CRIMINAL JUSTICE AND BEHAVIOR
Volume 30, Number 6, December 2003

Contents

SAGE Publications Thousand Oaks • London • New Delhi

CRIMINAL JUSTICE AND BEHAVIOR seeks contributions examining psychological and behavioral aspects of the juvenile and criminal justice systems. The concepts "criminal justice" and "behavior" should be interpreted broadly to include analyses of the etiology of delinquent or criminal behavior, the process of law violation, of victimology, offender classification and treatment, deterrence, and incapacitation. The journal will include analyses of both clientele and employees in the justice systems, and it will include analyses of the effects of differing sanctions or programs. The journal emphasizes reports of original empirical research, theoretical contributions, development and testing of innovative programs and practices, and critical reviews of literature or theory on central topics of criminal justice and behavior. Articles dealing with behavioral aspects of juvenile or criminal justice are welcomed from throughout the world.

MANUSCRIPTS should be submitted in triplicate to Curt R. Bartol, Ph.D., Editor, **CRIMINAL JUSTICE AND BEHAVIOR,** Department of Psychology, Castleton State College, Castleton, VT 05735, typewritten double-spaced with tables, charts, footnotes, and references on separate pages. The format described in the *Publication Manual of the American Psychological Association* (5th edition) must be followed. Questions concerning manuscript submissions may be directed to curt.bartol@castleton.edu.

BOOKS FOR REVIEW should be sent to David J. Simourd, Ph.D., Book Review Editor, **CRIMINAL JUSTICE AND BEHAVIOR**, Providence Continuing Care Center, Forensic Unit, 752 King St. W., Kingston, Canada K7L 4X3; e-mail: simourdd@pccc.kari.net. Potential contributors of book reviews should correspond with the Book Review Editor prior to submission of manuscripts.

CRIMINAL JUSTICE AND BEHAVIOR (ISSN 0093-8548)(J233) is published six times annually—in February, April, June, August, October, and December—by Sage Publications, 2455 Teller Road, Thousand Oaks, CA 91320. Telephone (800) 818-SAGE (7243) and (805) 499-9774; fax/order line (805) 499-0871; e-mail order@sagepub.com; http://www.sagepub. com.

Subscriptions: Annual subscription rates for institutions and individuals are based on the current frequency. Prices quoted are in U.S. dollars and are subject to change without notice. Canadian subscribers add 7% GST (and HST as appropriate). Outside U.S. subscription rates include shipping via air-speeded delivery. Institutions: $551 (within the U.S.) / $575 (outside the U.S.) / single issue: $106 (worldwide). Individuals: $84 (within the U.S.) / $108 (outside the U.S.) / single issue: $25 (worldwide).

Orders from the U.K., Europe, the Middle East, and Africa should be sent to the London address (below). Orders from India and South Asia should be sent to the New Delhi address (below). Noninstitutional orders must be paid by personal check, VISA, MasterCard, or Discover.

Periodicals postage paid at Thousand Oaks, California, and at additional mailing offices.

This journal is abstracted and indexed in **Academic Abstracts, Academic Search, Applied Social Sciences Index & Abstracts (ASSIA), Asia Pacific Database, Biological Sciences Abstracts, Central Asia: Abstracts & Index, Corporate ResourceNET, Criminal Justice Abstracts, Current Citations Express, Current Contents: Social & Behavioral Sciences, Expanded Academic Index, Health & Safety Sciences Abstracts, ISI Basic Social Sciences Index, Legal Periodicals in English, Legal Trac, MasterFILE FullTEXT, Mental Health Abstracts, Middle East: Abstracts & Index, Military Library FullTEXT, NCJRS Abstracts Database, North Africa: Abstracts & Index, Periodical Abstracts, Prevention Evaluation Research Registry for Youth (PERRY), Psychological Abstracts, PsycINFO, PsycLIT, Public Library FullTEXT, Risk Abstracts, Safety Science & Risk Abstracts, Sage Urban Studies Abstracts, Social Science Source, Social Sciences Citation Index, Social Sciences Index Full Text, Social Services Abstracts, Sociological Abstracts, Southeast Asia: Abstracts & Index, Standard Periodical Directory (SPD), TOPICsearch, Violence & Abuse Abstracts, Wilson OmniFile V,** and **Wilson Social Sciences Index/Abstracts** and is available on microfilm from University Microfilms, Ann Arbor, Michigan.

Back Issues: Information about availability and prices of back issues may be obtained from the publisher's order department (address below). Single-issue orders for 5 or more copies will receive a special adoption discount. Contact the order department for details. Write to London office for sterling prices.

Inquiries: All permissions requests and subscription inquiries, orders, and renewals with ship-to addresses in North America, South America, Australia, China, Indonesia, Japan, Korea, New Zealand, and the Philippines must be addressed to Sage Publications, 2455 Teller Road, Thousand Oaks, CA 91320, U.S.A., telephone (800) 818-SAGE (7243) and (805) 499-9774, fax (805) 499-0871; e-mail order@sagepub.com; http://www.sagepub.com. All subscription inquiries, orders, and renewals with ship-to addresses in the U.K., Europe, the Middle East, and Africa must be addressed to Sage Publications, Ltd, 6 Bonhill Street, London EC2A 4PU, England, telephone +44 (0)20 7374 0645, fax +44 (0)20 7374 8741. All subscription inquiries, orders, and renewals with ship-to addresses in India and South Asia must be addressed to Sage Publications India Pvt Ltd, B-42 Panchsheel Enclave, Post Box 4109, New Delhi 110-017, telephone (91-11) 2649-1290, fax (91-11) 2649-1295. Address all permissions requests to the Thousand Oaks office.

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Sage Publications, for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of 50¢ per copy, plus 10¢ per copy page, is paid directly to CCC, 27 Congress St., Salem, MA 01970. 0093-8548/2003 $.50 + .10.

Advertising: Current rates and specifications may be obtained by writing to the Advertising Manager at the Thousand Oaks office (address above).

Claims: Claims for undelivered copies must be made no later than six months following month of publication. The publisher will supply missing copies when losses have been sustained in transit and when the reserve stock will permit.

Change of Address: Six weeks' advance notice must be given when notifying of change of address. Please send old address label along with the new address to ensure proper identification. Please specify name of journal. POSTMASTER: Send address changes to Criminal Justice and Behavior, c/o 2455 Teller Road, Thousand Oaks, CA 91320.

Printed on acid-free, recycled paper

CHILD MALTREATMENT AND RECIDIVISM AMONG ADOLESCENT DETAINEES

J. B. KINGREE
Clemson University
DEBRA PHAN
Emory University
MARTIE THOMPSON
Clemson University

Many studies have examined predictors of recidivism among adolescent detainees. A recent meta-analysis of these predictors indicated that child maltreatment is associated with recidivism. This study expanded on prior work on this topic by using a well-validated self-report instrument to assess abuse and neglect experiences. Results revealed that emotional neglect predicted recidivism during a 6-month follow-up period. Implications and limitations of the findings as well as suggestions for future research are discussed.

Keywords: adolescent detainees; child abuse; child neglect; recidivism

Child maltreatment involves a wide range of harmful behaviors directed toward children. It includes physical (e.g., hitting with force), sexual (e.g., fondling), and emotional (e.g., belittling statements) abuse as well as physical (e.g., withholding provisions) and emotional (e.g., withholding love) neglect. Physical, sexual, and emo-

AUTHORS' NOTE: *This study was conducted with financial support from a center grant (No. 5 P30 DA 12121-02) funded by the National Institute of Drug Abuse. The authors thank the Honorable Robin S. Nash of the DeKalb Juvenile Court for assisting with the design and conduct of this study.* *Correspondence concerning this article should be addressed to J. B. Kingree, Department of Public Health Sciences, 517 Edwards Hall, Clemson University, Clemson, South Carolina, 29634; e-mail: kingree@clemson.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 30 No. 6, December 2003 623-643
DOI: 10.1177/0093854803256460

tional abuse are acts of commission that may be perpetuated by an adult or another juvenile. Physical and emotional neglect are acts of omission that are perpetuated by parents or legal guardians.

Child maltreatment is a major social problem in the United States. For every 1,000 children in the country, there were 2.5 cases of physical abuse, 1.3 cases of sexual abuse, 6.5 cases of neglect, and .9 cases of psychological maltreatment reported officially to state agencies in 1999 (Administration on Children, Youth, & Families, 2001). Considering both reported and unreported cases, the National Incidence Study of Child Abuse and Neglect estimated that there were more than 1.5 million maltreated children in the United States in 1993. Research has shown that victims of child maltreatment are vulnerable for a wide range of psychosocial and behavioral problems, including depression, substance abuse, and low educational attainment (Kaplan, Pelcovitz, & Labruna, 1999; National Research Council, 1993; Perez & Widom, 1994; Widom, Weiler, & Cottler, 1999).

One of the more disturbing consequences of child maltreatment is that victims themselves often become perpetrators of crime. Although determining the sequence of occurrence between child maltreatment and criminal offending can be challenging, numerous studies have documented that these two experiences are strongly associated (National Research Council, 1993; Smith & Thornberry, 1995; Widom, 1989). One very noteworthy, longitudinal study that compared individuals with substantiated cases of child maltreatment against controls found that abused and/or neglected children had a 27% likelihood of being arrested as juveniles and a 42% likelihood of being arrested as adults (Widom & Maxfield, 2001).

Because maltreatment is strongly associated with delinquency, one would also expect it to be associated with recidivism among adolescent offenders. A recently published meta-analysis on recidivism among juveniles included five studies that examined child maltreatment in relation to recidivism (Cottle, Lee, & Heilbrun, 2001). Collectively, the five studies indicated that child maltreatment was a significant predictor of recidivism (Archwamety & Katsiyannis, 1998; Dembo et al., 1998; Katsiyannis & Archwamety, 1997; Myner, Santman, Capelletty, & Perlmutter, 1998; Towberman, 1994). However, child maltreatment had a weaker association with recidivism than did many other risk factors, which was surprising given its strong effects on delinquency.

Due to its multidimensional nature, child maltreatment is best assessed with instruments that tap experiences with different types of abuse and neglect (National Research Council, 1993). Among the five studies cited previously that examined child maltreatment in relation to recidivism, only one included measures of both abuse and neglect (Dembo et al., 1998). This study found that recidivism was more strongly predicted by neglect than by different types of abuse. Thus, the meta-analysis may have found a stronger association between child maltreatment and recidivism if the four other studies had employed more comprehensive measures of abuse and neglect.

The assessment of maltreatment has been significantly facilitated in recent years by the introduction of the Childhood Trauma Questionnaire (CTQ). Various versions of this self-report instrument have appeared in the literature during the past decade, each with adequate psychometric properties (Bernstein, Ahluvalia, Pogge, & Handelsman, 1997; Bernstein et al., 1994; Fink, Bernstein, Handelsman, Foote, & Lovejoy, 1995; Wolfe, Scott, Wekerle, & Pittman, 2001; Wright et al., 2001). The versions have differed primarily in the number of constituent factors and items. Recent versions have included five separate subscales (i.e., Physical Abuse, Physical Neglect, Sexual Abuse, Emotional Abuse, and Emotional Neglect) and fewer items than did the earlier versions. The CTQ was developed primarily for adult samples, but it has been effectively used with adolescents (Bernstein et al., 1997), including those who come into contact with the juvenile justice system (Carrion & Steiner, 2000). The CTQ and other self-report instruments are generally advantageous for research studies as they provide more sensitive measures of abuse and neglect than do official records. As many victims of child maltreatment—and especially victims of emotional neglect—never come to the attention of the authorities, relying on official records can result in many youths being misclassified as nonvictims.

