102)
NEUT ^e	.92 / (106)	.89 / (105)
YCSDS ^f	.69 / (79)	.75 / (62)

Note. The n fluctuates as a result of missing data; ^a Young Offender Level of Service Inventory, ^b 3-item criminal attitude component of the YO-LSI comprised of item 69 (supportive of delinquency), item 70 (poor attitude towards sentence) and item 73 (intends to continue crime), ^c Criminal Sentiments Scale, ^d Pride in Delinquency Scale, ^e Neutralization Scale, ^f Social Desirability Scale.

To establish convergent validity, a series of Pearson product-moment correlations were calculated among the total scores for each instrument. All four criminal attitude measures were found to be highly correlated with one another for both female and male youthful offenders (see Table 2). Results from both the reliability and convergent validity analyses are

Table 2. Inter-scale correlations (*r*) among predictor variables

	Female				Male			
	PID	YO-LSI	CSS	YO-LSI Att. ¹	PID	YO-LSI	CSS	YO-LSI Att. ¹
NEUT	.37**	.21*	.48**	.27**	.34**	.33**	.39**	.40**
PID	-----	.41**	.51**	.39**	-----	.38**	.50**	.41**
YO-LSI	-----	-----	.46**	.50**	-----	-----	.48**	.49**
CSS	-----	-----	-----	.48**	-----	-----	-----	.40**

Note. No significant correlations were found between the Young Children’s Social Desirability Scale and any other scale. * *p* < .01. ** *p* < .001.

consistent with previous findings examining several of these same instruments (Shields & Simourd, 1991; Shields & Whitehall, 1994; Simourd & Van de Ven, 1999).

Lastly, the predictive validity of these instruments was assessed including both their ability to predict re-offending and to predict program completion (see Table 3). Overall, four of the five criminal attitude measures predicted program completion among the females. However, none the measures predicted re-offending among the females. Interestingly, no significant predictors of either program completion or re-offending emerged for the male subset.

Discussion

Having designed this study to assess whether gender-neutral criminal attitude measures (originally constructed using predominantly male samples) are equally valid for girls, a clear cut answer can not be given. Based upon these results, it appears that three of the four assessment instruments tested, specifically the Pride in Delinquency scale, Neutralization scale, and Young Offender Level of Service Inventory (including the 3-item criminal attitude

Table 3. Predictive validity

	Female		Male	
	Re-offended (<i>n</i>)	Completed Program (<i>n</i>)	Re-offended (<i>n</i>)	Completed Program (<i>n</i>)
YO-LSI	.18*(132)	-.24**(132)	.11 (130)	-.16 (130)
YO-LSI: Att. ¹	-.01 (133)	-.20* (133)	-.13 (130)	-.06 (130)
CSS	.03 (107)	-.18 (107)	-.08 (105)	-.02 (105)
NEUT	.01 (128)	-.18* (128)	-.07 (115)	-.13 (115)
PID	.10 (128)	-.23**(128)	-.04 (116)	-.07 (116)
YCSDS	-.02 (93)	.20 (93)	.18 (83)	.04 (83)

Note. The *n* fluctuates as a result of missing data; ¹ YO-LSI 3-item criminal attitude component; * *p* < .05. ** *p* < .01.

subcomponent) are moderately effective assessment tools for predicting program outcomes for youthful female offenders. Surprisingly, this is not true for the male sample in this study as none of the instruments were useful in predicting either program completion or re-offending.

In an attempt to explain these unexpected results, there were several limitations of this study. The first limitation, and most detrimental to assessing predictive validity, was the low base rate of re-offending for both girls (10%) and boys (16%). Due to this factor, the true predictive abilities of these instruments may have been underestimated. A follow-up study is underway, using official recidivism data, to determine the long-term predictive ability of the assessment battery. A second limitation to this study was missing data. Given the nature of archival data, many of the files were missing scores for one or more assessment battery and as a result, sample size fluctuated for each instrument. Although complete files may not have influenced the results, they may have allowed for a more accurate evaluation of the predictive validity of the instruments assessed.

Overall, this study found strong support for the reliability and convergent validity of these instruments in both genders. Interestingly, while none of the measures predicted re-offending among the boys, only the total YO-LSI score predicted re-offending among the girls. While the majority of the measures predicted program completion for the girls none the measures predicted program completion for the boys. An examination of the extant juvenile risk assessment literature reveals a significant gap between what is known to be valid for females in comparison to males. Although this study provides some evidence of convergent and predictive validity for the risk assessment battery examined when used on youthful female offenders, it is essential that further research be conducted to investigate the predictive ability of these measures in both genders in order to draw firm conclusions regarding their efficacy.

References

Andrews, D. A., & Bonta, J. (2003). *The Psychology of Criminal Conduct*, 3rd ed. Cincinnati: Anderson Publisher Ltd.

- Andrews, D. A., & Wormith, J. S. (1984). *Criminal sentiments and criminal behavior: A construct validation*. Ottawa, ON: Ministry of the Solicitor General of Canada.
- Engels, R., Luijpers, E., Landsheer, J., & Meeus, W. (2004). A longitudinal study of relations between attitudes and delinquent behaviour in adolescents. *Criminal Justice and Behaviour*, 31, 244-260.
- Ford, L. H., & Rubin, B. M. (1970). A social desirability questionnaire for young children. *Journal of Consulting and Clinical Psychology*, 35 (2), 195-204.
- Funk, S. J. (1999). Risk assessment for juveniles on probation: A focus on gender. *Criminal Justice and Behaviour*, 26, 44-68.
- Hoge, R. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behaviour*, 29, 380-396.
- Hoge, R., Andrews, D. A., & Leschied, A. W. (1994). Tests of three hypotheses regarding the predictors of delinquency. *Journal of Abnormal Psychology*, 22, 547-559.
- Hubbard D. J., & Pratt, T. C. (2002). A meta-analysis of the predictors of delinquency among girls. *Journal of Offender Rehabilitation*, 34, 1-13.
- Shields I. W., & Simourd, D. J. (1991). Predicting predatory behaviour in a population of incarcerated young offenders. *Criminal Justice and Behaviour*, 48, 180-194.
- Shields, I. W., & Whitehall, G. C. (1991, December). *The Pride in Delinquency Scale*. Paper presented at the eastern Ontario correctional psychologists' winter conference, Burritts Rapids, Canada.
- Shields, I. W., & Whitehall, G. C. (1994). Neutralization and delinquency among teenagers. *Criminal Justice and Behaviour*, 21 (2), 223-235.
- Simourd, D. J., & Andrews, D. A. (1994). Correlates of delinquency: A look at gender differences. *Forum on Corrections Research*, 6 (1), 26-31.
- Simourd, D. J., & Olver, M. E. (2002). The future of criminal attitudes research and practice. *Criminal Justice and Behavior. Special Issue: Risk and assessment in contemporary corrections*, 29(4), 427-446.
- Simourd, D. J., & Van de Ven, J. (1999). Assessment of criminal attitudes: Criterion-related validity to the criminal sentiments scale-modified and pride in delinquency scale. *Criminal Justice and Behavior*, 26, 90-126.

The Jesness Inventory-Revised (JI-R) As a Measure of Psychopathology in a Sample of Adjudicated Juvenile Offenders

Arthur Eisenbuch
Metropolitan College of New York
E-mail: vitali@ptd.net

Donna Martin
Private Practice

Michael R. Schneider
Northampton County Juvenile Probation Department

James Menapace
Northampton County Juvenile Probation Department

Pamela Reube
Northampton County Juvenile Probation Department

Every year in the United States more than 2,000,000 youth are arrested for juvenile delinquency (75% males, 25% females). Research has shown that a significant percentage of these delinquent youth, compared to nondelinquents, present with one or more co-occurring mental health problems. These disorders are in order of descending frequency: Conduct Disorder; Substance Abuse-Dependence; Attention-Deficit/Hyperactivity Disorder; Mood Disorders; Anxiety Disorders; Psychotic Disorders and Autistic Spectrum Disorder (Abram, Teplin & McClelland, 2003; Cocozza & Skowrya, 2000; Office of Juvenile Justice and Delinquency Prevention, 2006). Concomitantly, research studies have demonstrated the essential role screening/assessment plays in relation to the successful treatment, facility management and public safety for both juvenile and adult offenders (Eisenbuch & Martin, 2000; Grisso, T. 2005).

The Jesness Inventory-Revised (JI-R) is the most recent version of the Jesness Inventory, a long-standing, delinquency-specific psychometric instrument. The original and revised versions of the Inventory both have reportedly acceptable levels of reliability and validity (Jesness, 2003; Jesness & Wedge, 1984).

The Northampton County Juvenile Justice Center was opened in January 2001, after two years of planning and development by the Court and its staff. The Short Term Offender Program, aka, STOP, is a 29-day group home program, housed within the Juvenile Justice Center. It was created for male youth aged 10 to 18, who have been adjudicated delinquent in the Northampton County Juvenile Court. The purpose of the program is to provide graduated sanctions to hold these youth accountable for their actions. It is highly structured

and consequence oriented. The program focuses on the youth reestablishing their priorities and accessing their immediate goals. The short term programming includes seminars and groups that are learning tools aimed at addressing decision making, peer influence and providing alternative positive thinking to the belief that juveniles in trouble are victims. These seminars are conducted by facility staff, as well as contracted vendor agencies. Discipline is taught by having clear rules, which are enforced with strong and immediate consequences. A strict regime of physical training exercises and tasks are utilized to help instill structure and discipline.

Because the program is short term and consequence oriented, treatment services are limited to those that the youth was already receiving prior to entering the program. If during a juvenile's placement it becomes evident that he is in need of ongoing treatment, a recommendation for follow-up services will be given to the supervising Juvenile Probation Officer.

Educationally, juveniles in the program attend classes in the on grounds program, provided by the Colonial Intermediate Unit #20. Every effort is made to provide schoolwork from the juvenile's home school.

Residents of STOP maintain contact with their families through weekly phone calls and one personal visit, which occurs, approximately halfway through the program.

During 2006, thirty males were placed in the Short Term Offender Program. Fifteen of the juveniles were Caucasian; eight of the juveniles were African-American and seven of the juveniles were Hispanic.

Method

This study retrospectively examined 100 JI-R inventories completed by 100 adjudicated juvenile offenders as part of their screening/assessment process for consideration of admission to a residential treatment program, 94 males and 6 females, males for the STOP Program (mean age = 16.05 years), from 2002 to 2005. The obtained sample means were then compared with the Jesness published means obtained from a study of delinquent and nondelinquent youth.

Results

On the Social Maladjustment Scale (SM), this study's sample was significantly different than the delinquent group, in that they scored significantly lower; but they were also significantly different than the nondelinquent group, scoring higher. On this 62 item scale, this sample is similar to the delinquent group in that they display less than adequate or disturbed socialization, and they have little regard or need to use socially approved methods for meeting their personal or environmental needs. This sample differs significantly from the nondelinquent group that does not share this set of attitudes.

For the 38 item Values Orientation Scale (VO), this sample is significantly different than the delinquent group, scoring lower, but not different than the nondelinquent group. On this measure of a tendency to share opinions characteristic of the lower socioeconomic class, this sample was not similar to the delinquent group but was very similar to the nondelinquent group.