Because problem behaviors tend to co-occur in the same individuals (Donovan & Jessor, 1985), studies that examine different types of maltreatment as risk factors for recidivism should also consider other variables that are tied to recidivism. Inclusion of these other variables or covariates in multivariate models will allow for determining if specific types of maltreatment add uniquely to the prediction of recidivism. The meta-analysis cited previously and studies published since

then point to the importance of considering demographic factors along with abuse and neglect. Although there has been some variability in the specific findings across different studies, the collective evidence indicates that men, minorities, and younger adolescents are more apt to recidivate than are women, Caucasians, and older adolescents, respectively (Dembo, 1998; Dembo et al., 1998; Giblin, 2002; Harrison, Maupin, & Mays, 2001; Hodges & Kim, 2000; Quist & Matshazi, 2000). Prior detention experiences should also be considered as various indicators of these experiences (e.g., age at first arrest or total amount of time detained) have been found to be strongly associated with recidivism (Cottle et al., 2001; Dembo, 1998). In addition, the personality variables of self-restraint and distress should be considered because a recently published study indicated they were predictive of recidivism even after controlling for prior detention experiences (Steiner, Cauffman, & Duxbury, 1999). Moreover, substance abuse should be considered not only because of its relation to recidivism (Cottle et al., 2001; Kataoka et al., 2001) but also because it is associated with many other risk factors for recidivism (Dembo, 1995; Wilson, Rojas, Haapanen, Duxbury, & Steiner, 2001).

With these issues in mind, this study examined if recidivism in a sample of adolescent detainees was uniquely predicted by different types of abuse and neglect. The CTQ and other study measures were administered to adolescents while they were being detained in a short-term facility. Recidivism was assessed through a subsequent review of official records. Given the findings discussed previously, it was hypothesized that recidivism would be more strongly predicted by neglect than by abuse experiences. It was also hypothesized that neglect would predict recidivism even after controlling for relevant sociodemographic and behavioral variables.

METHOD

SAMPLE

This study targeted adolescents who were being detained in a holding facility for juveniles in a county in metropolitan Atlanta. The holding facility was operated by a juvenile court and was used for youths who had been apprehended in the county for a status offense (e.g., tru-

ancy or running away) or delinquent offense (e.g., assault or theft). The adolescents were asked to enroll in the study after their cases had been adjudicated.

A total of 282 adolescents were invited to enroll in the study. Only 6 adolescents declined to enroll, and 4 others withdrew after initially agreeing to enroll. Thus, there were 272 adolescents (a 96% participation rate) in the initial sample.

In Georgia, offenders who are 17 years or older are processed in the adult correctional system unless they are charged with a status offense or violation of a probated sentence. Because individuals who were 16 years and 6 months or older at the time of the baseline interview (n = 42; 15%) could have been processed in the adult system if charged with reoffending in the follow-up period, they were excluded from the analyses as we did not have access to adult system records and could not assess recidivism for them. In addition, 13 individuals (5%) who were younger than 16 years and 6 months at baseline were excluded from the analyses because no information could be located for them in the database maintained by the juvenile court facility. Therefore, the final sample was composed of 217 respondents (80%) of the original sample.

Each of the adolescents was charged with one primary offense at baseline. The 217 adolescents were charged with a wide range of primary offenses. The most frequent primary offenses at baseline were burglary (n = 30; 14%), probation violation (n = 27; 12%), simple battery (n = 26; 12%), theft by taking or shoplifting (n = 22; 10%), and running away (n = 19; 9%).

PROCEDURES

Eligibility for participation in the study was based on whether adolescents were being detained when same-sex research interviewers were stationed at the facility. During the data collection period, one female and one male interviewer were present at the holding facility in 3-hour intervals on different weekdays. The specific weekdays and times that each interviewer was present varied during the data collection period. The age, gender, race, and primary criminal charge of potential respondents did not differ significantly in relation to the day of the week or the time of day that the interviews were conducted.

Each interviewer was introduced individually to potential respondents by a security officer. The researchers requested that the security officer include one female for every two male youth who were given an opportunity to participate. Otherwise, no specific eligibility criteria were used for participating in the study. Because background data were not available for adolescents in the facility who were not given an opportunity to participate, it cannot be determined if they differed systematically from those who were introduced to the researchers.

Once the introduction took place, the interviewer provided a description of the study and asked the adolescents to participate in it. The adolescents were assured that the interviewer was not an employee of the court and would not share information from the interviews with court staff members or anyone else without their permission. Adolescents who agreed to participate gave informed assent by reading and signing a form that listed their rights as research participants. Consent for the adolescents' participation was provided by an official of the juvenile court. The study procedures were approved by the Emory University Institutional Review Board.

Nine months after the last adolescent in the sample enrolled in the study, arrest records maintained by the court were reviewed to determine which adolescents had recidivated. The records were reviewed for each adolescent for a 6-month period following his or her baseline assessment.

VARIABLES

Demographics. The adolescents' gender (0 = male and 1 = female) and race (0 = Caucasian and 1 = African American) were coded as dichotomous variables. Age was treated as a continuous variable. The original sample included 169 males and 103 females as well as 232 African Americans and 40 adolescents who were not African American. The ages of the adolescents ranged from 13 to 17 ($M = 14.82$, $SD = 1.14$).

Prior detention. The adolescents' detention history was assessed by asking whether they had ever been "locked up" in a detention facility prior to the present episode (0 = no and 1 = yes).

Substance use consequences. A 12-item scale derived from the Problem Oriented Screening Instrument for Teenagers was used to assess various negative consequences (e.g., forgetting things or getting into arguments) that the adolescents may have experienced from using substances during their lifetimes (Latimer, Winters, & Stinchfield, 1997). The experience of such consequences is indicative that adolescents may have a substance abuse problem. The adolescents responded to each item by indicating whether it was true or false for them. Responses were summed to create continuous scores reflecting the number of negative consequences they had experienced. This measure was used instead of one that merely tapped the frequency or quantity of substance use because prior research indicated that a measure of substance use severity was a stronger predictor of recidivism (Cottle et al., 2001).

Self-restraint and emotional distress. The adolescents completed the condensed Weinberger Adjustment Inventory (Weinberger & Bartholomew, 1996), which included separate, 12-item measures of self-restraint ($\alpha = .77$) and emotional distress ($\alpha = .77$). A 5-point Likert-type response format was used for individual items. The two measures were coded continuously. High scores on the self-restraint measure were indicative of a nonimpulsive, responsible personality. High scores on the emotional distress measure were indicative of high distress.

Maltreatment. The adolescents completed a condensed version of the CTQ. This version was created from psychometric analyses that were conducted on the original 70-item version of the questionnaire (Bernstein & Fink, 1998; Bernstein et al., 1997). The psychometric analyses resulted in a 28-item measure that included the following six subscales: (a) Physical Abuse (5 items), (b) Physical Neglect (5 items), (c) Sexual Abuse (5 items), (d) Emotional Abuse (5 items), (e) Emotional Neglect (5 items), and (f) Minimization/Denial (3 items). Consistent with the approach taken by other researchers (Scher, Stein, Asmundson, McCreary, & Forde, 2001; Wright et al., 2001), the Minimization/Denial subscale was dropped because it was not a focus of this study. In addition, one item that was originally included in the

Emotional Neglect subscale was dropped because prior research indicated that it was not a strong indicator of the construct in this population (Kingree, Braithwaite, & Woodring, 2001).

Therefore, our measure included 24 items constituting five subscales. Responses to individual items were made on a 5-point Likert-type scale, with response options ranging from 1 (*never true*) to 5 (*very often true*). The five subscales asked about specific occurrences of abuse or neglect that the adolescents may have experienced in their lives and were coded as continuous measures.

Recidivism. Offense records maintained by the juvenile court provided two date-specific measures of recidivism for the 6-month period (i.e., 182 days) following the baseline interview. One measure of recidivism indicated whether the adolescents were charged with any offenses during the 6 months following the baseline interview. This "any recidivism" measure was coded dichotomously (0 = no and 1 = yes). The second measure of recidivism indicated the number of times the adolescents were charged with offenses during the follow-up period. It was coded continuously. Adolescents who did not recidivate were assigned a score of 0 on the times recidivated measure.

STATISTICAL ANALYSES

Factor structure and subscale reliabilities of the condensed CTQ. Initial analyses examined the factor structure and reliability of the 24-item version of the CTQ. The factor structure was evaluated through a single factor analysis with varimax rotation. Items were retained if they loaded greater than .50 on the designated factor and no greater than .40 on any other factor. Internal consistency analyses were then conducted on the retained items for each subscale.

Variable descriptives. Frequency analyses were used to illuminate the adolescents' responses to the study variables. These analyses provided percentiles for each category of the dichotomously coded variables as well as means and standard deviations for the continuously coded variables.

Associations between predictors and recidivism. Two sets of multiple regression analyses were used to examine associations between the predictor and recidivism variables. One set used logistic regression analyses and specified "any recidivism" as the criterion. The other set used linear regression analyses and specified "times recidivated" as the criterion. Due to the small sample size relative to the number of predictor variables, we did not test all of the predictors simultaneously in a single model. Instead, both sets of regression analyses included three models. The first model examined sociodemographic and behavioral variables. The second model examined maltreatment variables. The third model simultaneously examined the predictor variables that were significant in the first and second models to determine which of them were uniquely associated with each measure of recidivism.

RESULTS

FACTOR STRUCTURE AND SUBSCALE RELIABILITIES OF THE CONDENSED CTQ

Five factors with eigenvalues greater than one emerged from the factor analysis. Collectively, the five factors accounted for 69% of the variance in the total scale. As reflected in the factor loadings for the individual items presented in Table 1, 21 of the 24 items met the criteria for retention. Two items (Items 2 and 24) that loaded on the Emotional Neglect factor were dropped because they were developed for the Physical Neglect factor. Another item (Item 20) was dropped because it loaded less than desired on the designated factor (i.e., Physical Abuse) and more than desired on two other factors (i.e., Emotional Abuse and Emotional Neglect). The internal consistency was adequate for each of the subscales, although it was lower than desired for the Physical Neglect subscale.

DESCRIPTIVE VALUES FOR THE STUDY VARIABLES

Table 2 includes descriptive values for the study variables. As shown, 46% (n = 100) of the adolescents recidivated during the 6-

TABLE 1. Childhood Trauma Questionnaire Item Loadings and Internal Consistencies for Its Five Subscales

	Designated Factor	*PA*	*PN*	*SA*	*EA*	*EN*
1. I did not have enough to eat.	PN	.17	.79[a]	.07	.03	.12
2. I knew there was someone to take care of me and protect me.[b]	PN	.29	.14	.05	.05	.67
3. People in your family called you things like stupid, lazy, and ugly.	EA	.14	.12	.18	.84[a]	.12
4. My parents were too drunk or high to take care of the family.	PN	.21	.57[a]	.12	.12	.34
5. I had to wear dirty clothes.	PN	–.02	.76[a]	–.01	.14	–.01
6. There was someone in my family that helped me feel that I was important or special.[b]	EN	–.15	.06	–.01	.13	.74[a]
7. I was punished with a belt, a board, a cord, or some other hard object.	PA	.64[a]	.17	.03	.35	–.06
8. People in my family said hurtful or insulting things to me.	EA	.25	.16	.14	.79[a]	.21
9. I got hit or beaten so badly that it was noticed by someone like a teacher, neighbor, or doctor.	PA	.76[a]	.06	.30	.24	.22
10. I believe that I was physically abused.	PA	.68[a]	.11	.30	.25	.33
11. I felt loved.[b]	EN	.23	.08	.11	.29	.72[a]
12. Someone tried to touch me in a sexual way or tried to make me touch them.	SA	.16	.05	.89[a]	.11	.01
13. Someone threatened to hurt me or tell lies about me unless I did something sexual with them.	SA	.14	.09	.84[a]	.05	.10

14. People in my family looked out for each other.[b]	EN	.15	.13	.24	.30	.69[a]
15. Someone in my family hated me.	EA	.35	.02	.14	.54[a]	.24
16. I believe that I was emotionally abused (e.g., “Why can’t you be more like me?”) and made to feel small and unimportant.	EA	.41	.04	.18	.62[a]	.33
17. Someone tried to make me do sexual things or watch sexual things.	SA	.27	.05	.84[a]	.13	.15
18. Someone took advantage of me sexually.	SA	.17	.03	.88[a]	.20	.07
19. My family was a source of strength and support.[b]	EN	.15	.19	.14	.28	.74[a]
20. You thought your parents wished you had never been born.	EA	.41	.06	.21	.47	.37
21. I got hit so hard by someone in my family that I had to see a doctor or go to the hospital.	PA	.73[a]	.10	.32	.05	.22
22. People in my family hit me so hard it left me with bruises or marks.	PA	.72[a]	.11	.24	.39	.15
23. I believe that I was sexually abused.	SA	.15	.09	.88[a]	.16	.16
24. There was someone to take me to the doctor if I needed it.[b]	PN	.23	.26	.08	–.03	.70
Percentage of variance		14.68	7.52	18.12	12.62	16.05
Scale internal consistency with retained items (β)		.87	.63	.94	.83	.86

NOTE: PA = physical abuse; PN = physical neglect; SA = sexual abuse; EA = emotional abuse; EN = emotional neglect.
a. Items were retained in the subscale in which they loaded.
b. Responses to the item were reverse scored.

TABLE 2. Descriptive Values for Study Variables

	%	*Range*	*Mean*	SD
Gender				
Male	61			
Female	39			
Race				
African American	88			
Other	12			
Age		13 to 16	14.55	1.04
Prior detention	60			
Current detention status				
Released to parents	75			
Transferred	25			
Substance use consequences		0 to 11	2.31	2.69
Self-restraint		1.67 to 5.00	3.15	.68
Emotional distress		1.17 to 4.42	2.51	.73
PA		1.00 to 5.00	2.07	
PN		1.00 to 3.00	1.38	
SA		1.00 to 5.00	1.49	
EA		1.00 to 4.00	2.24	
EN		1.00 to 4.00	2.10	
Any recidivism	46			
Times recidivated		0 to 6	.86	1.30
0	54			
1	26			
2	9			
3 or more	10			

NOTE: Percentiles for times recidivated do not sum to 100 due to rounding. PA = physical abuse; PN = physical neglect; SA = sexual abuse; EA = emotional abuse; EN = emotional neglect.

month follow-up period. The number of times recidivated ranged from 0 to 6; the frequencies for categories ≥ 3 were small and therefore were collapsed into a single category.

REGRESSION ANALYSES ON ANY RECIDIVISM

Results of the multivariate analyses that regressed the predictor variables on any recidivism are presented in Table 3. The first model revealed that three variables (race, prior detention, and substance use consequences) were significantly associated with any recidivism. Adolescents who were African American and who had a prior deten-

TABLE 3: Associations Between Any Recidivism and Other Study Variables

	B	SE	*Wald*	*Adjusted Odds Ratio*	*95% Confidence Interval*
Model 1					
Gender	.12	.35	.12	1.13	.57 to 2.23
Race	.95	.49	3.78	2.57	1.00 to 6.67*
Age	–.23	.15	2.39	.80	.60 to 1.06
Prior detention	.80	.30	6.93	2.22	1.23 to 4.02*
Case deferral	–.14	.38	.14	.87	.41 to 1.82
Substance use consequences	.12	.06	3.83	1.13	1.00 to 1.28*
Self-restraint	–.26	.26	1.00	.78	.47 to 1.28
Emotional distress	–.38	.23	2.78	.69	.44 to 1.07
Model 2					
PA	.08	.04	3.46	1.08	.99 to 1.17
PN	–.19	.08	5.52	.83	.70 to .97*
SA	.06	.04	2.95	.94	.87 to 1.01
EA	.01	.05	.03	.99	.91 to 1.09
EN	.10	.04	5.60	1.10	1.02 to 1.20*
Model 3					
Race	.87	.49	3.12	1.79	1.00 to 3.22
Prior detention	.87	.30	3.81	2.38	.91 to 6.25
Substance use consequences	.10	.06	3.33	1.11	1.03 to 1.19
PN	–.14	.08	3.02	.87	.74 to 1.02
EN	.10	.04	7.04	1.11	1.03 to 1.19**

NOTE: PA = physical abuse; PN = physical neglect; SA = sexual abuse; EA = emotional abuse; EN = emotional neglect.
*$p \leq .05$. **$p < .01$.

tion history were more likely to recidivate than were their counterparts. Adolescents who reported relatively more substance use consequences also reported relatively high rates of recidivism.

The second model presented in Table 3 included the five maltreatment variables. Although Physical Abuse and Sexual Abuse approached statistical significance, Physical Neglect and Emotional Neglect were the only maltreatment variables that were significantly associated with any recidivism. Respondents who reported more emotional neglect at baseline showed relatively high rates of recidivism. It was surprising that physical neglect was negatively associated with recidivism as respondents who reported less physical neglect showed relatively high rates of recidivism.

TABLE 4: Associations Between Times Recidivated and Other Study Variables

	B	SE	*Beta*	t
Model 1				
Gender	.03	.16	.01	.16
Race	.36	.20	.12	1.78
Age	–.03	.07	–.04	–.52
Prior detention	.45	.14	.22	3.23**
Case deferral	–.15	.18	–.06	–.84
Substance use consequences	.04	.03	.11	1.46
Self-restraint	–.09	.12	–.06	–.76
Emotional distress	–.12	.10	–.09	–1.17
Model 2				
PA	.02	.02	.13	1.18
PN	–.08	.04	–.16	–2.11*
SA	–.01	.02	–.07	–.83
EA	–.01	.02	–.07	–.68
EN	.05	.02	.20	2.36*
Model 3				
Prior detention	.42	.14	.21	3.05**
PN	–.06	.04	–.12	– 1.65
EN	.04	.02	.17	2.34*

NOTE: PA = physical abuse; PN = physical neglect; SA = sexual abuse; EA = emotional abuse; EN = emotional neglect.
*$p < .05$. **$p < .01$.

The third model included race, prior detention, substance use consequences, physical neglect, and emotional neglect as predictors. When simultaneously examining those variables that were significant in the two previous models, Emotional Neglect was the only factor that was still uniquely associated with any recidivism.

REGRESSION ANALYSES ON TIMES RECIDIVATED

Results of the analyses that regressed times recidivated on the predictor variables are presented in Table 4. The first model revealed that prior detention was the only sociodemographic or behavioral variable that was significantly associated with times recidivated. Adolescents who had been detained on at least one occasion prior to baseline showed more recidivism than did those whose first detention experience occurred at baseline.

Results of the second model indicated that the two neglect variables again emerged as significant predictors. Reports of relatively more emotional neglect and less physical neglect at baseline were associated with more times recidivated. None of the three abuse variables were significant predictors of times recidivated.

The third model included prior detention, emotional neglect, and physical neglect as predictors. When simultaneously including the variables that were significant in the two previous models, emotional neglect and prior detention were still uniquely associated with recidivism.

DISCUSSION

This study found that relatively high levels of emotional neglect were associated with recidivism during a 6-month follow-up period. These findings emerged through a series of regression analyses that considered single and multiple episodes of recidivism as well as the simultaneous influence of different sociodemographic, behavioral, and maltreatment variables. It is important that the findings converge with those from other studies that have found neglect to be more deleterious than abuse (Kaplan et al., 1999), including one indicating neglect to be a more potent predictor of recidivism (Dembo et al., 1998).

The specificity of the neglect-recidivism association was further illuminated through a set of supplemental analyses that were conducted using the total CTQ score as our measure of maltreatment. These analyses mirrored the third models in the two sets of analyses that were presented previously. In the supplemental analyses, the total CTQ score was substituted for the Emotional Neglect and Physical Neglect factors. Thus, the analysis of any recidivism included race, prior detention, substance use consequences, and the total CTQ score as predictors. The analysis of times recidivated included prior detention and the total CTQ score as predictors. Results indicated that the total CTQ score was not significantly associated with either recidivism variable, which in turn points to the need to consider emotional neglect specifically in future studies.

The findings from this study are noteworthy for illuminating differential effects of abuse and neglect on functioning in a highly vulnerable sample. Whereas the CTQ and similar instruments have allowed researchers to assess different types of maltreatment, there has been relatively little attention given to the differential effects of abuse and neglect in recent research. Given that neglect and abuse often co-occur within the same individuals, ascertaining their differential effects is challenging but still feasible (Kaplan et al., 1999). There is some suggestion that abuse and neglect have differential effects on psychosocial functioning, but there is little available data to evaluate this possibility much less to determine whether the differential effects on psychosocial functioning are linked to delinquency (Finzi, Ram, Har-Even, Shnit, & Weizman, 2001; Wolfe et al., 2001).

The findings are consonant with the extensive empirical literature that exists on familial or parental risk factors for adolescent delinquency. Research on this topic has revealed that adolescents who experience little or no monitoring from their parents are at increased risk for delinquency (Forehand, Miller, Dutra, & Chance, 1997; Steinberg, 1987; Weintraub & Gold, 1991). Low parental monitoring is manifested in many ways, including the failure to set or enforce standards of conduct. It is also manifested by parents who ignore their children's emotional needs, perhaps by failing to seek services when they are indicated. Logically, adolescents with prior detention experience who receive little supervision or care from their parents would be expected to be at high risk for recidivism. Future research should aim to assess both low parental monitoring and neglect within at-risk families to determine if they are differentially associated with delinquency and recidivism.

It is surprising that the bivariate analyses revealed that physical neglect was negatively associated with the two recidivism measures. It cannot be determined from the available data why emotional neglect and physical neglect were associated with recidivism in opposite ways. We do not believe that much weight should be given to the bivariate findings because physical neglect was not significantly associated with either measure of recidivism in the multivariate tests. If negative associations between physical neglect and recidivism continue to emerge in future studies, then attention can be given to the possibility the adolescents with physical neglect have lower rates of

recidivism because they receive relatively high levels of therapeutic attention in the juvenile justice system. In contrast, it seems likely that the needs of emotionally neglected adolescents are not being fully recognized and addressed in this system.

It is important to evaluate research findings based not only on statistical significance but also on effect sizes (Cohen, 1994; Rosnow & Rosenthal, 1989). Whereas emotional neglect was significantly associated with recidivism, its effect on the two criterion measures was relatively small. The meta-analysis cited previously also found that maltreatment had significant but modest effects on recidivism. More research is needed to evaluate the practical or clinical significance of the findings regarding the role of emotional neglect in predicting recidivism.

In addition, research should continue to examine the psychometric properties of the CTQ in diverse samples. Use of the CTQ is advantageous for studies that aim to assess emotional neglect because this form of maltreatment is less likely than other forms to be documented in official records. The factor loadings produced in this sample were substantially similar to those generated through other studies, which was impressive considering that the CTQ was initially developed for use in adult samples.