On the 29 item Immaturity Scale (Imm), the sample was significantly different than the delinquent's group. They did not tend to endorse self/other perceptions or attitudes that would be expected to be displayed by much younger persons. They were not significantly different than the nondelinquent group.

The 26 item Autism Scale (Au) measures the tendency to use distorted thinking and perceiving in regard to reality, bending it to one's own needs and desires. This sample did not differ significantly from the delinquent group but differed extremely from the non-delinquent group.

On the 25 item Alienation Scale (AI), which purports to measure distrust and estrangement in attitudes, especially in attitudes toward authority, the sample is not significantly different from the delinquent group but is significantly different from the nondelinquent group.

On the 32 item Manifest Aggression Scale (MA), the sample is significantly different from both the delinquent group and the nondelinquent group, scoring much lower than both. This scale reflects a measure of awareness of uncomfortable feelings, most specifically frustration and anger. It further measures discomfort in the presence of the control of such feelings and also the tendency to act on them.

The Withdrawal-depression Scale (Wd) consists of 22 items that indicate the extent of an individual's self and other dissatisfaction and a tendency to isolate. The sample was significantly different from the delinquent group and significantly different from the nondelinquent group.

The 20 item Social Anxiety Scale (SA) is intended to measure feelings of anxiety and emotional discomfort of which one is aware in interpersonal relations. This sample was significantly different than both the delinquent and nondelinquent groups, scoring lower than both.

The 14 item Repression Scale (Rep) is intended to reflect the exclusion from conscious awareness, feelings a person would be expected to be aware of, or the failure to give them a label. There were no significant differences between the sample and either the delinquent or the nondelinquent group.

The 19 item Denial Scale (Den) purports to indicate a reluctance to admit to or acknowledge unpleasant conditions or events in day to day life. There were no significant differences between the sample and either the delinquent or the nondelinquent group.

The Asocial Index (AI) is a weighted total measure that reflects a generalized disposition that includes a disrespect and disregard for social customs and rules when resolving social or personal problems. This sample was significantly different than the delinquent group, scoring lower. However, they were also significantly different than the nondelinquent group, scoring significantly higher.

Discussion

The results of this study indicated that in the aggregate, the JI-R's Asocial Index was clearly able to distinguish between groups of delinquent and nondelinquent youth. Equally the study's sample was found to be significantly more maladjusted than nondelinquent youth. These findings are consistent with recent research studies which have found that well-designed psychometric instruments based on offender self-reports tended to be unexpectedly reliable and valid, despite offenders historical reputation for dissimulation and manipulation (Kroner & Loza, 2001; Schretlen & Arkowitz, 1990).

The sample was found to be significantly more mature than the Jesness delinquent group. However, the sample was also unexpectedly found to be significantly less maladjusted, with a significantly lower Asocial Index and tendency toward manifest aggression than the delinquent group. These unexpected findings hypothetically may reflect the study's more "sophisticated" sample engaging in inadequately controlled for, explained or interpreted "impression management" and "socially desirable responding," in relation to such JI-R items as *the police are pretty dumb(T) or people who run things are usually against me (T)*. See Mills and Kroner (2005) for a comprehensive review of the complex issues related to the interaction of offender impression management, socially desirable responding, concurrent and predictive validity. The primary limitations of this study would appear to center around its retrospective design, utilizing an almost exclusively male sample drawn from a predominantly urban/suburban Northeastern area. Primary implications for further research would center around the utilization of a prospective multi-factor research design to examine the relationships between the hypothesized impression management, socially desirable responding and the current and predictive validity of the JI-R.

References

- Abram, K. M., Teplin, L., & McClelland, G. M. (2003). Comorbid psychiatric disorders in youth in juvenile detention. *Archives of General Psychiatry*, *60*, 1097-1108.
- Cocozza, J. J., & Skowrya, K. (2000). Youth with mental disorders: Issues and emerging responses. *Juvenile Justice*, *7*, 3-13.
- Eisenbuch, A., & Martin, D. (2000). *The Carlson Psychological survey as a Measure of Psychopathology in Serious and Violent Juvenile Offenders*. Paper presented at the annual meeting of the American College of Forensic Examiners International, New York, New York.
- Grisso, T. (2005). Why we need mental health screening and assessment in juvenile justice programs. In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in Juvenile Justice*, (pp.3-18). New York: Guilford.
- Jesness, C. F. (2002). *The Jesness Inventory-Revised: Technical Manual*. North Towanda, NY: Multi-Health Systems.
- Jesness, C. F., & Wedge, R. F. (1984). Validity of a revised Jesness Inventory I-Level Classification with delinquents. *Journal of Consulting and Counseling Psychology* *52-6*, 997-1010.
- Kroner, D. G., & Loza, W. (2001). Evidence for the efficacy of self-report in predicting violent and nonviolent criminal recidivism. *Journal of Interpersonal Violence*, *16*, 166-177.
- Mills, J. & Kroner, D. G. (2005). An investigation into the relationship between socially desirable responding and offender self-report. *Psychological Services*, *2*, 70-80.
- Office of Juvenile Justice and Delinquency Prevention (2006). *Psychiatric disorders of youth in detention*. Washington, DC: U.S. Department of Justice.
- Schretlen, D., & Arkowitz, H. (1990). A psychological test battery to detect prison inmates who fake insanity or mental retardation. *Behavioral Sciences and the Law*, *8*, 75-84.

What's Not Working within 'What Works': Executive Cognitive Functioning Capacity of First Time Offenders, Return Offenders, and Controls

Erin H. Ross, Jenny Neil, & Peter, N.S. Hoaken
E-mail: edezell@uwo.ca

There is now a compelling literature demonstrating that imprisonment alone is insufficient to deter offenders from future criminal activities. Accordingly, Correctional Services Canada has, over the past several decades, implemented various remediation programs in an attempt to reduce recidivism rates. Reports of recidivism reduction in relation to remediation programs have, unfortunately, been somewhat inconsistent, with some studies reporting that remediation programs had no effect, but with others reporting reductions in recidivism as high as 58%. A variety of reasons have been presented to account for this variability, such as the type of offender sub-groups studied (e.g., sexual offenders, property offenders), and motivation level of the offender in question. However, one factor that may be related to propensity for re-offending that has received little attention is the constellation of cognitive abilities referred to as the "executive cognitive functions" (ECF). Executive function refers to higher order cognitive abilities involved in goal-directed behaviour, including cognitive flexibility, strategy formation, response monitoring, working memory, and inhibition. Generally speaking, these abilities represent behaviours relevant to effective problem-solving and decision-making. Extant research has consistently shown various ECF deficits in incarcerated offenders, both adult and adolescent, with some researchers linking these deficits to recidivism. Unfortunately, little is known about whether existing remediation programs have any impact on ameliorating executive dysfunction, and none appear designed specifically to do so.

The purpose of the current study was to investigate the ECF abilities of incarcerated males from two medium security federal institutions within Canada, in comparison to non-incarcerated male community controls. More specifically, the study sought to elucidate to what extent ECF abilities differentiated repeat offenders (offenders

with at least one prior federal prison sentence) from both first time offenders (offenders completing their first federal sentence) and controls. As previous research from our lab has demonstrated a link between ECF and criminality (Hoaken, 2006; Hoaken, Allaby & Earle, 2007), it was hypothesized that both offender groups would display ECF performance inferior to that of controls. We also hypothesized that repeat offenders would display more ECF impairments than first time offenders. Additionally, the current study also investigated whether a relationship existed between exposure to correctional programming and ECF. It was hypothesized that the number of remediation programs completed by offenders, most notably those including a 'problem-solving' skills training component, would have contributed positively to ECF abilities.

Of the 140 male participants, 95 were offenders incarcerated at either Springhill Institution (Springhill, Nova Scotia) or Fenbrook Institution (Gravenhurst, Ontario). Controls consisted of 45 males recruited from the community of London, Ontario. Participants completed four behavioural measures of ECF, including the Conditioned Nonspatial-Association Task, the Iowa Gambling Task, the Wisconsin Card Sorting Task, and Go/No-Go Task. In combination, these tests were thought to effectively and holistically assess the various elements of executive function. Offenders were recruited by posters, individual invitation (sent through institutional mailing system), and recruitment events. ECF measures were completed on a laptop computer in a private testing area within the institution. Demographic information and self-report of remediation program involvement was also obtained during the testing session. A review of institutional records provided additional information regarding prior offences, demographics, and program involvement. In order to avoid bias, file reviews were completed subsequent to testing. Controls,

who responded to a local advertisement, completed the same measures, including assessment of community program involvement (e.g., substance abuse programs), at the University of Western Ontario.

A multivariate analysis of variance, with post-hoc tests, revealed that return offenders performed significantly worse than both first time offenders and controls in the areas of strategy formation, working memory, response monitoring, and impulsivity. Surprisingly, first time offenders were not statistically different than controls on any of these variables. With respect to programming, it was not surprising that return offenders had completed more programs than first time offenders, who in turn had completed more programs than controls. However, contrary to the hypothesis, number of completed programs was positively correlated with poor performance on some ECF measures, indicating that being engaged in more programming was actually associated with poorer executive functioning.

Poorer performance by return offenders in areas of strategy formation, working memory, response monitoring, and impulsivity likely reflects a global problem-solving impairment in these individuals. These findings are quite relevant to recidivism. If an offender is not equipped with appropriate problem-solving abilities, navigating through situations with potential criminal components (e.g., solving financial strain; interpreting ambiguous social situations) will inevitably be more problematic, and could lead to further poorly-planned and impulsive, and indeed potentially criminal, activities. The public policy implications of these findings are numerous. For example, understanding the distinctions between first timer offenders and return offenders may further counsel against a 'one size fits all' approach to remediation; instead, program reformers could capitalize on the unique strengths evidenced by first time offenders. In addition, understanding the distinctions in executive abilities between subgroups of offenders may aid in the accurate classification for risk of recidivism. Furthermore, the results suggest that correctional remediation programming may be lacking a specialized element of ECF training. Other types of remediation which focus on attenuating executive dysfunction, such as those involved in

rehabilitating individuals with head injury or schizophrenia, may inform future research.

References

- Hoaken, P.N.S., Allaby, D., & Earle, J. (2007). Executive cognitive functioning and the recognition of facial expressions of emotion in incarcerated violent offenders, non-violent offenders, and controls. *Aggressive Behavior, 33*, 1-10.
- Hoaken, P.N.S. (2006). Investigations of social information processing and executive cognitive function in adult male offenders and community controls. In: Österman, K., & Björkqvist, K. (Eds.). *Contemporary research on aggression*. Åbo Akademi University, Finland.

Autonomous Classification of Sexual Offenders from Eye Pattern Behavior

Sylvain Chartier
University of Ottawa
Centre de recherche de l'Institut Philippe Pinel de
Montréal
E-mail : sylvain.chartier@uottawa.ca

Patrice Renaud
Centre de recherche de l'Institut Philippe Pinel de
Montréal
Université du Québec en Outaouais
Hexagram: Institut de recherche / création en arts et
technologies médiatiques

Sébastien Caron
University of Ottawa

Violent behaviours, like the ones associated with sexual crimes, are major social concerns and the cause of considerable psychological distress. For this reason, it is important to thoroughly assess and diagnose these behaviours. Current assessment of sexual offenders often involve the use of penile plethysmography (PPG). Since its introduction by Freund (1963), PPG has been the target of much criticism on both ethical and methodological grounds (Kalmus & Beech, 2005; Laws & Gress, 2004; Laws & Marshall, 2003; Marshall & Fernandez, 2003). A large part of the method's test-retest reliability and discriminating validity problems arise from PPG's vulnerability to strategies used by sex offenders to voluntarily control their penile response in order to fake a non-deviant preference profile (Quinsey & Chaplin, 1988; Seto & Barbaree, 1996). Participants generally manage to lower their scores through the use of aversive or anxiogenic thoughts and images, that is, by diverting their attention from the sexual stimuli to which they are exposed.