Future research can also expand on this study by examining emotional neglect in relation to problematic behaviors other than recidivism in this population. Adolescent detainees engage in high levels of risky health behaviors, including substance use, unprotected sex, and violence. Prior efforts to address these behaviors in juvenile justice samples have had limited success (Blechman, Hile, & Fishman, 2001; Cottle et al., 2001; Wilson et al., 2001). Future efforts may be more successful if they include assessments and interventions for emotional neglect.

There are several limitations with this study that should be acknowledged. Because the study was conducted in a single detention facility in the southeastern United States, the findings may not generalize to adolescent detainees sampled in other locations. Female and African American youths composed 14% and 40%, respectively, of all adolescents detained in residential facilities in the United States in 1997 (Gallagher, 1999), but they made up 39% and 88%, respectively, in this sample. Female youths were purposively oversampled to allow

for meaningful and valid analyses of the effects of gender. The proportion of African Americans in the sample was consistent with the overall racial composition of adolescent detainees at the holding facility where this study was conducted and partly reflected that this racial group constituted a majority of the overall population in the county where the study was conducted.

Limitations also existed in the methods used to assess maltreatment and recidivism. Both the self-report instrument used to assess maltreatment and the official records used to assess recidivism were imperfect. The self-report instruments were subject to various recall, interpretational, and presentational biases. These biases could have led to underreporting of maltreatment if the adolescents did not recall specific incidents of maltreatment, did not interpret abusive or neglectful treatment as such, or were wary of implicating their parents as perpetrators or themselves as victims. These biases could have led to overreporting if the adolescents exaggerated the hardships of their lives, although this seems unlikely.

The measure of recidivism was also inexact. Human errors in record keeping can affect the accuracy of official data. The measures of recidivism did not adjust for the time the adolescents may have been detained during the follow-up period because we did not have access to this information. It also did not reflect the behavior of the older adolescents because had they recidivated, the cases would have been processed in the adult system and we did not have access to adult records. The high rate of recidivism (46%) found in this study was partly a function of the sample being composed of relatively young adolescents as age tends to be inversely related to recidivism in samples of juvenile delinquents (Cottle et al., 2001).

Furthermore, the study was limited by its correlational research design. This limitation was partly offset by assessing the predictors before the adolescents had an opportunity to recidivate, thereby precluding the possibility that recidivism during the follow-up period influenced our measures of abuse and neglect. Nonetheless, even when correlational designs are prospective, they cannot control for the influence of unmeasured variables that could have confounded the association between emotional neglect and recidivism. To gain a better understanding about possible causal connections between these variables, future research can use experimental designs to investigate

whether risk for recidivism is reduced by participation in an intervention that is designed to offset the effects of emotional neglect.

In sum, this study yielded important findings that have potential implications for preventing recidivism among adolescent detainees. The findings can be used to design future studies and interventions with the ultimate aim of reducing offending behavior and health-related problems in a highly vulnerable population.

REFERENCES

Administration on Children, Youth, and Families, U.S. Department of Health and Human Services. (2001). *Child maltreatment, 1999*. Washington, DC: Government Printing Office.

Archwamety, T., & Katsiyannis, A. (1998). Factors related to recidivism among delinquent females at a state correctional facility. *Journal of Child and Family Studies*, *7*, 59-67.

Bernstein, D., Ahluvalia, T., Pogge, D., & Handelsman, L. (1997). Validity of the Childhood Trauma Questionnaire in an adolescent psychiatric population. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 340-348.

Bernstein, D., & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report manual*. San Antonio, TX: The Psychological Corporation.

Bernstein, D., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., & Wenzel, K., et al. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *American Journal of Psychiatry*, *151*, 1132-1136.

Blechman, E., Hile, M., & Fishman, D. (2001). Restorative justice and the prosocial communities solution. *Youth and Society*, *33*, 272-295.

Carrion, V., & Steiner, H. (2000). Trauma and dissociation in delinquent adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 353-359.

Cohen, J. (1994). The earth is round ($p < .05$). *Psychological Bulletin*, *49*, 997-1003.

Cottle, C., Lee, R., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior*, *28*, 367-394.

Dembo, R. (1995). On the poignant need for substance misuse services among youths entering the juvenile justice system. *International Journal of Addictions*, *30*, 747-751.

Dembo, R. (1998). Predictors of recidivism to a juvenile assessment center: An expanded analysis. *Journal of Child and Adolescent Substance Abuse*, *5*, 27-53.

Dembo, R., Schmeidler, J., Nini-Gough, B., Sue, C., Borden, P., & Manning, D. (1998). Predictors of recidivism to a juvenile assessment center: A three year study. *Journal of Child and Adolescent Substance Abuse*, *7*, 57-77.

Donovan, J., & Jessor, R. (1985). Structure of problem behavior in adolescence and young adulthood. *Journal of Consulting and Clinical Psychology*, *53*, 890-904.

Fink, L., Bernstein, D., Handelsman, L., Foote, J., & Lovejoy, M. (1995). Initial reliability and validity of the Childhood Trauma Interview: A new multidimensional measure of childhood interpersonal trauma. *American Journal of Psychiatry*, *152*, 1329-1335.

Finzi, R., Ram, A., Har-Even, D., Shnit, D., & Weizman, A. (2001). Attachment styles and aggression in physically abused and neglected children. *Journal of Youth and Adolescence*, *30*, 769-786.

Forehand, R., Miller, K., Dutra, R., & Chance, M. (1997). Role of parenting in adolescent deviant behavior: Replication across and within two ethnic groups. *Journal of Consulting and Clinical Psychology*, *65*, 1036-1041.

Gallagher, C. (1999). *Juvenile offenders in residential placement, 1997* (OJJDP Fact Sheet No. 96). Washington, DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.

Giblin, M. (2002). Using police officers to enhance the supervision of juvenile probationers: An evaluation of the Anchorage CAN Program. *Crime and Delinquency*, *48*, 116-137.

Harrison, P., Maupin, J., & Mays, G. (2001). Teen court: An examination of processes and outcomes. *Crime and Delinquency*, *47*, 243-264.

Hodges, K., & Kim, C. (2000). Psychometric study of the Child and Adolescent Functional Assessment Scale: Prediction of contact with the law and poor school attendance. *Journal of Abnormal Child Psychology*, *28*, 287-297.

Kaplan, S., Pelcovitz, D., & Labruna, V. (1999). Child and adolescent abuse and neglect research: A review of the past 10 years. Part 1: Physical and emotional abuse and neglect. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 1214-1222.

Kataoka, S., Zima, B., Dupre, D., Moreno, K., Yang, X., & McCracken, J. (2001). Mental health problems and service use among female juvenile offenders: Their relationship to criminal history. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 549-555.

Katsiyannis, A., & Archwamety, T. (1997). Factors related to recidivism among delinquent youths in a state correctional facility. *Journal of Child and Family Studies*, *6*, 43-55.

Kingree, J., Braithwaite, R., & Woodring, T. (2001). Psychosocial and behavioral problems in relation to recent experience as a runaway among adolescent detainees. *Criminal Justice and Behavior*, *28*, 190-205.

Latimer, W., Winters, K., & Stinchfield, R. (1997). Screening for drug abuse among adolescents in clinical and correctional settings using the Problem Oriented Screening Instrument for teenagers. *American Journal of Drug and Alcohol Abuse*, *23*, 79-98.

Myner, J., Santman, J., Capelletty, G., & Perlmutter, B. (1998). Variables related to recidivism among juvenile offenders. *International Journal of Offender Therapy and Comparative Criminology*, *42*, 65-80.

National Research Council. (1993). *Understanding child abuse and neglect*. Washington, DC: National Academy Press.

Perez, C., & Widom, C. (1994). Childhood victimization and long-term intellectual and academic outcomes. *Child Abuse and Neglect*, *18*, 617-633.

Quist, R., & Matshazi, D. (2000). The Child and Adolescent Functional Assessment Scale (CAFAS): A dynamic predictor of juvenile recidivism. *Adolescence*, *35*, 181-192.

Rosnow, R., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, *44*, 1276-1284.

Scher, C., Stein, M., Asmundson, G., McCreary, D., & Forde, D. (2001). The Childhood Trauma Questionnaire in a community sample: Psychometric properties and normative data. *Journal of Traumatic Stress*, *14*, 843-857.

Smith, C., & Thornberry, T. (1995). The relationship between childhood maltreatment and adolescent involvement in delinquency. *Criminology*, *33*, 451-477.

Steinberg, L. (1987). Familial factors in delinquency: A developmental perspective. *Journal of Adolescent Research*, *2*, 255-268.

Steiner, H., Cauffman, E., & Duxbury, E. (1999). Personality traits in juvenile delinquents: Relation to criminal behavior and recidivism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*, 256-262.

Towberman, D. (1994). Psychosocial antecedents of chronic delinquency. In N. Pallone (Ed.), *Young victims, young offenders: Current issues in policy and treatment* (pp. 151-164). New York: Haworth.

Weinberger, D., & Batholomew, K. (1996). Social-emotional adjustment and patterns of alcohol use among young adults. *Journal of Personality*, *64*, 495-527.

Weintraub, K., & Gold, M. (1991). Monitoring and delinquency. *Criminal Behavior and Mental Health*, *1*, 268-281.

Widom, C. (1989). The cycle of violence. *Science*, *244*, 160-166.

Widom, C., & Maxfield, M. (2001). *An update on the "cycle of violence"* (National Institute of Justice Research in brief). Washington, DC: National Institute of Justice.

Widom, C., Weiler, B., & Cottler, L. (1999). Childhood victimization and drug abuse: A comparison of prospective and retrospective findings. *Journal of Consulting and Clinical Psychology*, *67*, 867-880.

Wilson, J., Rojas, N., Haapanen, R., Duxbury, E., & Steiner, H. (2001). Substance abuse and criminal recidivism: A prospective study of adolescents. *Child Psychiatry & Human Development*, *31*, 297-312.

Wolfe, D., Scott, K., Wekerle, C., & Pittman, A. (2001). Child maltreatment: Risk of adjustment problems and dating violence in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 282-289.

Wright, K., Asmundson, G., McCreary, D., Scher, C., Hami, S., & Stein, M. (2001). Factorial validity of the Childhood Trauma Questionnaire in men and women. *Depression and Anxiety*, *13*, 179-183.

OFFENDER ETHNICITY AND MENTAL HEALTH SERVICE REFERRALS FROM JUVENILE COURTS

CAROLYN S. BREDA
Center for Mental Health Policy, Vanderbilt University

Juvenile courts play a critical role in facilitating access to mental health services for the large number of offenders with or at risk of emotional or behavioral problems. This research examines the decision of the court to refer offenders to treatment and whether offender ethnicity affects the referral decision. Results suggest that ethnicity has no independent effect on treatment recommendations. Rather, ethnicity modifies the effects of other variables, particularly legalistic variables such as petitioning and adjudication. Implications of the findings for research on ethnic disparities in juvenile court decision making as well as for Black offenders who are disproportionately involved in the juvenile justice system are discussed.

Keywords: race/ethnicity; juvenile offenders; service referrals; juvenile courts

Among the various decisions juvenile courts make is whether to refer a young offender to formal mental health services. This research examines the decision of juvenile court judges to refer offenders to treatment and whether offenders' ethnicity plays any part in the referral. The significance of the research is multifold. First, it

AUTHOR'S NOTE: *The author gratefully acknowledges the National Institute for Mental Health for funding the research through Grant No. MH54638-01A2, the Tennessee Council of Juvenile and Family Court Judges for making data available, and Dr. Warren Lambert and Dr. Mark Lipsey for their helpful comments and encouragement. Correspondence concerning this article should be addressed to Carolyn S. Breda, Center for Mental Health Policy, Vanderbilt University, 1207 18th Ave. S, Nashville, TN, 37212; e-mail: carolyn.breda@vandebilt.edu.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 30 No. 6, December 2003 644-667
DOI: 10.1177/0093854803256451

examines service referrals, an outcome previously overlooked in the literature, thereby adding to what we think we know about racial disparity in juvenile court decision making. Second, it uses a multiplicative statistical model that tests for interaction effects related to ethnicity, which prior research has not typically done. Third, it sheds light on the degree to which juvenile courts, historically rooted in therapeutic approaches to delinquency, make clinical interventions available to particularly at-risk social groups. Last, the research begins to specify the nature of the relationship between ethnicity and the courts' consideration of treatment options and in so doing, helps guide others who address ethnic disparity in court decisions in general and in decisions to refer offenders to services in particular.

CONTEXT OF MENTAL HEALTH REFERRALS THROUGH JUVENILE COURTS

In 1997, courts with juvenile jurisdiction disposed of nearly 1.8 million cases nationwide for delinquency alone (U.S. Department of Justice, 2000), representing more than one million individual juveniles who came into contact with the courts that year. Many of these youth have emotional or behavioral disorders. Estimates of co-occurring emotional disorder and delinquency within juvenile justice populations have ranged from 10% to 22% (Harstone & Cocozza, 1984; Otto, 1991, in Rogers, Powell, & Strock, 1998) to virtually 100% (e.g., McManus, Alessi, Grapentine, & Brickman, 1984; Rogers et al., 1998), depending on what definitions of offender and disorder are used (Fagan, 1991; Otto, Greenstein, Johnson, & Friedman, 1992). Moreover, young offenders' problems are not limited to externalizing crime-related behaviors but also include internalizing problems such as depression (Armistead, Wierson, Forehand, & Frame, 1992; Dembo, LaVoie, Schmeidler, & Washburn, 1987; McManus et al., 1984; Van Ness, 1984). These data indicate that clinical interventions would be appropriate for a very large number of youth engaged in the juvenile justice system.

Although many offenders need mental health services, juvenile offenders and their needs have generally been ignored (Cocozza, 1992; Knitzer, 1982). Several reasons may account for this inatten-

tion, including public stigmatization of those with mental health needs, a belief that offenders with mental health needs may not deserve therapeutic interventions, or an overall dissatisfaction with treatment models of jurisprudence (Otto et al., 1992). Regardless of the reasons, failure to deliver potentially effective treatment to this population can result in behaviors that continue to place the youth and their communities at risk and increase youth's chances of a criminal career and exposure to an already overburdened adult criminal justice system (Teplin, 2001).

Amid this general inattention to the mental health needs of offenders are data that suggest that a large number of young offenders have received services prior to their contact with the courts. In their review of the literature, Otto and colleagues (1992) reported that between 3% and 26% of youth within the juvenile justice system have been hospitalized and between 38% and 66% have received outpatient care prior to their court involvement. Rates vary by the study considered. These findings suggest that the courts are not necessarily or generally the first opportunity to respond to the mental health needs of juvenile offenders. Together, findings underscore the vital role courts can play in facilitating services either for offenders whose needs may not be addressed without court intervention or for offenders whose needs have not been adequately addressed previously in other service sectors.

One tool available to courts for facilitating services for offenders is a referral to treatment. Treatment referral is one of the decisions courts can make during the dispositional phase of proceedings. A referral decision can be the sole disposition or it can be combined with other outcomes such as restitution, community service, or probation. Data suggest that service referrals are most commonly (about 50% of the time) used in conjunction with probation, indicating that courts may use compliance with the referral as a condition of probation (Breda, 2001a). The authority of the courts to mandate treatment in this way further underscores its importance in effecting the delivery of services to youth in need.

ETHNIC BIAS IN COURT DECISION MAKING

If referring offenders to needed services is critical, so too is ensuring that all offenders with mental health needs have equal access to

services through the courts. For example, evidence suggests that the need for therapeutic services is at least as great for Black offenders as for White offenders (Martin & Grubb, 1990). Glisson (1996) found that about 80% of youth placed in state custody for delinquency had levels of symptomatology that would indicate a need for treatment and that symptomatology did not vary by the youth's race. Rogers et al. (1998) found that Black and White juvenile offenders incarcerated in a short-term correctional facility showed similar rates of internalizing behavior, although Black youth were more likely than White youth to be diagnosed with conduct disorder. Clinician observations have also suggested that the need for psychiatric help is at least as great for Black delinquents as for White delinquents (Lewis, Balla, & Shanok, 1979).

A search of the empirical literature on race and its effect on juvenile courts' treatment referral decisions suggests this area has been neglected. What research has been done has been based on restrictive samples of offenders, is dated, or does not consider ethnic differences in referrals from juvenile courts. For example, Rogers et al. (1998) found that 9% of Black offenders and 13% of White offenders were referred for clinical evaluation after they were incarcerated in a short-term correctional facility. However, this research did not control for confounding effects of other variables that can affect referral decisions such as the nature of the juveniles' offense. Glisson (1996) found that service referrals for youth already in state custody were unaffected by race; however, the results were not broken down by whether the reason for custody was delinquency or dependency neglect.

Cohn (1963) examined correlates of probation officers' recommendations for psychiatric assessments, but the effects of race are not reported clearly. Black offenders may have been referred less often than were White offenders because they demonstrated fewer personality problems, which did affect referral. The small number of Black offenders in Cohn's referral group ($N = 8$) and the use of bivariate analyses (the standard when the article was published) also limit findings. Finally, Breda (1999) found that the overall rate of treatment referrals through juvenile courts was about 3%; however, ethnic differences in referral rates were not considered.

Research on juvenile offender's ethnicity within the mental health profession has found that therapists underestimate the need for treat-

ment after Black offenders are referred to them by the courts (Lewis et al., 1979). Others (Warren, Jackson, Nugaris, & Farley, 1973) have found that Blacks are less often accepted for therapy, more often assigned to inexperienced therapists, and seen for shorter periods of time. However, these studies do not address the role of ethnicity in the initial decision of the courts to recommend services for juvenile offenders.

Although empirical studies of ethnic disparity in courts' treatment referral decisions remain scarce, a great deal of research has been conducted on other court decisions such as detention and adjudication. Results have been mixed and often contradictory, with some studies showing harsher outcomes for Black youth (Dannefer & Schutt, 1982; Frazier & Bishop, 1985; Marshall & Thomas, 1983; McCarthy & Smith, 1986; Thomas & Cage, 1977; Thomson & Zingraff, 1981), some showing harsher outcomes for Whites (Scarpitti & Stephenson, 1971), and others finding no racial differences (Bell & Lang, 1985; Cohen & Kluegel, 1978; Horwitz & Wasserman, 1980; Minor, Hartmann, & Terry, 1991; Niarhos & Routh, 1992). In short, the weight of the evidence on ethnic disparity in juvenile court decision making shows no clear trend (Fagan, Slaughter, & Harstone, 1987; Marshall & Thomas, 1983).

The lack of consistent findings may be due to a variety of methodological shortcomings such as inadequate measures (Bishop & Frazier, 1988), failure to control for salient covariates or to consider multiplicative effects (Bell & Lang, 1985; Bishop & Frazier, 1988; Cohen & Kluegel, 1978; Myers & Tallarico, 1986), small or restrictive samples (Bishop & Frazier, 1988), lack of comparable data across multiple jurisdictions (Sampson & Laub, 1993), and focus on a single court encounter or a single decision point within an encounter, which cannot detect any cumulative effect of race (Bishop & Frazier, 1988; McCarthy & Smith, 1986; Thornberry & Christenson, 1984). This study overcomes several of these issues by examining a wide range of offenders across multiple jurisdictions using multivariate techniques that allow the assessment of direct, indirect, and moderating effects of ethnicity on treatment referrals through juvenile courts.

THEORETICAL EXPECTATIONS

Previous research on race bias in court outcomes has generally adopted a conflict perspective (Chambliss & Seidman, 1971; Leiber, 1994; Quinney, 1970). Hypothetically, Blacks would receive harsher dispositions than would Whites because they lack the social, economic, or political resources to resist them (Bell & Lang, 1985; Marshall & Thomas, 1983; Quinney, 1970; Schur, 1971) or they pose a threat to White hegemony, which must be suppressed by the dominant, White group (Frazier, Bishop, & Henretta, 1992).

Alternatively, Martin and Grubb (1990) have suggested that differential experiences by Blacks and Whites may be less a function of overt racism in the system than of more symbolic, cultural differences between groups. For example, data have shown that symptoms that would have been recognized as pathological in White juvenile offenders were either ignored or incorrectly assessed in Black juveniles by predominantly White mental health professionals (Lewis et al., 1979). Such disparities are attributed to ethnocentrism among White professionals whose clinical standards for assessment exclude "the social reality of Blackness" (Martin & Grubb, 1990, p. 264). In sum, any ethnic disparity in mental health–related decisions, including service referral decisions made by predominantly White juvenile court judges, may stem from either cross-cultural misunderstanding or subtle or overt racism.

Based on the prior research on ethnic disparity in decision making, it is hypothesized in this research that juvenile courts will be more likely to refer White than Black offenders to mental health services. This hypothesis rests on the following four theoretical assumptions: (a) Treatment may represent a valuable and scarce resource that judges (who are predominantly White) may reserve for White offenders, (b) juvenile courts may view mental health services as a lenient yet still legitimate sanction they can apply to White offenders, (c) court officials may think that White youth are more suitable candidates for rehabilitation than are their Black counterparts, and (d) community-based treatment requires financial resources that White offenders may be in a better position than Black offenders to obtain.

METHOD

SAMPLE

The sample includes more than 33,000 youth between the ages of 5 and 18 who were referred to any one of the 98 juvenile courts throughout Tennessee in 1997 for a criminal or status offense (*Juvenile Court Data*, 1997). The courts preside over urban, suburban, and rural areas throughout the state, with metropolitan courts likely to be representative of other metropolitan courts in the United States (Stapleton, Aday, & Ito, 1982). The racial, gender, and offense-type distributions among referrals in Tennessee closely resemble those in the nation (U.S. Department of Justice, 1990).

This study is based only on White and Black offenders. Black juveniles compose 36% of the sample; Whites compose the other 64%. Offenders are included regardless of how far they penetrated the system. So for example, youth who were released upon intake are included as are those who proceeded to a judicial hearing. This avoids a sample selection bias that occurs when court decisions that restrict the range (and alter the composition) of the sample are examined (Frazier & Bishop, 1985).

MEASURES

Six legal factors that have been used routinely in studies of court decision making are assessed. Current offense is coded as status offense, proceedings violation (e.g., of probation), illegal conduct (e.g., disorderly conduct), alcohol or drug related, property, or violent (person), with the most serious offense coded in the case of multiple offenses. Prior record indicates whether youth had a previous encounter with the courts during the target year. Police indicates who took the youth to court—a police officer or someone else, notably, a family member or school official. Detention indicates whether youth were detained at any time prior to adjudication. Petition indicates whether a formal petition was filed against the youth at intake. Adjudication status identifies youth who were not recommended for judicial review, those who had judicial review but were found not to be delinquent, and those who were adjudicated delinquent during the review.