The use of immersive video-oculography makes it possible to prevent this problem (Duchowski et al., 2002; Renaud et al., 2002, 2003, 2005). This technique allows researchers to observe a crucial part of the participants' experience - that is, their global visual perspective coupled with the exact position of their gaze as it scans and dwells upon the various simulated objects in virtual reality, particularly, the different corporal segments of a virtual sexual character. The visual scanning of virtual stimuli produces a highly complex geometry from which we can analyze certain recurrences patterns, including transitions from one virtual area to another (Renaud et al., 2002, 2003). These transition patterns constitute visual

routines that are the dominant characteristic of the automatic cognitive processes implicated in the appraisal of the emotional and sexual signification of the stimuli. They present the major advantage of not being as transparent to the subject as PPG measurement. The analyses of oculomotor patterns alone may possess sufficient discriminating power to determine deviant profiles without needing to resort to PPG.

How oculomotor patterns should be classified remains an open question. Several classification algorithms exist, but not all of them work with nonlinear and dynamic data. Support vector machine (SVM) present a good choice because it can cope with problems that are not linearly separable (Haykin, 1999) and it is easily implementable.

The remainder of the paper is divided as follows: a) the avatars used in the virtual environment, b) the eye-tracker recording system, c) the invariant features extracted from the eye movement time series, d) the Support vector machine used for classification, e) general methodological issues, f) results, and g) conclusions.

Avatars in a Virtual Environment

The major asset of using synthetic characters (avatars) is that they guards against the victimization of real models (Renaud et al. 2004). It is necessary to have avatars depicting naked characters of both genders and of clinically significant age phases to prompt sexual attraction and arousal. These avatars have been developed to represent the required sexual properties to assess sexual preference (Renaud et al. 2005). Each avatar was assigned two virtual measurement points (VMP) placed over the head and pubic

areas for the analysis of eye-movements via an eye-tracker.

Gaze Behaviour Measurement in Virtual Environment

Eye-tracking has been used in a wide variety of applications as a window to cognitive process (e.g., Albert et al., 2005; Duchowski et al., 2002; Land & Lee, 1994; Satava, 1995). The eye-tracker system relies on the corneal reflection of an infrared source that is measured relative to the pupil center location. From this information a 2 degrees-of-freedom computation of the point of regard (POR) can be performed (Figure 1). From that information, gaze radial deviation (GRAD) between a given virtual measurement point and the point of regard is computed (Renaud et al. 2003).

From those data, it can be possible to determine the number of eye fixations, saccades, average saccade speed, etc. Figure 2 illustrates an example of 1 minute GRAD. From this time series, features can then be extracted before it is used for classification by a SVM.

Extracted Features

The GRAD measurement for each subject is composed of 3600 data points. Usually, mean time fixation and saccade variability are used. However, such variables removed the temporal aspect present in the data (Renaud et al. 2005). Therefore, in this study we focussed on nonlinear time series analysis tools: the dimension correlation and the largest Lyapunov exponent (Spratt, 2003), which are among the most popular features for representing a given GRAD time series.

Dimension Correlation. Being able to distinguish random process from a chaotic (but not random) system has received much attention lately. The most widely know nonlinear dynamic statistic is the correlation dimension (CD). The correlation dimension is a measure of self-similarity. This measure is consistent with non-fractal objects, in which lines have a dimension of one while planes have a dimension of two, etc. CD expresses invariance amidst nonlinear dynamical processes by shedding light on the self-similarity of the underlying attractors' geometry (Spratt, 2003).

Largest Lyapunov Exponent. One of the most popularized properties of chaotic systems is their sensitive dependence on initial condition. The largest Lyapunov exponent (LLE) is a quantity that characterizes the rate of separation of infinitesimally close trajectories. LLE determines the predictability of a dynamical system. A positive LLE is usually taken as an indication that the system is chaotic (Spratt, 2003).

Support Vector Machine

Support Vector Machines (SVMs) is a technique for classification and regression tasks (Haykin, 1999). In a binary classification task like the one in this study, the aim is to find an optimal separating hyperplane between sexual offenders and non offenders. Figure 3 shows two examples of optimal separating hyperplanes (OSH) that generates the maximum margin (dashed line) between the two arbitrary data sets. SVM finds this OSH by maximizing the margin between the classes. As shown by the Figure 3, SVM offer good flexibility to establish the classification function; SVM can classify both linearly separable data and non-linearly separable data. To accomplish the classification, SVM first transforms input data into a higher dimensional space by means of a kernel function and then constructs a linear OSH between the two classes in the transformed space. Those data vectors nearest to the constructed line in the transformed space are called the support vectors. There are several different kernel functions; for this study, we used the most popular, the radial-basis function (RBF).

Method

Ten participants were selected for the study. They were 6 sexual offenders and 4 non offenders. For each participant there were four conditions that represent the different avatars (girl, woman, boy and man). Each participant were immerse in the virtual environment for 1 minute per condition. Raw data were collected by an ASL 504 series eye-tracker combined into a V8 Virtual Research head mounted display. The single eye-tracker returns 2 degrees of freedom, i.e. variations in x and y plane, at 60 Hz with a margin of error ± 0.5 degrees. From the data file generated with ASL software, POR angular deviation was computed

Table 1. Input variables (correlation dimension and largest Lyapunov exponent) used to train the SVM for each participant as a function of each virtual character.

Participant	Female adult		Female children		Male adult		Male children	
	CD	LLE	CD	LLE	CD	LLE	CD	LLE
Control 1	1.7	0.0154	1.5	0.0189	1.64	0.0172	1.7	0.018
Control 2	1.58	0.015	1.76	0.0096	1.67	0.0108	1.68	0.011
Control 3	1.93	0.019	2.05	0.0153	1.61	0.091	1.7	0.0198
Control 4	1.8	0.0202	2.16	0.0085	2.13	0.011	1.77	0.0124
Offender 1	1.67	0.0155	1.61	0.02	1.58	0.017	1.42	0.0159
Offender 2	1.56	0.0176	1.42	0.0149	1.29	0.0215	1.58	0.0181
Offender 3	1.78	0.0141	1.65	0.0111	1.71	0.0113	1.72	0.0141
Offender 4	1.81	0.014	1.73	0.0193	1.78	0.0248	1.78	0.0133

between the POR and a given target, which was the genital area. Consequently, for each condition a vector composed of 3600 GRAD was obtained. From the signals, two features were extracted. The first feature was the dimension correlation and the second feature was the largest Lyapunov exponent. Consequently, for each participant, eight data points were obtained (4 variables \times 2 features), giving a total of eight patterns (four offenders and four non offenders) (see Table 1). To train the SVM, four sexual offenders and four non offenders were selected. The remained two sexual offenders were used to test the network generalization capacity.

Results

Table 2 shows the classification result. The SVM was able to find a function that allows the classification of sexual offenders from non offenders. This function enables the network to correctly classify each participant using solely the correlation dimension and largest Lyapunov exponent. Generalization capacity of the network was tested with 2 new sexual offender patterns (offender G1 and G2). In each case, those generalization patterns were correctly assigned to the offender class. Therefore, even though the network has never seen the given input it nevertheless was able to assign it to the correct class.

Conclusion

Table 2. SVM classification results according to the training patterns and its generalization to new inputs.

	Inputs (CD, LLE)	Classification	
		Desired	SVM
Training Patterns	Control 1	-1	-1
	Control 2	-1	-1
	Control 3	-1	-1
	Control 4	-1	-1
	Offender 1	1	1
	Offender 2	1	1
	Offender 3	1	1
	Offender 4	1	1
Generalization Patterns	Offender G1	-	1
	Offender G2	-	1

It has been shown that oculomotor measurements can be used for automatic classification of sexual offenders from non offenders. The network classification is based on inputs coming from angle deviations between the participant gaze and the genital target area. Once trained, the network is able to generalize its classification function to new participants. This offline classification demonstrated perfect performance.

Before the system can be used for clinical classification, further investigation is required. First, the network must be trained with many more participants, both sexual offenders and non offenders. Given the small sample of data in this study, results are preliminary at this point. In addition, the avatars need more realism. A new generation of avatars is being developed, which will be used in futures studies. Finally, the SVM should be modified in way that it can be used for real time classification. This online implementation would allow feedback in helping to better understand deviant behavior. In this way, this new tool would not only help in diagnostics and risk assessment, but it could also be an active ingredient in behaviour modification therapy.

References

- Albert, G., Renaud, P., Chartier, S., Renaud, L., Sauvé, L. & Bouchard, S. (2005). Scene Perception, Gaze Behavior and Perceptual Learning in Virtual Environments, *CyberPsychology and Behavior*, 8, 592-600.
- Duchowski, A. T., Medlin, E. Courmia, N. Murphy, H. Gramopadhye, A. Nair, S. Vorah, J. & Melloy, B. (2002). 3D Eye Movement Analysis, *Behavior Research Methods, Instruments, & Computers (BRMIC)*, 34, 573-591.
- Freund, K. (1963). A laboratory method for diagnosing predominance of homo- and hetero-erotic interest in the male. *Behaviour Research and Therapy*, 1, 85-93.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*, Upper Saddle River, NJ: Prentice Hall.
- Kalmus, E. & Beech, A. R. (2005). Forensic assessment of sexual interest: A review. *Aggression and Violent Behavior*, 10, 193-217.
- Laws, D.R., & Gress, C.L.Z. (2004). Seeing things differently: The viewing time alternative to penile plethysmography. *Legal and criminological psychology*, 9, 183-196.
- Laws D. R. & Marshall, W. L. (2003). A brief history of behavioral and cognitive-behavioral approaches to sexual offenders, Part 1, Early developments. *Sexual Abuse: A Journal of Research and Treatment*, 15, 75-92.
- Marshall, W. L., & Fernandez, Y. M. (2003). Sexual preferences: Are they useful in the assessment and treatment of sexual offenders? *Aggression and Violent Behavior*, 8, 131-143.
- Sprott, J. L. (2003). *Chaos and Time-Series Analysis*, New York, NJ: Oxford University Press.
- Land, M. F. & Lee, D. N. (1994). Where we look when we steer, *Nature*, 369, 742-744.
- Quinsey, V. L. & Chaplin, T.C. (1988). Preventing faking in phallometric assessments of sexual preference. In: R. A. Prentky & V.L. Quinsey (Eds.). *Human Sexual Aggression: Current Perspectives*. New York: New York Academy of Sciences, 49-58.
- Renaud, P., Bouchard, S. & Proulx, R. (2002). Behavioral avoidance dynamics in the presence of a virtual spider. *IEEE Transactions in Information Technology and Biomedicine*, 6(3), 235-243.
- Renaud, P., Albert, G. Sauvé, L., Renaud, L., Décarie, J., & Bouchard, S. (2004). Assessing perceptual learning dynamics during visual search in virtual immersion using eye-tracking technologies. *Proceeding of the IADIS International Conference on Cognition and Exploratory Learning in Digital Age*, Lisbon, Portugal: IADIS Press, 413-418.
- Renaud, P., Décarie, J., Gourd, S.-P., Paquin, L.-C., & Bouchard, S. (2003). Eye-tracking in immersive environments: A general methodology to analyze affordance-based interactions from oculomotor dynamics. *CyberPsychology and Behavior*, 6, 519-526.
- Renaud, R., Proulx, J., Rouleau, J.-L., Bouchard, S., Madrigano, G., Bradford, J. & Fedoroff, P. (2005). The recording of observational behaviors in virtual immersion: A new clinical tool to address the problem of sexual preferences with paraphiliacs. *Annual review of Cybertherapy and Telemedicine*, 3, 85-92.
- Satava, R. M. (1995). Medical applications of virtual reality, *Journal of medical Systems*, 19, 275-280.
- Seto, M.C. & Barbaree, H. E. (1996). Sexual aggression as antisocial behavior: A developmental model. In D. Stoff, J. Brieling & J. D. Maser (Eds.), *Handbook of antisocial behavior* (pp. 524-533). New York: Wiley.

Is Sexual Assault Committed by Hebephiles Different from that Committed by Pedophiles and Rapists?

Sophie Desjardins
Université du Québec à Trois-Rivières
E-mail: sophie.desjardins@uqtr.ca

Luc Granger
Université de Montréal

Given that sex offenders are recognized as a heterogeneous group (Erickson, Luxenberg, Walbek & Seely, 1987; Hall, Maiuro, Vitaliano & Proctor, 1986; Levin & Stava, 1987; Looman, Gauthier & Boer, 2001), a number of typologies have been proposed over the years in an effort to increase our understanding of the group and have a positive impact on both the prevention of sex crimes and the reduction of recidivism rates.

Although victim age is a widely accepted preliminary discriminant variable in typological systems (Bard et al., 1987), the fact remains that sex offenders who target adolescents, referred to as hebephiles, constitute a group generally absent from studies on sexual aggression.

According to recent estimates, however, post-pubertal minors represent a large proportion of victims of sexual aggression, i.e., over one third of sexual prey in Canada (Motiuk & Belcourt, 1996) and over one quarter in Quebec (Proulx et al., 1999). The number of studies on hebephiles does not currently reflect these figures. Fewer than 15 scientific articles, some of which include only members of the clergy in their sample (e.g., Camargo, 1997), have been written on this group.

According to some authors (Laws, 1989; Pithers, Kashima, Cumming & Beal, 1988), the development of effective strategies is largely dependent on an understanding of the offenders' modus operandi. Repucci and Haugaard (1989), for example, recommend focusing on the behaviour patterns of sex offenders and incorporating the resulting findings into prevention programs. Kalichman (1991) argues that different offender groups should receive different treatment. If it turns out that hebephiles act differently from pedophiles or rapists, the most appropriate interventions should be reconsidered

for them. If not, they can justifiably be included in the pedophile or rapist group without distinction.

The purpose of this study is to contribute to the meagre corpus of knowledge currently available on hebephilia by comparing the modus operandi of the three sex-offender subgroups formed on the basis of victim age. Specifically, the research attempts to determine whether hebephiles are, in fact, a separate group from pedophiles and rapists in terms of their modus operandi.

Method

Participants

Data relating to a total of 134 sex offenders, consisting of 49 pedophiles, 44 hebephiles and 41 rapists incarcerated in a correctional centre and serving a sentence of two years or more, were taken from a database developed by a research group within the Correctional Service of Canada. The participants in question were grouped according to victim age. Thus, pedophiles were defined in this study as offenders convicted of sexual assault on children under 13 years of age, hebephiles as offenders convicted of sexual assault on adolescents between 13 and 17, and rapists as individuals convicted of sexual assault on women over 17.

The use of the terms "pedophile," "hebephile" and "rapist" is based not on the sexual preference of these groups but solely on the behaviours exhibited. Consequently, we use the terms "pedophile," "hebephile," and "rapist" to refer to offenders charged with and convicted of having had sexual contact with prepubescent children, pubescent adolescents and non-consenting adult females, respectively.

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 162-165. Public Safety Canada.

The authors wish to thank the Fonds québécois de recherche sur la société and la culture (FQRSC) and the Correctional Service of Canada for their support.

All the offenders selected to participate in the study denied having assaulted a victim in an age category other than that in which they were included and denied having committed incest with or having killed a victim. Furthermore, none of the offenders had been convicted of any of the offences in the other categories. At the time of assessment, none of the offenders showed signs of serious mental or organic disorders.

Statistical Analyses

Fifteen variables relating to behaviours and fantasies prior to the sexual assault (alcohol consumption and drug use, presence of fantasies in the 48 hours prior to the sexual assault), victim characteristics (gender, level of acquaintance), how the sexual assault was carried out (accomplice involved, degree of force used, victim humiliation, use of a weapon, touching of the genitals, fondling, rubbing and interfemoral penetration, coitus, masturbation on and by the victim) and level of acknowledgement of sex problems were simultaneously entered in a discriminant function analysis to predict membership in the three offender groups. We chose simultaneous data entry because we had no reason to give certain predictive variables priority over others and because this procedure minimized the likelihood of chance being a factor (Stevens, 1996).

Following the procedure recommended by Dillon and Westin (1982), category variables involving more than two levels were subjected to dummy binary variable transformation before being included in the analysis. Degree of force used (none, minimal, more than necessary), acknowledgement of sex problems (absent, partial, present) and level of acquaintance with the victim (unknown, previously seen, previously spoken to) were the three variables transformed in this manner.

As recommended (Tabachnick & Fidell, 2001), the discriminant function analysis considered the probabilities of belonging to each group, which were 36.6% for pedophiles, 32.8% for hebephiles and 30.6% for rapists.

Results

Two significant discriminant functions were identified from the analysis, the first with a Wilks' lambda of .33 and a $\chi^2(36) = 137.83, p < .001$, and the second with a Wilks' lambda of .74 and a $\chi^2(17) = 37.50, p < .001$. The first and second functions represented 78% and 22% respectively of the between-group variance. The first discriminant function achieved maximal separation between pedophiles and the other two offender groups, and the second distinguished hebephiles from rapists, with pedophiles being divided between the two groups.

Table I provides a detailed illustration of the modus operandi of the three offender groups. The analyses show that pedophiles are less likely than hebephiles or rapists to use force. When they choose to do so, they have a greater tendency to use only the degree of force necessary to commit their offences, rather than excessive force. Pedophiles are also more likely than hebephiles and rapists to choose male victims, less likely to humiliate their victims, more likely to acknowledge their sex problems and less likely to use a weapon. They are also more likely than the other groups to fondle their victims, rub against them and move their penis back and forth between the victim's thighs.

The five predictive variables that best distinguish hebephiles from rapists on the second discriminant function are masturbation of the victim, victim gender, masturbation by the victim, interest in humiliating the victim and involvement of an accomplice. Hebephiles are more likely than rapists to sexually assault male prey, to masturbate their victims, to ask their victims to masturbate them and to involve an accomplice in the commission of their offence. They are also less interested in humiliating their male and female adolescent victims.

As shown in Table II, discriminant function analysis was successfully used to classify participants in their respective groups, as 76.1% of the offenders were classified correctly, by comparison with 33% who would have been classified correctly by chance alone.

Like the rapists, the incorrectly classified pedophiles were found primarily in the hebephile category. The incorrectly classified hebephiles

Tableau I. Profile of the modus operandi of the three offender groups

	P (%) ^b	H (%) ^b	R (%) ^b
No force used ^a	76	23	14
More force used than necessary ^a	2	37	52
Male victim	54	32	0
Victim humiliation	2	16	42
Sex problems not acknowledged ^a	28	63	70
Sex problems acknowledged ^a	52	18	18
Fondling, rubbing, interfemoral penetration	39	10	4
Weapon use	4	23	38
Alcohol consumption prior to offence	18	48	54
Coitus	37	78	71
Fantasies 48 hours prior to offence	52	22	23
Drug use prior to offence	12	35	40
Had previously spoken to victim ^a	74	55	44
Victim unknown to offender ^a	18	33	44
Masturbation of victim	30	28	2
Masturbation by victim	46	40	18
Accomplice involved	2	22	14
Touching of victim's genitals	86	80	66

Note. ^a Dummy Variable. ^b P = pedophile, H = hebephile, R = rapist.

Tableau II. Classification Matrix

		Predicted Group Membership			
		Pedophiles N (%)	Hebephiles N (%)	Rapists N (%)	Total N (%)
Initial Groups	Pedophiles N (%)	39 (79.6)	7 (14.3)	3 (6.1)	49 (100)
	Hebephiles N (%)	5 (11.4)	33 (75.0)	6 (13.6)	44 (100)
	Rapists N (%)	3 (7.3)	8 (19.5)	30 (73.2)	41 (100)

were divided almost equally between the pedophile and rapist groups.

Discussion

The results obtained support the hypothesis that hebephiles, like pedophiles and rapists, are a separate group in terms of their modus operandi. Hebephiles may be distinguished from pedophiles primarily by the fact that they are more likely to use force to obtain sexual gratification from their victims, to use excessive force and to use a weapon in the commission of the offence. As a group, hebephiles are more likely than pedophiles to choose female adolescents as their sexual prey and to be interested in humiliating them. Hebephiles are also differentiated from pedophiles by the fact that they are less likely to fondle their victims and rub against them. Hebephiles are distinguished from rapists by the fact that they choose a substantially higher proportion of male victims and practise mutual masturbation with them. They are also less likely than rapists to humiliate their prey and more likely to involve an accomplice in the assault.

Although the presence of fantasies, alcohol consumption and drug use prior to the assault, touching, coitus and the level of acquaintance with the victim all help to distinguish between the three sex-offender groups, those variables are secondary to the ones just mentioned.

This research highlights the advisability of considering hebephiles as a separate group in future studies or, at least, of specifying that the samples studied included this category of sex offender. On their own, the results obtained do not allow us to make recommendations concerning the choice of therapeutic strategies for the three groups of sex offenders studied, although they do provide a strong incentive to continue studies in this area.