Social variables most commonly used in prior research are included, specifically, youth's age, gender, and ethnicity (Marshall & Thomas, 1983). Age includes four categories to approximate developmental groupings—youth younger than 10 years, those 10 to 12, those 13 to 15, and youth 16 to 18. The youth's living arrangement has been less used in studies of court outcomes, although some (Barton, 1976; Thomas & Cage, 1977) suggest it may be relevant for decision making. Here, youth who lived with both biological parents at the time they were taken to court, in a blended household (e.g., with a biological and stepparent), with a single parent, with relatives, or in some other living arrangement (e.g., group home) are compared. A direct measure of social class is unavailable. The database also does not include a measure of offenders' mental health status. Although it is reasonable to assume that mental health status would affect referral decisions, others (e.g., Glisson, 1996; Kelley, 1978) have found that service-related decisions for offenders are largely unaffected by information on their clinical profiles.

The dependent variable, treatment referral, is a dichotomous variable (referred versus not referred) that identifies whether youth were referred for mental health counseling (more than 98% of all treatment referrals), placed in a private mental health facility, or voluntarily placed with the Department of Mental Health and Mental Retardation (DMHMR). The few youth committed involuntarily by the courts to the Department of Mental Health and Mental Retardation are excluded from the treatment referral group because their circumstances of treatment are different and interest is in the largely noncustodial group of offenders referred to court. Other dispositions (e.g., probation, restitution, or community service) can accompany a treatment referral. Here, referral is indicated if any of multiple dispositions included a referral.

All variables are "dummy" coded so that for each variable, there is one missing category or referent group in the analyses to which results for the included categories or groups are compared. For example, gender is coded as 1 if offenders are male and as 0 if they are female. Results that relate to gender, then, will show the difference in the odds of being referred to services for male youth compared with female youth, who constitute the missing or referent group.

DESIGN AND PROCEDURE

The design for the analysis is based on the work of Cohen and Kluegel (1978), who offer the following three criteria for identifying racial bias in court outcomes: (a) race directly affects outcome, (b) race indirectly affects outcome through other factors considered stereotypical in nature, or (c) race moderates the effect of another variable.

The second and third criteria may warrant some elaboration. Consider the following premise: Youth raised by a single parent are more prone toward criminality than are youth raised by both biological parents. Blacks, who are more likely than Whites to live in single-headed households, may experience discriminatory outcomes because of preconceived notions among court officials about youth raised by single parents. Thus, the effect of race is indirect through its relationship to a variable—living arrangement—that can stereotype a particular social group (Cohen & Kluegel, 1978; Schur, 1973). The third criterion is met, for example, when the effect of any variable (e.g., offense type) on the referral decision is found to differ for Black and White offenders. In this case, race is said to moderate the effect of offense type on outcome.

Logistic regression is used to assess the additive and multiplicative effects of ethnicity on the courts' decision to refer juvenile offenders to services. Two equations are modeled. First, treatment referrals are modeled by youth's social and legal profiles. This allows a direct test of whether offenders' ethnicity affects the chances of referral while simultaneously controlling for other potentially relevant social and legal variables. It also provides a basis for determining whether ethnicity indirectly affects the referral decision through an association with living arrangement. Second, interaction terms are included in the model to assess whether ethnicity modifies the effects of other social or legal variables on the courts' decision to recommend services.

The significance of direct, indirect, or moderating effects (shown as superscripts in columns 2 and 4 of Table 2) is based on alpha levels adjusted for the number of variables in each model. This approach accounts for the possibility that with multiple significance tests (i.e., variables in the model), some variables may appear significant by chance alone. For example, with 20 significance tests, as in the first

model, we expect one variable to be significant by chance. The adjusted alpha for each of the two models presented in Table 2 is $p < .003$ and $p < .001$, respectively.

RESULTS

Table 1 shows results from a series of cross-tabulations between offender ethnicity and other social and legal variables. Black juvenile offenders seem to be somewhat less likely than White offenders to be referred to formal mental health services—2.4% of Blacks compared with 3.3% of Whites. Blacks also differ from Whites on most social and legal variables frequently relevant for court decisions. For example and consistent with other research, Black youth are significantly more likely than White youth to be brought to court for more serious offenses, especially violent crimes against persons and to be brought to court by police rather than by family or school officials. Blacks are also more likely to have a prior record and to have been detained prior to adjudication. However, Whites are more likely than Blacks to have a formal petition filed against them at intake and to be adjudicated delinquent during judicial review. Socially, Black offenders tend to be somewhat younger than are Whites and are less likely to live with both biological parents. Notably, however, living with a single parent is the most common living arrangement for offenders of both ethnic groups. Although a few differences between Black and White offenders (e.g., in referral rates) may seem substantively small, all meet chi-square tests for statistically significant differences at $p < .001$, which can be expected with such a large sample.

Table 2 shows the results of a two-step logistic regression that assesses the effects of ethnicity on the treatment referral decision while controlling for potentially confounding social and legal covariates. The first step assesses the direct effect of ethnicity on treatment referrals, with results shown in columns 1 and 2. The second step assesses moderating effects of ethnicity, with results shown in columns 3 and 4.

Column 2 shows that the log odds of referral are somewhat greater for Black offenders than for White offenders, increasing by a factor of 1.20, when all other variables are held constant. However, the signifi-

TABLE 1: Cross-Tabulations Between Ethnicity and Other Social and Legal Factors (*N* = 33,349)

	Black (36%)	*White (64%)*
Treatment referral rate	2.4%	3.3%
Social profile		
Gender		
Male	69%	66%
Female	31%	34%
Age		
5 to 9 years	1%	1%
10 to 12 years	6%	4%
13 to 15 years	33%	25%
16 to 18 years	60%	70%
Living arrangement		
Both biological parents	12%	35%
Biological & stepparent	7%	11%
One parent	65%	44%
Other relatives	12%	5%
Other	4%	5%
Legal profile		
Offense type		
Status offense	25%	27%
Proceeding violation	3%	4%
Illegal conduct	24%	27%
Alcohol- or drug-related offenses	9%	14%
Property	24%	20%
Person	15%	8%
Police	79%	62%
Prior record	34%	23%
Detention	55%	12%
Petition	38%	63%
Adjudication status		
No adjudicatory hearing	5%	12%
Hearing but not adjudicated	68%	34%
Hearing and adjudicated	27%	54%

NOTE: All ethnic differences meet chi-square tests of significance at $p < .001$.

cance of this effect does not meet the Bonferroni-adjusted alpha level ($p < .003$). Moreover, youth living in single-headed households, which could indirectly affect the chances for Black youth to be referred to services, are as likely to be referred as are youth living with both biological parents, Exp(B) = 1.25. The small proportion of youth who live in other living arrangements are about half as likely to be recommended for treatment than youth living with both biological par-

TABLE 2: Logistic Regression of Referral to Mental Health Services on Social and Legal Characteristics of Juvenile Offenders

	Direct Effects		*Direct + Modifying Effects*	
	B	*Exp*(B)	B	*Exp*(B)
Social variables				
Ethnicity				
Black	.18	1.20	–.56	.57
Gender				
Male	–.04	.96	–.07	.93
Age				
5 to 9 years	.14	1.16	–.48	.62
10 to 12 years	.77	2.16*	.74	2.10**
13 to 15 years	.56	1.75*	.53	1.70**
Living arrangement				
Biological parent & stepparent	.29	1.34	.31	1.36
One parent	.22	1.25	.22	1.25
Other relatives	.04	1.04	–.04	.96
Other	–.62	.54*	–.55	.58
Legal variables				
Offense type				
Proceeding violation	–.80	.45*	–.51	.60
Illegal conduct	–1.13	.32*	–1.36	.26**
Alcohol- or drug-related offenses	–.98	.38*	–.92	.40**
Property	–1.05	.35*	–.93	.39**
Person	–.16	.85	.30	1.34
Prior record	.58	1.79*	.69	1.99**
Police	–.41	.66*	–.18	.83
Detention	–.10	.90	.29	1.34
Petition	1.83	6.24*	1.11	3.02**
Adjudication status				
No adjudication	.64	1.90*	.60	1.82**
Adjudicated delinquent	.47	1.60*	.70	2.01**
Interaction terms				
Black × Male			.08	1.09
Black × Younger Than 10			1.43	4.20
Black × 10 to 12			.17	1.18
Black × 13 to 15			.17	1.18
Black × Biological Parent & Stepparent			–.27	.76
Black × One Parent			.01	1.01
Black × Relatives			.20	1.22
Black × Other			–.52	.60
Black × Proceeding Violation			–1.35	.26**
Black × Illegal Conduct			.20	1.22
Black × Alcohol- or Drug-Related Offenses			–.69	.50
Black × Property			–.61	.54
Black × Person			–1.61	.20**

TABLE 2: (continued)

	Direct Effects	*Direct + Modifying Effects*	
	B	*Exp*(B)	B	*Exp*(B)
Black × Prior Record			–.32	.73
Black × Police			–.69	.51**
Black × Detention			–.39	.68
Black × Petition			2.29	9.84**
Black × No Adjudication			–.29	.75
Black × Adjudicated Delinquent			–.79	.45**
Constant	–5.03		–4.73	
Nagelkerke *R*-squared	.16		.19	
Df	20		39	

NOTE: The referent group for each Variable is as follows: White (ethnicity), female (gender), 16 to 18 years (age), both biological parents (living arrangement), no prior record in 1997 (prior record), offender referred to court by others (police), not detained prior to adjudication (detention), status offense (offense type), not formally petitioned (petition), and adjudicatory hearing held but youth not adjudicated delinquent (adjudication status).
*$p < .003$ (Bonferroni adjusted). **$p < .001$ (Bonferroni adjusted).

ents, Exp(B) = .54. However, as shown in column 4 of the table, this effect (Black × Other Living) does not vary significantly by youth's ethnic background.

Legal circumstances are far more relevant for treatment referral decisions than is youth's ethnicity. Of the six legal factors considered, the following five are significant: offense type, prior record, whether police take the youth to court, whether a petition is filed, and adjudication status. For example, the odds for referral almost double for offenders with person offenses and offenders with status offenses (the referent group) compared with juveniles with more intermediate types of offenses. A prior record increases the odds of referral by about 80%, Exp(B) = 1.79. Being taken to court by police rather than by parents or school authorities reduces the chances of referral by somewhat less than half, Exp(B) = .66, controlling for other variables such as offense type.

Decisions of the courts prior to final disposition also have a strong effect on service referrals, specifically decisions associated with intake and adjudication. Youth formally petitioned at intake are more than six times as likely to be referred to services than those not peti-

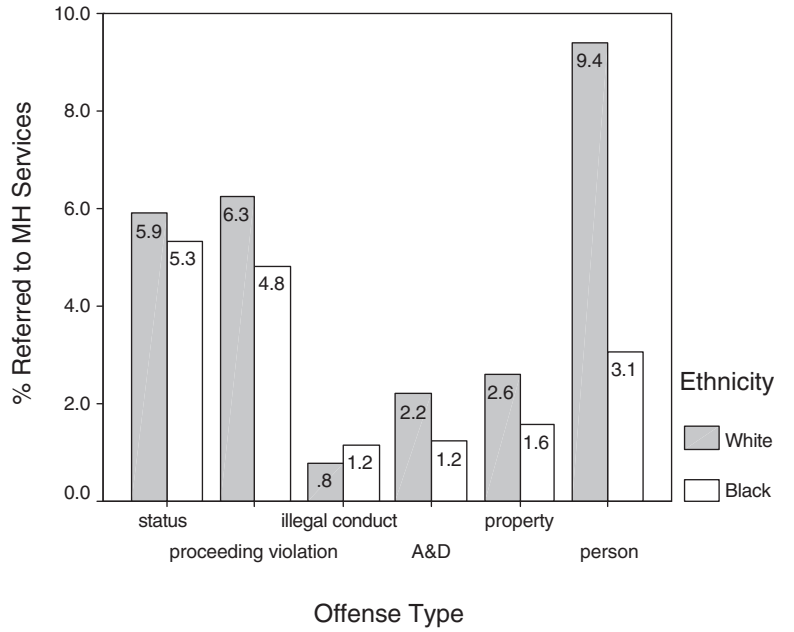

Figure 1: Implication of Offense Type for Treatment Referral

tioned by the courts. All else being equal, youth who leave the courts without a judicial hearing, Exp(B) = 1.90, and those who proceed to the hearing phase and are found to be delinquent, Exp(B) = 1.60, are significantly more likely to be sent for services than those who are found not to be delinquent during judicial review.