References

- Bard, L. A., Carter, D. L., Cerce, D. D., Knight, R. A., Rosenberg, R., & Schneider, B. (1987). A descriptive study of rapists and child molesters: Developmental, clinical, and criminal characteristics. *Behavioral Sciences & the Law*, 5, 203-220.
- Camargo, R. (1997). Factor, cluster, and discriminant analyses of data on sexually active clergy : the molesters of youth identified. *American Journal of Forensic Psychology*, 15, 5-24.
- Dillon, W. R. & Westin, S. (1982). Scoring frequency data for discriminant analysis : Perhaps discrete procedures can be avoided. *Journal of Marketing Research*, 19, 44-56.
- Erickson, W., Luxenberg, M., Walbek, N., & Seely, R. (1987). Frequency of MMPI two-point code types among sex offenders. *Journal of Consulting and Clinical Psychology*, 55, 566-570.
- Hall, G. N., Maiuro, R., Vitaliano, P., & Proctor, W. (1986). The utility of the MMPI with men who have sexually assaulted children. *Journal of Clinical and Consulting Psychology*, 54, 493-496.
- Kalichman, S. C. (1991). Psychopathology and personality characteristics of criminal sexual offenders as a function of victim age. *Archives of Sexual Behavior*, 20, 187-197.
- Laws, D. R. (1989). *Relapse prevention with sex offenders*. New York : Guilford.
- Levin, S. & Stava, L. (1987). Personality characteristics of sex offenders : A review. *Archives of Sexual Behavior*, 16, 57-79.
- Looman, J., Gauthier, C., & Boer, D. (2001). Replication of the Massachusetts Treatment Center child molester typology in a Canadian sample. *Journal of Interpersonal Violence*, 16, 753-767.
- Motiuk, L. & Belcourt, R. (1996). Profil des délinquants sexuels sous responsabilité fédérale au Canada. *FORUM – Recherche sur l'Actualité Correctionnelle*, 8, 3-8.
- Pithers, W. D., Kashima, K. M., Cumming, G. F., & Beal, L. S. (1988). Relapse prevention : A method of enhancing maintenance of change in sex offenders. In A. C. Salter (Ed.), *Treating child sex offenders and victims*. Newbury Park, CA : Sage.
- Proulx, J., Granger, L., Ouimet, M., Guay, J.-P., McKibben, A., St-Yves, M., Bigras, J., Perreault, C., Brien, T., & Pellerin, B. (1999). Profil descriptif d'un échantillon de délinquants sexuels incarcérés. *FORUM – Recherche sur l'Actualité Correctionnelle*, 11, 11-14.
- Repucci, N. D. & Haugaard, J. J. (1989). Prevention of child sexual abuse : Myth or reality. *American Psychologist*, 44, 1266-1275.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ : Lawrence Erlbaum.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA : Allyn & Bacon.

Personal and Criminal Characteristics Distinguishing Pedophiles, Hebephiles, and Rapists

Sophie Desjardins
Université du Québec à Trois-Rivières
E-mail : sophie.desjardins@uqtr.ca

Luc Granger
Université de Montréal

Few systematic empirical studies have been conducted for the purpose of obtaining a better understanding of sex offenders. Much of the data accumulated on the subject has been descriptive, as is generally the procedure when dealing with an emerging field of research. However, many of the studies providing these data are characterized by significant limitations relating, for example, to the failure to apply statistical analyses; the absence of operational definitions; very small sample sizes; confusion between legal, moral, clinical and scientific constructs; and faulty methodological procedures. This probably explains why so many contradictory results are found in the research that followed and focussed on formulating etiological theories, developing a variety of preventive and curative interventions and assessing their effectiveness.

Despite the prevalence of many antinomies, there appears to be a consensus that sex offenders are a heterogeneous group (Erickson, Luxenberg, Walbek & Seely, 1987; Hall, Maiuro, Vitaliano & Proctor, 1986; Levin & Stava, 1987; Looman, Gauthier & Boer, 2001). The search for narrow typological systems therefore makes sense, if only to allow valid between-group comparisons. Among the elements forming the basis of such systems, victim age appears to be an important discriminant variable (Bard et al., 1987), although its potential requires further development.

As evidence of this, although 27.5% of Canada's sex offenders target adolescents (Proulx et al., 1999), we know little or nothing about the group, referred to as hebephiles. Generally absent from sex offender studies, hebephiles are seldom distinguished from the pedophile group when they are included in those studies.

However, according to Baxter, Marshall, Barbaree, Davidson and Malcolm (1984), certain characteristics distinguish rapists from pedophiles and could help distinguish pedophiles from hebephiles. These characteristics include both personal factors and those relating to their criminal record. By examining these elements, we can, in fact, observe between-group differences.

Although some data supports the notion that pedophiles, hebephiles and rapists may be distinguished by personal and criminal variables, the importance of the respective contributions of those variables, where applicable, has yet to be determined. As well, a study by Baxter et al. (1984) found little difference between hebephiles and rapists. These conclusions call into question the widespread practice of including hebephiles in the category of pedophiles rather than rapists, which raises the following question: are hebephiles pedophiles, rapists or an entirely separate category of sex offender?

The purpose of this study is to begin to answer this question by investigating whether pedophiles, rapists and hebephiles in particular can be differentiated via personal and criminal variables and, if so, which of those variables contribute the most to that distinction. It is also hoped that this research will contribute to the meagre body of knowledge currently available on hebephilia.

Method

Participants

Data relating to 145 sex offenders incarcerated in a correctional centre and serving a sentence of two years or more were taken from a database developed by a research group within the Correctional Service of Canada. Participants were

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 166-169. Public Safety Canada.

The authors wish to thank the Fonds québécois de recherche sur la société and la culture (FQRSC) and the Correctional Service of Canada for their support.

divided into three groups according to victim age. The first group consisted of 49 pedophiles convicted of sexual assault on children under 13 years of age, the second of 46 hebephiles convicted of the same offence on adolescents between 13 and 17, and the third of 50 rapists convicted of sexual assault on women over 17.

For the purposes of this study, the use of the terms “pedophile,” “hebephile” and “rapist” is based on the offence committed and not on sexual preference. Although deviant sexual desires may be of interest as a subject of study, they offer fewer practical and theoretical implications than does the commission of the offence. We therefore chose observable behaviour rather than attraction, which is less easily observed.

The offenders who participated in the study denied having assaulted a victim in an age category other than from that in which they were included, and denied being related to or having killed any of their victims. Furthermore, none of the participants had been convicted of any of these offences. At the time of assessment, none of the offenders showed signs of serious mental or organic disorders.

Statistical Analyses

Prior to the discriminant function analysis to determine membership in the three groups of sex offenders, two categorical variables involving three different levels were subjected to dummy binary variable transformation, following the procedure described by Dillon and Westin (1982). Those variables were sexual orientation (homosexual, bisexual, heterosexual) and antisocial personality diagnosis (presence, traits, absence).

The data selected for analysis related to the offenders’ personal characteristics (age, employment status, sexual orientation, antisocial personality diagnosis, victimization as a child or adolescent, treatment for personal problems, hard drug use and prostitute use patterns) and criminal record (sexual assault, sexual assault with a weapon, kidnapping and/or forcible confinement).

These 11 personal and criminal variables were simultaneously entered in the discriminant function analysis, both to minimize the likelihood

of chance being a factor (Stevens, 1996) and because we had no reason to give some variables priority over others. The discriminant function analysis also considered the probabilities of belonging to each group, as recommended by Tabachnick and Fidell (2001). The probabilities were 33.8% for pedophiles, 31.7% for hebephiles and 34.5% for rapists.

Results

Two significant discriminant functions were identified from the analysis, the first with a Wilks’ lambda of .45 and a $\chi^2(26) = 107.69$, $p < .001$, and the second with a Wilks’ lambda of .84 and a $\chi^2(12) = 24.37$, $p < .05$. The first and second functions represented 81% and 19% respectively of the between-group variance. The first discriminant function achieved maximal separation between pedophiles and the other two offender groups, while the second distinguished hebephiles from rapists, with pedophiles being divided between the two groups.

Table 1 presents all the raw data, apart from those concerning offender age, which is provided directly in the text.

The first function from the analysis indicates that the variables that best differentiate pedophiles are age at time of assessment, treatment for sex problems, absence of antisocial personality diagnosis, sexual victimization as a child or adolescent, and a criminal record of sexual assault and sexual assault with a weapon.

As it turns out, pedophiles are older (44.4 ± 11.8 years old, age range = 24-71), on average, than hebephiles (35.5 ± 11.0 years old, age range = 18-64) and rapists (33.0 ± 7.7 years old, age range = 20-49). Pedophiles are more likely than other groups to have received treatment for their sex problems and to have claimed that they were sexually assaulted as children or adolescents. Pedophiles are also more likely than hebephiles and rapists to have a criminal record of sexual assault but less likely to have a record of sexual assault with a weapon. They are also less likely than the other two groups in the study to be diagnosed with antisocial personality disorder.

In terms of the second discriminant function, the three predictive variables most effective in

Table I. Personal and Criminal Characteristics of Sex Offenders

	P (%) ^b	H (%) ^b	R (%) ^b
Treatment for sex problems	50	10	14
Absence of antisocial personality diagnosis ^a	82	39	46
Antisocial personality diagnosis ^a	6	31	22
Sexual victimization before age 18	66	44	22
Record of sexual assault	52	25	20
Record of sexual assault with a weapon	4	20	32
Record of kidnapping and/or forcible confinement	8	18	36
Prostitute use	26	43	54
Employed when assessed	38	16	42
Heterosexual orientation ^a	64	74	94
Homosexual orientation ^a	20	17	4
Drug use	20	33	50

^a Dummy variable. ^b P = pedophile, H = hebephile, R = rapist.

distinguishing hebephiles from rapists are employment status at time of assessment, absence of antisocial personality diagnosis, and heterosexual orientation. Specifically, hebephiles are more likely to be unemployed and less likely to exhibit a heterosexual orientation than rapists are. Hebephiles are also less likely than rapists to be diagnosed with antisocial personality disorder.

As shown in Table II, discriminant function analysis was successfully used to classify participants in their respective groups, as 71.0% of the offenders were classified correctly, by comparison with 33.3% who would have been classified correctly by chance alone.

Like rapists, the incorrectly classified pedophiles were found primarily in the hebephile category. The incorrectly classified hebephiles were divided almost equally between the pedophile and rapist groups.

Discussion

The 11 personal and criminal variables included in this study successfully distinguished pedophiles from hebephiles and rapists, as well as hebephiles from rapists. However, hebephiles proved the most difficult to categorize, as was asserted by Harry, Pierson and Kuznetsov (1993), since barely half were classified correctly. However, the hypothesis that they represent a separate group of sex offenders is supported by the results obtained.

Of the three groups studied, the pedophiles turned out to be the least antisocial, the most likely to have been sexually assaulted as children or adolescents, but also the most likely to have obtained help for their sex problems. Data on this group suggests that, for pedophiles, sexual assault is not simply one of many behaviours contributing to an overall profile of delinquency or criminality; in fact, it represents a way of fulfilling emotional needs (McKay, Chapman & Long, 1996) and is accompanied by stronger guilt feelings than in the case of other offender groups (Blumenthal, Gudjonsson & Burns, 1999).

Hebephiles are also distinguished from pedophiles and rapists by the fact that over 80% of them were unemployed at the time of assessment.

In general, although a criminal record involving kidnapping and forcible confinement, prostitute use, hard drug use and homosexuality were useful in classifying a large group of offenders accurately, they remain secondary to the factors discussed above.

This study highlights the advisability of distinguishing hebephiles from pedophile and rapist groups in future studies, in light of the personal and criminal traits characterizing them. According to Grubin and Kennedy (1991), well-defined groups are indeed necessary in order to make decisions regarding the inclusion of essential therapeutic components in treatment programs, determine who is at risk of recidivism

Tableau II. Classification Matrix

		Predicted Group Membership			
		Pedophiles N (%)	Hebephiles N (%)	Rapists N (%)	Total N (%)
Initial Groups	Pedophiles N (%)	44 (89.8)	4 (8.2)	1 (2.0)	49 (100)
	Hebephiles N (%)	10 (21.7)	25 (54.3)	11 (23.9)	46 (100)
	Rapists N (%)	5 (10.0)	11 (22.0)	34 (68.0)	50 (100)

and establish the importance of different etiological factors.

Research is needed to determine whether the results of this study are also observed in sex offenders having a familial or other relationship with their victims, in sex killers, in adolescent sex offenders and in volunteers outside the prison system.

Références

Bard, L. A., Carter, D. L., Cerce, D. D., Knight, R. A., Rosenberg, R., & Schneider, B. (1987). A descriptive study of rapists and child molesters : Developmental, clinical, and criminal characteristics. *Behavioral Sciences & the Law*, 5, 203-220.

Baxter, D. J., Marshall, W. L., Barbaree, H. E., Davidson, P. R. et Malcolm, P. B. (1984). Deviant sexual behavior : Differentiating sex offenders by criminal and personal history, psychometric measures, and sexual response. *Criminal Justice and Behavior*, 11, 477-501.

Blumenthal, S., Gudjonsson, G. et Burns, J. (1999). Cognitive distortions and blame attribution in sex offenders against adults and children. *Child Abuse and Neglect*, 23, 129-143.

Dillon, W. R. & Westin, S. (1982). Scoring frequency data for discriminant analysis : Perhaps discrete procedures can be avoided. *Journal of Marketing Research*, 19, 44-56.

Erickson, W., Luxenberg, M., Walbek, N., & Seely, R. (1987). Frequency of MMPI two-point code types among sex offenders. *Journal of Consulting and Clinical Psychology*, 55, 566-570.

Grubin, D. H. et Kennedy, H. G. (1991). The classification of sexual offenders. *Criminal Behaviour and Mental Health*, 1, 123-129.

Hall, G. N., Maiuro, R., Vitaliano, P., & Proctor, W. (1986). The utility of the MMPI with men who have sexually assaulted children. *Journal of Clinical and Consulting Psychology*, 54, 493-496.

Harry, B., Pierson, T. R. et Kuznetsov, A. (1993). Correlates of sex offender and offense traits by victim age. *Journal of Forensic Sciences*, 38, 1068-1074.

Levin, S. & Stava, L. (1987). Personality characteristics of sex offenders : A review. *Archives of Sexual Behavior*, 16, 57-79.

Looman, J., Gauthier, C., & Boer, D. (2001). Replication of the Massachusetts Treatment Center child molester typology in a Canadian sample. *Journal of Interpersonal Violence*, 16, 753-767.

McKay, M. M., Chapman, J. W. et Long, N. R. (1996). Causal attributions for criminal offending and sexual arousal : Comparison of child sex offenders with other offenders. *British Journal of Clinical Psychology*, 35, 63-75.

Proulx, J., Granger, L., Ouimet, M., Guay, J.-P., McKibben, A., St-Yves, M., Bigras, J., Perreault, C., Brien, T., & Pellerin, B. (1999). Profil descriptif d'un échantillon de délinquants sexuels incarcérés. *FORUM – Recherche sur l'Actualité Correctionnelle*, 11, 11-14.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Hillsdale, NJ : Lawrence Erlbaum.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA : Allyn & Bacon.

An Empirical Taxonomy of Incarcerated Male Sexual Offenders Using Finite Mixture Modeling: Adult Victims

Jamison D. Fargo
Utah State University
E-mail: jamison.fargo@usu.edu

The sexual victimization of women remains an endemic social, criminal, and public health problem. Much research has sought to identify risk and protective factors related to the sexual victimization of women so that prevention and intervention strategies can be more informed and targeted. Modern criminology has recognized the heterogeneous nature of many criminal behaviors in terms of their etiology, offender-, offense-, and victim-related characteristics. Such an approach has been labeled criminal profiling or criminal investigative analysis and yields richer information about the nature of crime than reliance on aggregate statistics (Hazelwood & Burgess, 2001). Knight (1999, p. 304) stated that understanding the taxometric structure of a deviant population is the 'keystone to theory building and the cornerstone of intervention'. One area where our understanding of the heterogeneity of criminal behavior is lacking is the offender-offense-victim triad in cases of sexual victimization. Marshall (1997) identified the reduction of heterogeneity among sexual offenders into manageable proportions as a priority research area. Unfortunately, this challenge has largely been met with little to no empirical effort.

Existing taxonomies of sexual offenders are based on their psychological characteristics or motivations or on offense-related characteristics (e.g., Amir, 1971; Groth, Burgess, & Holstrom, 1977; Prentky & Knight, 1991; Rada, 1978). Existing taxonomies differ in terms of the labels applied to individual sexual offender subtypes, but each system shares underlying themes. Perhaps two of the best known taxonomies are those

proposed by 1) Groth *et al.* (1977), consisting of Power Reassurance, Power Assertive, Anger Retaliatory, and Anger Excitation subtypes and 2) Prentky and Knight's (1991) MTC:R3 (Massachusetts Treatment Center, 3rd revision) system, consisting of Opportunistic with high or low social competence; Pervasively Angry; Sexual that is either sadistic with overt or muted behaviors or non-sadistic with high or low social competence; and Vindictive with moderate or low social competence.

Although most of the aforementioned taxonomies exhibit face validity, only two investigations to date have evaluated the construct validity of these theoretical systems using quantitative methods, and none have done so using a sample that is broadly generalizable. McCabe and Wauchope (2005) used multidimensional scaling with offense-related police record data and found support for Power Reassurance and Sexual-Sadistic offender subtypes, and weak evidence for Power Assertive and Anger Retaliatory. However, no demographic, psychological, or criminal history data were available for analysis. Using cluster analysis, Rosenberg and Knight (1988) examined the role of substance abuse, life management skills, antisocial behavior, sexual aggression, and offense-related impulsivity in a treatment center sample to validate an early form of the MTC:R3. Nine meaningful sexual offender subtypes were culled from these analyses. However, limitations of this study included combining data on offenders whose victims were an admixture of adults and children and use of sexual offender participants from a single

Proceedings of the North American Correctional & Criminal Justice Psychology Conference, 2008, pp 170-174. Public Safety Canada.

The author would like to thank Dr. James Lynch for his helpful insights and his encouragement in this work. The data used in this publication were made available by the National Archive of Criminal Justice Data (NACJD; <http://www.icpsr.umich.edu/NACJD/>), located in Ann Arbor, MI. The study entitled *Study of Inmates in State and Federal Correctional Facilities, 1997* was conducted by the US Bureau of the Census. The US Bureau of the Census, the funding agency, and the NACJD are not responsible for the analyses and interpretation presented herein.

treatment center. Clearly, empirical evidence in support of sexual offender subtypes is lacking.

Therefore, the purpose of the present investigation was to develop an empirical taxonomy of male sexual offenders who victimized adult women. Taxometric systems should be theoretically driven and based on etiological and criminological variables that enable discrimination of sexual offenses into reliable (i.e., consistent across populations) and valid (i.e., clear distinction between subtypes, generalizability) categories. To meet this goal, the present study utilized concrete, easily obtainable offender-, offense-, and victim-related descriptive indices measured on a nationally representative sample of sexual offenders.

Method

Participants

The present study is based on a secondary analysis of data originally collected by the Bureau of the Census in their study entitled *Survey of Inmates in State and Federal Correctional Facilities, 1997*. Data were retrieved from the National Archive of Criminal Justice Data (<http://www.icpsr.umich.edu/NACJD>) on June 28, 2006. The study consisted of personal interviews with a nationally representative sample of state and federal prison inmates selected by 2-stage cluster sampling (first prisons, then inmates) between June and October 1997. Of the 280 prisons selected, 275 participated, resulting in a total sample size of 14,285 randomly selected inmates who agreed to participate and were interviewed (79% male). Data from male inmates incarcerated for the sexual assault or abuse (i.e., rape, sexual abuse or assault) of an adult (>17 years) female were extracted from the sample. Inmates whose controlling sexual offense included multiple victims were excluded as responses related to victim characteristics were ambiguous. A preliminary sample of 235 inmates was reduced to a final sample of 207 due to missing data.

Measures

Participants were interviewed about victim and offense characteristics as well as current and past 1) socio-demographic/economic characteristics, 2)

family dynamics and childhood experiences, 3) substance use/abuse; and 4) criminal behaviors. All variables were categorical except for age at offense and time since incarceration.

Statistical Analyses

Finite mixture modeling (FMM) as implemented in Latent Gold® 4.0 (Vermunt & Magidson, 2005) was used to examine whether the sample of sexual offenders should be best considered unitary or heterogeneous. FMM is a probabilistic clustering technique that addresses two related questions: 1) which variables distinguish latent classes or subtypes? and 2) what number of latent classes or subtypes best accounts for population heterogeneity? In FMM, the predictor variable is an unobserved multinomial latent (class) variable that 'causes' scores on the observed (indicator) variables. Based on the characteristics of each class, a description or label for each class was developed. It was hypothesized that more than one subtype of sexual offender would be identified and, therefore, a series of 1 through 10 nested latent class FMMs were compared using the AIC3 and bootstrap likelihood ratio test. Non-significant indicator variables were iteratively removed. Time served in months was covaried to (potentially) adjust for threats to internal validity associated with historical or recall biases. The 2-stage cluster sampling design was accounted for in the analysis by specifying prison ID as the primary sampling unit so that correct standard errors could be calculated.

Results

A four-class FMM possessed the best fitting, most parsimonious relationship to the data based on the AIC3. Entropy R^2 for the 4-class model was .95 and classification error was 2%. Figures 1 and 2 illustrate the percentage of occurrence of given characteristics within each sexual offender subtype. An examination of the characteristics of each subtype revealed strong similarities to the taxonomy proposed by Prentky and Knight (1991).

Class 1 (39% of sample) was most similar to the Opportunistic subtype, without differentiating between high and low social competency. This

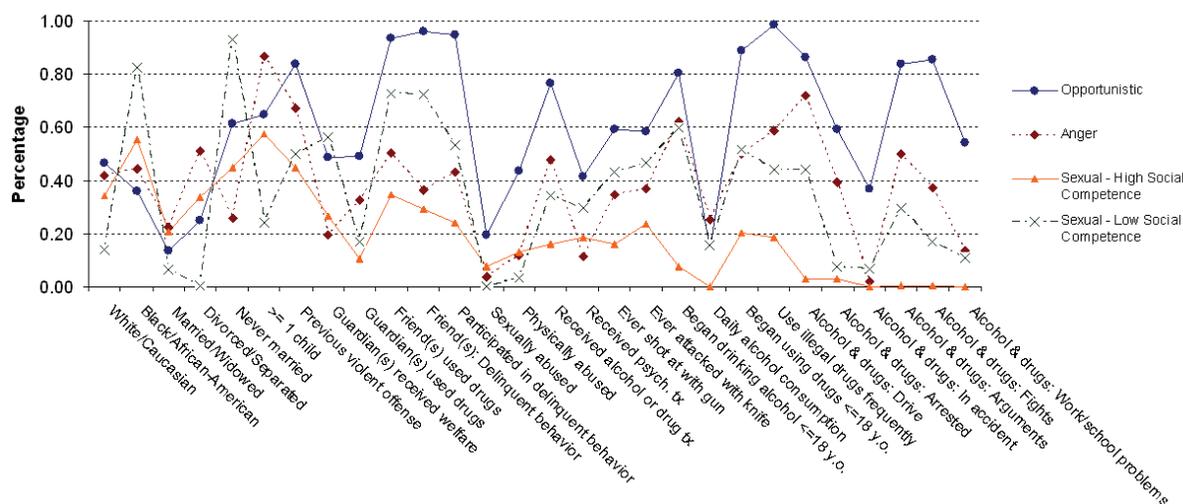


Figure 1. Offender characteristics by sexual offender subtype.

group consisted of mostly young ($M = 27$ years at arrest), single (60%) individuals, who engaged heavily in juvenile unsocial behavior (~98%), were of low SES (50% receiving welfare during childhood), had extensive prior (violent) criminal histories (>80%), were most likely to have experienced childhood sexual and physical abuse (20% and 42%, respectively), received alcohol and drug or psychiatric treatment (80% and 40%, respectively), had guardians that abused substances (50%), were victims of violent crime (60%), were frequent drug users (99%) and moderate alcohol users, had many problems associated with alcohol and drug use (40-90%), were most likely to use a knife (19%) and alcohol and drugs (80% and 60%, respectively) during the offense, were likely to commit the offense in the victim's residence (~40%), and were equally likely to perpetrate against victims of any relationship type (e.g., 34% stranger, 35% casual acquaintance, 31% well known). Thus, sexual perpetrators in this class were most likely substance abusing criminal opportunists whose assaults were part of a larger criminal background and were perpetrated against victims almost indiscriminately during sexual or other non-sexual crimes.