Thus far, the findings do not fulfill Cohen and Kluegel's (1978) first two criteria for racial bias in court decisions. The question remains whether ethnicity modifies any effect other variables have on the referral decision. Columns 3 and 4 of Table 2 show that the effect on referral decisions of legal variables, notably, offense type, police involvement, petitioning, and adjudication, depends significantly on whether offenders are White or Black.

For example, Black and White violent offenders are significantly more likely than intermediate types of offenders to be referred for counseling. However, Black violent offenders are about 20% less

likely than White violent offenders to be referred (Black × Person in Table 2). Figure 1 graphically illustrates the point. Although the figure does not control for all the covariates in the regression model, it suggests that about 3% of Black violent offenders are referred for treatment compared with about 9% of their White counterparts.

The ethnic difference in referrals among violent offenders may be partially attributable to differences in the specific kind of violent crime Whites and Blacks allegedly commit (Breda, 2001b). Whites represent 64% of the total sample and 48% of the violent offenders. Yet, they constitute 87% of sexual abuses against children, 68% of aggravated rapes, and 59% of other sex offenses. The availability of specialty mental health services such as sex offender programs for rapists and sex offenders may account for the courts' proclivity to refer these types of offenders, who are disproportionately White, to treatment. This possibility remains tentative, however, because of the small number of cases available for significance testing once referral, ethnicity, and specific offense are considered together. Among the most common specific offenses against persons—aggravated assaults and assaults—Blacks are still less likely than Whites to be referred to treatment. A three-way cross-classification (which does not control for all the covariates in the regression model) suggests that referral rates are about three times as high for White assaulters as for Black assaulters, $\chi^2(1, n = 2{,}183) = 27.61$, $p < .01$. A four-way cross-tabulation shows that referral rates are higher for White assaulters whether the youth has a prior record, $\chi^2(1, n = 573) = 6.30$, $p = .01$, or not, $\chi^2(1, n = 1{,}610) = 26.91$, $p < .01$.

Another significant interaction is found between offender ethnicity and adjudication status. The direct effect reported earlier suggests that offenders, Black and White, who are adjudicated delinquent during the judicial review or who are not recommended for judicial review are more likely than those reviewed but not adjudicated delinquent to be referred for services. However, the implication of adjudication status for treatment referral differs for Blacks and Whites. Blacks adjudicated delinquent are about half as likely, $\text{Exp}(B) = .45$, as their White counterparts to be considered for mental health services (Black × Adjudicated Delinquent). As shown in Figure 2, this effect appears to be particularly relevant in the case of violent offenses when about

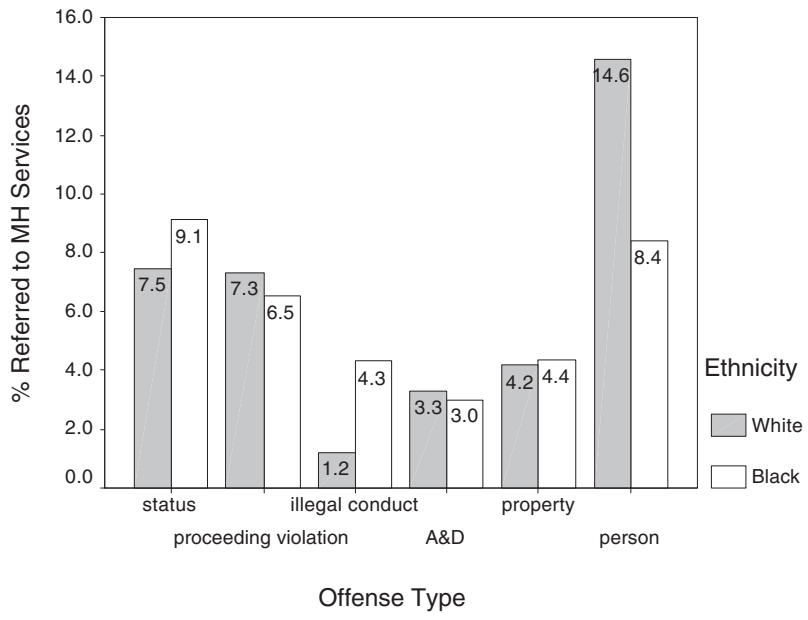

Figure 2: Implication of Adjudication for Treatment Referral (*N* = 14,796)

15% of Whites adjudicated for an offense against a person are referred compared with about 8% of their Black counterparts.

Another key finding is the significant interaction between ethnicity and the courts' decision at intake to file a formal petition. Filing a formal petition rather than handling the case more informally increases the odds for treatment referral for White and Black offenders. However, these odds are significantly greater for Blacks than for Whites (Black × Petition). Figure 3 suggests that this effect largely pertains to offenders who are petitioned for status offenses when the rate of treatment referral for Blacks is about 20% compared with 8% for Whites.

Finally, the chances for treatment referral are reduced for White and Black offenders when police are involved at intake. However, the odds are even lower for Blacks in this situation when about one Black youth is considered for services for every two Whites, Exp(B) = .51.

Overall, these significant interaction effects suggest that ethnicity is significant for courts' consideration of formal mental health ser-

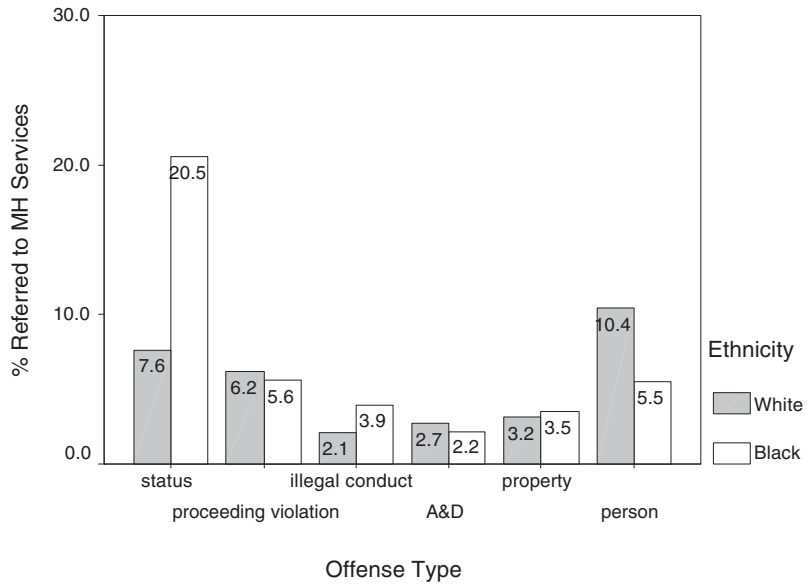

Figure 3: Implication of Petitioning for Treatment Referral (*N* = 17,950)

vices for offenders; however, its significance depends on legal considerations associated with the case. Moreover, any differential effect of ethnicity on treatment referral does not disadvantage one group exclusively.

DISCUSSION

This research addresses the important question of ethnic disparity in the juvenile courts' decision to refer offenders to treatment. In its design, the study is able to overcome some of the major limitations of earlier work by using a multivariate analysis of direct, indirect, and ethnicity-based interaction effects for a large sample of various types of offenders across multiple jurisdictions.

The results suggest that decisions to refer offenders to services are not substantially affected by offenders' ethnicity, at least not in a direct or simple way. Rather, as others have found regarding other types of

court outcomes, legalistic factors explain the treatment referral decision more directly than does youth's social capital. Specifically, type of offense, prior record, police involvement, petitioning, whether to hold a judicial hearing, and if so, whether to officially adjudicate the youth delinquent significantly affect the courts' referral decision. Among social variables, only age is relevant, with offenders between 10 and 15 years of age most likely to be referred once other factors are taken into account. Youth's living arrangement, a variable that may stereotype Black offenders who disproportionately live with single parents, is unrelated to a referral outcome.

Although offenders' ethnicity does not substantially affect service referral decisions in a direct way, the data show that ethnicity conditions the effect other variables, largely legal variables, have on the referral decision. Referrals are influenced strongly by offense type, whether police are involved in the proceedings, petitioning, and adjudication. But the effect these four variables have on the courts' decision to refer for services depends in part on whether the offenders are White or Black. So whereas the findings fail to support Cohen and Kluegel's (1978) first two criteria for racial bias in court outcomes, they support the third criterion—Ethnicity does moderate the effect of other variables.

For example, violent offenders (and status offenders) are significantly more likely than intermediate types of offenders to be referred for counseling. However, among violent offenders, the referral rate for Whites to Blacks is roughly 5 to 1. Serious White offenders disproportionately commit sex-related offenses for which specialized mental health services exist. This may account for courts' greater use of service options for them. On the other hand, Blacks are still less likely than Whites to be referred for the most common types of person offenses, including aggravated assaults. Perhaps other variables not available to the study may account for this discrepancy. For example, the measure of prior record used here considers previous court encounters within the year of the study; therefore, it cannot assess whether ethnic differences in more extended offense histories might help explain the lesser use of a therapeutic approach for serious Black offenders. Perhaps future research can address this possibility. Until then, this study finds that courts do not consider treatment options for

violent offenders similarly for otherwise comparable Black and White offenders.

The odds of being referred for services decline significantly for Blacks and Whites but especially for Blacks when police rather than others take the youth to court. Perhaps police are less inclined than others to either request services for youth or be amenable to courts' inclinations to use treatment options. Or, it may be that courts tend to forego treatment approaches when police are involved in the complaint. For Black offenders in particular, this finding suggests the significant role police play in decisions made by the courts (Dannefer & Schutt, 1982) and supports other research that finds police practices vary by offenders' ethnicity, with harsher responses directed toward Black youth (Dannefer & Schutt, 1982; Ferdinand & Luchterhand, 1970; Thornberry, 1973).

Several scholars (Bishop & Frazier, 1988; Fagan et al., 1987; McCarthy & Smith, 1986; Thornberry & Christenson, 1984) note that the significance of ethnicity for court outcome may be tied to other decisions made throughout the judicial process. This study supports this view. At the outset, Black offenders taken to court by police are significantly less likely than Whites to be referred to treatment. Then at intake, a formal petition increases the chances of treatment referral significantly more for Blacks than for Whites. Subsequently, at adjudication, a finding of guilt decreases the chances for Blacks while increasing the chances for Whites. The decision to detain is an exception and does not seem to affect referral decisions significantly for either Blacks or Whites. Overall, however, a court's final disposition, at least that of service referral, seems to be a function of inextricable relationships between offender ethnicity and the courts' administrative proceedings. Future research should continue to consider multiple decision points in the judicial process.

The heightened effect of petitioning on referrals for Black offenders, specifically for Black status offenders, can be elaborated. Most (about 40%) of the Black offenders in the sample who are younger than 10 are status offenders. It may be that the higher referral rate for these youth reflects the courts' effort to provide early therapeutic interventions to the youngest of offenders while offenses are minor for that social group at greatest risk for involvement with the juvenile jus-

tice system (U.S. Department of Justice, 1999). The finding that courts tend to use treatment options more for the youngest offenders when Black than when White supports this possibility, although this effect (Black × 5 to 9) fails to meet the adjusted level for statistical significance.