Class 2 (28% of sample) was most appropriately aligned with the Pervasively Angry sexual offender subtype. These were mostly older ($M = 33$ years at time of arrest), divorced (50%) individuals and more than 80% had at least one child. These individuals were low to moderate in self-reported participation in juvenile unsocial

behavior (~40%), were highest in daily alcohol consumption (25%) and moderate to high in frequent drug use (60%), were moderate to high in terms of experiencing problems associated with alcohol or drug use (~40-70%; particularly fighting and arguments), had the second highest rate of alcohol (60%) and drug (30%) use at the time of offense, had one of the highest rates of gun use during offense (10%), victimized individuals well-known to them (80%; friend or casual acquaintance), and had one of the highest rates of victim alcohol or drug use at the time of offense (>50%). The anger and difficulties reported by these individuals, particularly associated with heavy substance use and abuse, and the use of violence or force during the offense, underscores the angry nature of these sexual offenders.

Class 3 (18% of sample) was best described as possessing features of the Sexual (non-sadistic) subtype with high social competence. Individuals in this class were distinguished from others as being mostly single (~80%), slightly older ($M = 31$ years at time of arrest), and were the least likely to have a prior (violent) offense (42%). These individuals were also the least likely to have engaged in juvenile unsocial behavior (~20-30%), low in terms of welfare assistance during childhood (25%) and guardian substance abuse (10%), were virtually absent of problems associated with substance use, were also likely to know their victim (~70%; friend or casual acquaintance), had one of the highest rate of gun use during offense (10%), and were most likely to

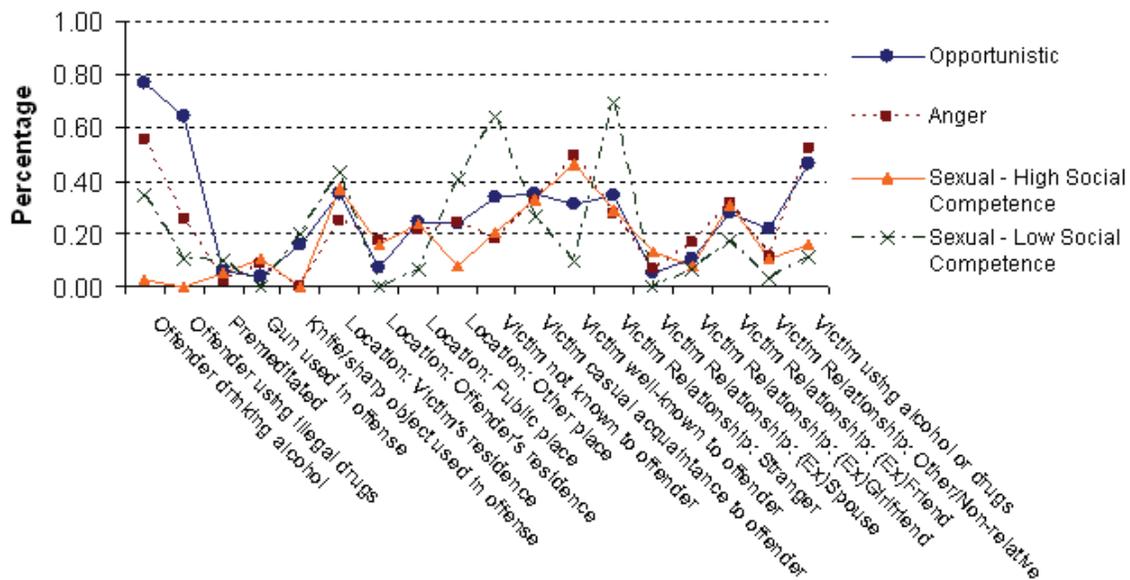


Figure 2. Offense and victim characteristics by sexual offender subtype.

commit the offense in the victim's residence (40%). Offenders in this subtype were older and likely victimized for sexual reasons, victimizing individuals that were close to or trusted them, thus demonstrating a higher level of social competence.

Class 4 (15% of sample) was also most closely associated with the Sexual (non-sadistic) subtype, but with low social competence. These individuals were the youngest subtype ($M = 23$ years at time of offense), were the most likely to be single (95%), and to have never had children (only 20%). These individuals had moderately high (violent) criminal histories (>50%) and came from the lowest SES backgrounds (~60% received welfare in childhood). Participation in juvenile unsocial behavior was moderate to high (~60-70%) and they reported moderate levels of substance use/problems (~50% frequently used drugs). These individuals were most likely to plan their sexual victimizations in advance (10%) and use violence (20% used a knife) during the offense as compared to the other subtypes. Most victims were either strangers to the offender (70%) or only a casual acquaintance (20%) and most sexual offenses took place in the victim's residence (45%) or another, unspecified, location perhaps at work or school (45%). Individuals in this subtype were younger and likely acted out for sexual reasons, but targeted mostly strangers, thus

highlighting their inability to gain trust or intimacy with their victims as compared to those in the 3rd subtype, and hence, were conjectured to be of low social competence.

Discussion

Results of the present study suggest that sexual offenses can be distinguished empirically into meaningful subtypes. The identified classes correspond to those proposed in prior research and findings lend support specifically to the MTC:R3 system. Such consistency of findings is particularly important as most research on sexual offender subtypes has not come from sources outside the original authors.

Other strengths of this study include enhanced generalizability through use of a nationally representative (random) sample, increased internal validity from use of data collected in a standardized manner, and individual-level information on offender, victim, and offense. An advantage of the present taxonomy is that it is devoid of psychoanalytic/psychological constructs that are part of many classification systems as it only relies on descriptive offender-, victim-, and offense-related characteristics. Such crime scene and criminal background variables are often available to individuals in law enforcement who often do not have the luxury of procuring psychological evaluation data when processing a

crime-scene. However, as a result, the more psychodynamically defined subtypes identified in other systems were not observed in the present study, such as the Vindictive subtype (Prentky & Knight, 1991).

An understanding of the characteristics of sexual offense classes can assist in identifying risk factors for sexual victimization. For instance, offenders in the Sexual-High Social Competence and Angry were mostly individuals intimate with their victims. Such information can be used to identify vulnerabilities to sexual victimization. Treatment for offenders may also be tailored according to the specific subtype to which they belong.

A limitation of the present study is that data were based on self-report and corresponding threats to internal validity would include recall biases, comprehension errors, and the offender faking good or bad. Also, due to the size of the sample versus the number of parameters estimated, the final FMM may be overly sample specific. Therefore, the proposed taxonomy should be considered disconfirmable and subject to further testing and modification. Taxonomies should also be interpreted narrowly, bearing in mind their purpose and the variables used in their construction. Taxometric analyses may be inaccurate for or unable to detect small classes (e.g., Sexual-Sadistic), or if a prominent subclass is represented by only a few individuals in a given sample.

In conclusion, the present study lends strong empirical support to the empirical classification of sexual offenders into distinguishable subtypes. Evidence for the existence of the Opportunistic, Angry, Sexual-High Competence, and Sexual-Low Competence subtypes proposed by Prentky and Knight (1991) is provided from a representative sample of sexual offenders. Examination of sexual offender subtype characteristics can be used to inform strategies for the prevention of sexual violence.

References

- Amir, M. (1971). *Patterns in forcible rape*. Chicago: University of Chicago Press.
- Bickley, J. A., & Beech, A. R. (2001). Classifying child abusers: Its relevance to theory and clinical practice. *International Journal of Offender Therapy and Comparative Criminology*, 45, 51-69.
- Groth, A. N., Burgess, A. W., & Holstrom, L. L. (1977). Rape: Power, anger, and sexuality. *American Journal of Psychiatry*, 134, 1239-1243.
- Hazelwood, R. R., & Burgess, A. W. (2001). *Practical aspects of rape investigation: A multidisciplinary approach* (3rd ed.). Boca Raton, FL: CRC Press.
- Knight, R. A. (1999). Validation of a typology for rapists. *Journal of Interpersonal Violence*, 14, 303-330.
- Marshall, W. L. (1997). Pedophilia: Psychopathology and theory. In D. R. Laws & W. T. O'Donahue (Eds.), *Sexual deviance: Theory, assessment, and treatment* (pp. 152-174). New York: Guilford Press.
- McCabe, M., & Wauchope, M. (2005). Behavioral characteristics of men accused of rape: Evidence for different types of rapists. *Archives of Sexual Behavior*, 34, 241-253.
- Prentky, R. A., & Knight, R. A. (1991). Identifying critical dimensions for discriminating among rapists. *Journal of Consulting and Clinical Psychology*, 39, 643-661.
- Rada, R. T. (1978). Classification of the rapist. In R. T. Rada (Ed.), *Clinical aspects of the rapist*. New York: Grune & Stratton.
- Rosenberg, R., & Knight, R. A. (1988). Determining male sexual offender subtypes using cluster analysis. *Journal of Quantitative Criminology*, 4, 383-410.
- Vermunt, J. K., & Magidson, J. (2005). *Latent Gold® 4.0 user's guide*. Boston: Statistical Innovations.
- U.S. Dept. of Justice, Bureau of Justice Statistics, & U.S. Dept. of Justice, Federal Bureau of Prisons. *Survey of Inmates in State and Federal Correctional Facilities, 1997* [Computer file]. Compiled by U.S. Dept. of Commerce, Bureau of the Census. ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 2000.

Restorative Justice: What Role Can Psychologists Play?

Tanya Ruge
Corrections Research, Public Safety Canada
E-mail: tanya.ruge@ps-sp-gc.ca

Restorative Justice (RJ) is an approach to justice that has been gaining momentum in Canada. The values that underpin restorative justice, such as holding the offender accountable to the victim and the community, and providing a voice to all those affected by crime, have been welcomed as a positive alternative to the current traditional justice system. Research has shown positive impacts of restorative justice processes on participants. However, restorative justice models, programming, and research are still, to some degree, in the developmental stage. Questions such as what model will work best at what time and with whom are yet to be determined. Areas where psychology may play an active role, such as risk assessment and treatment services to offenders and victims, still appear to be on the outskirts of restorative justice programming. Given the growing popularity of restorative justice, it is appropriate to discuss what role psychology can play in the following areas: in the development and application of restorative justice programs at various stages in the criminal justice system with differing populations of victims and offenders, in effectively meeting the needs of restorative justice participants, and in assisting with the reparation and healing of affected parties. This paper will address these questions. But, before delving into these questions, recent research findings will be reviewed to familiarize those who are relatively new to the area of restorative justice.

Restorative Justice Research

Victims, offenders, and the community are all affected by crime. Restorative justice is an approach that addresses crime using a holistic approach, involving all affected parties in a process of communication towards restoration, transformation, and reparation, with the goal of attempting to repair the harm caused, and healing

relationships. Dr. Robert Cormier described restorative justice as follows:

An approach to justice that focuses on repairing the harm caused by crime while holding the offender responsible for his or her actions, by providing an opportunity for the parties directly affected by crime – victim(s), offender and community – to identify and address their needs in the aftermath of a crime, and seek a resolution that affords healing, reparation and reintegration, and prevents future harm (Cormier, 2002, p.1).

Restorative justice is certainly not a new idea and although it has been gaining momentum in Canada in the past decade, in many aspects, research is still in its infancy. That said, research to date has established that participants of restorative justice programs are generally more satisfied than those who experience the traditional criminal justice system (Beven, Hall, Froyland, Steels, & Goulding, 2005; Bonta, Jesseman, Ruge, & Cormier, 2006; Bonta, Wallace-Capretta, Rooney, & McAnoy, 2002; Clairmont, 2005; Latimer, Dowden, & Muise, 2001; Ruge, Bonta, & Wallace-Capretta, 2005; Strang, 2002; Umbreit, 1994). Research has also found that participants express high levels of perceived fairness from the restorative justice process (Ruge, 2005) and that restorative justice processes have a strong effect on both victims and offenders in humanizing the system (Umbreit, 1994). Regarding victims specifically, research suggests that victims who participate in restorative justice programs experience increased empowerment, a lessening of anger, and decreased levels of fear (Ruge et al., 2005; Strang, 2002; Wemmers & Cyr, 2005). Restorative justice practices have also been found to have positive impacts on both victims' and offenders' overall wellbeing, as well as

psychological and physical health (Rugge, 2006). For offenders, restorative justice has been found to lower the likelihood of offenders reoffending by up to 12% (Bonta et al., 2006; Latimer et al., 2001; Rugge et al., 2005). Furthermore, offenders who participate in restorative justice programs are more likely to comply with restitution orders (Bonta et al., 2006; Bonta et al., 2002; Latimer et al., 2001).

These research highlights are important as they clearly indicate that restorative justice is having an impact and is worthy of further exploration. Although research in this area is promising, there are many areas where it has not yet been “proven” that restorative justice works in all contexts with all people. There is still a need to test the various combinations of restorative justice processes (e.g., various restorative justice models, unique populations of participants, different stages of the criminal justice process, etc.). We need to learn what works best for whom before it is likely that restorative justice will become an established option in our Canadian criminal justice system.

In order to work towards this end, testing of specific restorative justice models is necessary. Experience working with restorative justice practitioners and program staff has shown how challenging it can be for practitioners to stay within a restrictive pilot model when their philosophy is to help people who have been affected by crime. The most appealing aspect of restorative justice, its flexibility, is also the aspect that can make empirical research very difficult. However, testing specific models is essential in order to advance the field of restorative justice. This is one area where psychology can play an important role. Past experience has also indicated that psychological researchers and restorative justice practitioners can work collaboratively and effectively through open communication and understanding to examine various program models. Yet, there is a continued need for innovative ideas and collaboration.

Potential Areas of Involvement for Psychology

Before reviewing the areas where the field of psychology could contribute to restorative justice, it is critical to say that these suggestions do not imply that the field of restorative justice needs

assistance. The purpose of this piece is to highlight the skills and areas of psychology that may be of potential use to restorative justice. Those who work in the area of restorative justice are strongly dedicated to their work and there is no doubt that they offer the best of their expertise to their clients every day. However, regardless of how informed and knowledgeable each discipline may be, our common desire to always strive for improvement brings us to the recognition that sharing across disciplines can be very fruitful and ultimately may benefit everyone.

Although there are many potential areas of involvements, three have been identified where the field of psychology can contribute to restorative justice: (1) program development, (2) program evaluation and research, and (3) program effectiveness. Program development is the first area where psychology’s expertise may be put to good use. Psychologists and those in the psychology field have been designing pilot programs based on research findings for decades. Our experiences could be useful to the area of restorative justice, a field that is relatively young in the program development arena, especially for designing programs based on solid research. Within this area, psychologists are also quite effective at disseminating research findings and may therefore be of use translating the research and sharing the effects of restorative justice to others. In order for this to happen, we need to connect with, educate, and increase awareness of our skill set in the area of program development. For example, we are the leaders in risk assessment and the core principles of effective correctional treatment, but it is likely that few restorative justice practitioners and program developers know this about psychology. A partnership between the two disciplines to design and test pilot programs based on research findings could certainly be beneficial.

The second area where psychology could play an important role is in the area of evaluation and research more generally. Many of those in the field of psychology have been conducting independent evaluations and are dedicated to advancing research. Struggling with methodological challenges and conundrums is what brings some of us joy. It would prove beneficial if we could educate others about the

importance of independent program evaluation, and how research is necessary to further expand our knowledge base as we continue to learn and ultimately improve our criminal justice programs and processes. A partnership between the fields of psychology and restorative justice could allow psychology to usefully assist those in restorative justice in designing and conducting high-quality program evaluations of existing programs, as well as designing research projects to advance restorative justice research.

The third area of potential contribution is regarding program effectiveness. The overarching question here is can our skills and expertise increase the effectiveness of restorative justice programs? Psychologists have knowledge, skills, and expertise that would be useful to the discipline of restorative justice by enhancing its effects. Psychologists aim to be experts in human behaviour. Many of us who work in the area of criminal justice and correctional psychology work with offenders, victims, and the community in situations of crime, conflict, and/or trauma. Our expertise and research could be useful in several areas but only two specific domains will be examined in more detail: assessment and treatment.

A major area within the discipline of psychology is “assessment” of individuals. Canada leads the world in terms of risk assessment of offenders (e.g., assessing an offender’s risk factors, needs, and overall risk level to reoffend). Psychologists could assist restorative justice practitioners in assessing offenders’ risk levels (e.g., low-risk to reoffend, medium-risk, high-risk). This information could be beneficial to caseworkers and facilitators in order to address concerns about potential revictimization but also may inform facilitators of the most effective means to work with offenders at various risk levels. For example, higher risk offenders will have multiple need areas and will require high levels of intervention for change to occur. Many psychologists are also skilled in trauma assessment of victims (e.g., assessing a victim’s needs after trauma) and could be informative in cases where victims have suffered severe trauma and the timing of the victim-offender meeting needs to be delicately determined. Accurate and comprehensive assessment of individuals involved

in restorative justice processes may increase overall restorative justice program effectiveness.

The second item under program effectiveness where psychology could contribute to the field of restorative justice is “treatment”. Again, psychology is leading the way in terms of the core principles for effective correctional treatment of offenders (e.g., determining appropriate treatment for offenders based on the risk, need, and responsivity principles). Many of us may not have considered the commonalities across the two disciplines when it comes to treatment. In reality, the goals of restorative justice programming and the core principles for effective correctional treatment for offenders attempt to accomplish the *same* tasks - both attempt to effectively meet the needs of the offender, the victim(s), and the community. The ultimate goal of both restorative justice and the core principles is to decrease the likelihood of future criminal activity on the part of the offender and to prevent future harm. In short, restorative justice processes can be interpreted as a type of “treatment” and the area of psychology could be valuable in assisting to improve treatment effectiveness by applying our extensive knowledge in this area. Psychologists could also assist in determining appropriate treatment plans for victims and the best means to address psychological aspects of their recovery. Also falling under the discussion of treatment and program effectiveness is the principle of responsivity. Responsivity is ensuring the type and delivery of treatment is matched to client’s needs and learning style (e.g., gender, risk level, learning style, race, etc.). Attending to responsivity issues can increase program engagement generally and program success for all participants.

The disciplines of psychology and restorative justice both possess impressive and important knowledge and skills within their respective areas of expertise. Now, imagine what could be accomplished if these skills and knowledge were combined. A collaborative partnership between the two disciplines, exchanging ideas and knowledge, may lead to practitioners improving their already effective practices, and may lead to increasing overall program success. Restorative justice practices, holding the offender accountable to victims and the community, while adhering to

the principles of effective correctional treatment – the effects could be stunning.

Next Steps

Although many feel that a partnership and collaboration is a worthy goal, the next question is how can those in the area of psychology become involved. At a minimum, the creation of partnerships and networks is necessary. Our field needs to share information about our areas of expertise and seek out opportunities for networking and training. For those of us who want to have a foot in each discipline, we need to learn more about the other. For those of us in psychology, we need to engage in methods to learn more about restorative justice and then highlight and demonstrate how our skills can advance the growing field of restorative justice. After all, we share the common goals of wanting to effectively meet the needs of the offender, the victim, the community, and ultimately, decrease or prevent further harm. Let's work together.

References

- Beven, J. P., Hall, G., Froyland, I., Steels, B., & Goulding, D. (2005). *Psychiatry, Psychology and Law*, 12(1), 194-206.
- Bonta, J., Jesseman, R., Rugge, T., & Cormier, R. (2006). Restorative justice and recidivism: Promises made, promises kept? In D. Sullivan & L. Tift (Eds.), *Handbook of Restorative Justice: A Global Perspective* (pp. 108-120). New York, NY: Routledge.
- Bonta, J., Wallace-Capretta, S., Rooney, J., & McAnoy, K. (2002). An outcome evaluation of a restorative justice alternative to incarceration. *Contemporary Justice Review*, 5(4), 319-338.
- Clairmont, D. (2005). *The Nova Scotia Restorative Justice Initiative: Final Evaluation Report*. Ottawa: National Crime Prevention Centre. Government Report.
- Cormier, R. B. (2002). *Restorative justice: Directions and principles – developments in Canada*. Solicitor General Canada, User Report 2002-02.
- Latimer, J., Dowden, C., & Muise, D. (2001). *The effectiveness of restorative justice practices: A meta-analysis*. Ottawa: Department of Justice Canada.
- Rugge, T., Bonta, J., & Wallace-Capretta, S. (2005). *Evaluation of the Collaborative Justice Project: A Restorative Justice Program for Serious Crime*. Ottawa: Public Safety and Emergency Preparedness Canada, User Report 2005-02.
- Rugge, T. (2006). *The Impact of Restorative Justice Practices on Participants*. Unpublished doctoral dissertation, Carleton University, Ottawa.
- Strang, H. (2002). *Repair or Revenge: Victims and Restorative Justice*. Oxford: Clarendon Press.
- Umbreit, M. S. (1994). *Victim Meets Offender: The Impact of Restorative Justice and Mediation*. Monsey, NY: Criminal Justice Press.
- Wemmers, J., & Cyr, K. (2005). Can mediation be therapeutic for crime victims? An evaluation of victims' experiences in mediation with young offenders. *Canadian Journal of Criminology and Criminal Justice*, 47(3), 527-544.