As discussed at the outset, conflict perspectives have been used often to predict ethnic differences in court outcomes. In this study, as in others before it, findings are mixed. In several circumstances, ethnicity has no bearing on court outcome; in others, service referrals are less likely for Blacks; and in others, they are more forthcoming for Blacks. These mixed findings as well as the absence of a strong, independent effect of ethnicity on the referral decision suggest that courts may not overtly discriminate in their consideration of treatment options. Instead, any bias may be subtle rather than direct, bound up as it is here in a complex web of interrelationships involving offender ethnicity, offense type, police involvement, petitioning, and adjudication status.

Effort must continue to identify factors that can account for ethnic bias in court decision making. This study highlights some legal variables relevant for the decision to refer offenders to services, but others can be considered. Offenders' demeanor (Bell & Lang, 1985), socioeconomic status (McCarthy & Smith, 1986; Thomas & Cage, 1977), or the relationship between the victim and the offender (Thomson & Zingraff, 1981) may help to account for results found here. For example, predominantly White judges may conclude from the demeanor of certain Black offenders that they are unsuitable candidates for treatment. Future research that can incorporate such measures as these as well as data from courts presided over by Black judges may help further specify conditions that seem to differentially affect court outcomes for Black and White offenders.

Data on the mental health status of offenders not available in this dataset and generally unavailable to juvenile courts prior to disposition (Breda, 2000) are also needed to understand better the impact of health status on the courts' service-related decisions. A few studies (e.g., Glisson, 1996; Kelley, 1978) that have examined the relationship between offenders' mental health status and service decisions surprisingly have found little relationship between the two variables.

Much more research is needed before we can understand fully the impact of mental health status on court outcomes and whether any effect this status may have varies by the ethnicity of offenders.

Future research on court outcomes may also want to consider statistical models more complex than ones used here. For example, although two-way interactions were formally tested, additional analyses suggest that three-way (or higher order) interactions may better fit the data (e.g., Ethnicity × Offense × Other Court Decisions). Furthermore, it is important that a multiplicative model was necessary to observe the impact of ethnicity on the referral outcome. More complex multiplicative models will likely be required to specify more fully the conditions in which the ethnicity of offenders affects decisions of the courts.

This research suggests that courts underuse treatment options for Black and White offenders given estimates of the need for services for both groups within the juvenile justice population. The courts' limited use of treatment options may be especially problematic for Blacks, however, who are consistently overrepresented at all stages of the juvenile justice system including arrests, court processing, and confinement (Allen-Hagan, 1991; McGarrell, 1993; U.S. Department of Justice, 1999). Community-based services may help to circumvent the escalation of involvement in the system for this particularly at-risk group and reduce the personal, familial, and social costs associated with emotional disorder and delinquency.

Limited use of service options for Black (and other) offenders also raises critical questions about barriers that may impede courts' consideration of therapeutic options. Barriers may be pragmatic, such as the cost and financing of formal services or the lack of appropriate treatment programs in a community. Although limited research suggests that the availability of services within a community has little bearing on juvenile courts' referrals to formal mental health services (Breda, in press), more research in this area is needed.

Barriers to treatment referrals through the courts also may be ideological. Currently, courts operate in a political environment that calls for harsher sentencing for juvenile offenders and referrals to adult courts (Schwartz, 1989), despite data that show the rate of serious and violent crime among juveniles has declined during the past 5 years (Dodge, 1999; U.S. Department of Justice, 1999). Mental health

approaches may seem too lenient a response in an era of "get-tough" public policy. Given the large number of offenders with emotional or behavioral disorders, it would seem prudent and compassionate to identify and remove any barriers that diminish the capacity of the courts to facilitate services for a troubling and troubled group of youth.

REFERENCES

Allen-Hagan, B. (1991). *Children in custody 1989: Public juvenile facilities* (OJJDP Update on Statistics). Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Armistead, L., Wierson, M., Forehand, R., & Frame, C. (1992). Psychopathology in incarcerated juvenile delinquents: Does it extend beyond externalizing problems? *Adolescence*, *27*, 309-314.

Barton, W. (1976). Discretionary decision-making in juvenile justice. *Crime and Delinquency*, *22*, 470-480.

Bell, D., Jr., & Lang, K. (1985). The intake dispositions of juvenile offenders. *Journal of Research in Crime and Delinquency*, *22*, 309-328.

Bishop, D. M., & Frazier, C. E. (1988). The influence of race in juvenile justice processing. *Journal of Research in Crime and Delinquency, 25*, 242-263.

Breda, C. (1999). Determinants of juvenile court decisions to use mental health interventions. In J. Willis, C. Liberton, K. Kutash, & R. Friedman (Eds.), *The 11th annual research conference proceedings, a system of care for children's mental health: Expanding the research base* (pp. 267-273). Tampa: University of South Florida.

Breda, C. (2000). The organizational context of courts' treatment referrals for juvenile offenders. In J. Willis, C. Liberton, K. Kutash, & R. Friedman (Eds.), *The 12th annual research conference proceedings, a system of care for children's mental health: Expanding the research base* (pp. 157-162). Tampa: University of South Florida.

Breda, C. (2001a). [Treatment referrals with probation]. Unpublished data.

Breda, C. (2001b). [Ethnicity and specific offenses]. Unpublished data.

Breda, C. (in press). The impact of external environment on service-related decisions of juvenile courts. In J. Willis, C. Liberton, K. Kutash, & R. Friedman (Eds.), *The 13th annual research conference proceedings, a system of care for children's mental health: Expanding the research base.* Tampa: University of South Florida.

Chambliss, W. J., & Seidman, R. (1971). *Law, order, and power.* Reading, MA: Addison-Wesley.

Cocozza, J. J. (1992). Introduction. In J. Cocozza (Ed.), *Responding to the mental health needs of youth in the juvenile justice system* (pp. 1-6). Seattle: The National Coalition for the Mentally Ill in the Criminal Justice System.

Cohen, L. E., & Kluegel, J. R. (1978). Determinants of juvenile court dispositions: Ascriptive and achieved factors in two metropolitan courts. *American Sociological Review*, *43*, 162-176.

Cohn, Y. (1963). Criteria for the probation officer's recommendations to the juvenile court judge. *Crime and Delinquency*, *9*, 262-275.

Dannefer, D., & Schutt, R. K. (1982). Race and juvenile justice processing in court and police agencies. *American Journal of Sociology*, *87*, 1113-1132.

Dembo, R., LaVoie, L., Schmeidler, J., & Washburn, M. (1987). The nature and correlates of psychological emotional functioning among a sample of detained youths. *Criminal Justice and Behavior*, *14*, 311-334.

Dodge, D. C. (1999). Early prevention and intervention equals delinquency prevention. Research and Training Center on Family Support and Children's Mental Health, Portland, OR. *Focal Point*, *13*, 17-21.

Fagan, J. (1991). Community-based treatment for mentally disordered juvenile offenders. *Journal of Clinical Child Psychology*, *20*, 42-50.

Fagan, J., Slaughter, E., & Harstone, E. (1987). Blind justice? The impact of race on the juvenile justice process. *Crime and Delinquency*, *33*, 224-258.

Ferdinand, T. N., & Luchterhand, E. G. (1970). Inner-city youths, the police, the juvenile court, and justice. *Social Problems*, *17*, 510-526.

Frazier, C. E., & Bishop, D. M. (1985). The pretrial detention of juveniles and its impact on case dispositions. *Journal of Criminal Law and Criminology*, *76*, 1132-1152.

Frazier, C. E., Bishop, D. M., & Henretta, J. C. (1992). The social context of race differentials in juvenile justice dispositions. *The Sociological Quarterly*, *33*, 447-458.

Glisson, C. (1996). Judicial and service decisions for children entering state custody: The limited role of mental health. *Social Science Review*, *70*, 257-281.

Harstone, E., & Cocozza, J. J. (1984). Providing services for the mentally ill, violent juvenile offender. In R. Mathias, P. DeMuro, & R. A. Allison (Eds.), *Violent juvenile offenders: An anthology* (pp. 157-175). Newark, NJ: National Council on Crime and Delinquency.

Horwitz, A., & Wasserman, M. (1980). Formal rationality, substantive justice, and discrimination. *Law and Human Behavior*, *4*, 103-115.

Juvenile court data [Data file]. (1997). Nashville: Tennessee Council of Juvenile and Family Court Judges.

Kelley, T. M. (1978). Clinical assessment and the detention, disposition, and treatment of emotionally disturbed delinquent youths. *Journal of Criminal Justice*, *6*, 315-327.

Knitzer, J. (1982). *Unclaimed children: The failure of public responsibility to children and adolescents in need of mental health services.* Washington, DC: Children's Defense Fund.

Leiber, M. (1994). A comparison of juvenile court outcomes for Native Americans, African Americans and Whites. *Justice Quarterly*, *11*, 257-279.

Lewis, D. O., Balla, D. A., & Shanok, S. S. (1979). Some evidence of race bias in the diagnosis and treatment of the juvenile offender. *American Journal of Orthopsychiatry*, *49*, 53-61.

Marshall, I. H., & Thomas, C. W. (1983). Discretionary decision-making and the juvenile court. *Juvenile and Family Court Journal*, *34*, 47-59.

Martin, T. W., & Grubb, H. J. (1990). Race bias in diagnosis and treatment of juvenile offenders: Findings and suggestions. *Journal of Contemporary Psychotherapy*, *20*, 259-272.

McCarthy, B. R., & Smith, B. L. (1986). The conceptualization of discrimination in the juvenile justice process: The impact of administrative factors and screening decisions on juvenile court dispositions. *Criminology*, *24*, 41-64.

McGarrell, E. F. (1993). Trends in racial disproportionality in juvenile court processing: 1985-1989. *Crime and Delinquency*, *39*, 29-48.

McManus, M., Alessi, N. E., Grapentine, W. L., & Brickman, A. (1984). Psychiatric disturbance in serious delinquents. *Journal of the American Academy of Child Psychiatry*, *23*, 602-615.

Minor, K. I., Hartmann, D. J., & Terry, S. (1991). Predictors of juvenile court actions and recidivism. *Crime and Delinquency*, *43*, 328-344.

Myers, M. A., & Tallarico, S. M. (1986). The social contexts of racial discrimination in sentencing. *Social Problems*, *33*, 236-251.

Niarhos, F. J., & Routh, D. K. (1992). The role of clinical assessment in the juvenile court: Predictors of juvenile dispositions and recidivism. *Journal of Clinical Child Psychology*, *21*, 151-159.

Otto, R. K., Greenstein, J. J., Johnson, M. K., & Friedman, R. M. (1992). Prevalence of mental disorders among youth in the juvenile justice system. In J. Cocozza (Ed.), *Responding to the mental health needs of youth in the juvenile justice system* (pp. 7-48). Seattle, WA: The National Coalition for the Mentally Ill in the Criminal Justice System.

Quinney, R. (1970). *The social reality of crime*. Boston: Little, Brown.

Rogers, K. M., Powell, E., & Strock, M. (1998). The characteristics of youth referred for mental health evaluation in the juvenile justice system. In J. Willis, C. Liberton, K. Kutash, & R. Friedman (Eds.), *The 10th annual research conference proceedings, a system of care for children's mental health: Expanding the research base* (pp. 329-334). Tampa: University of South Florida.

Sampson, R. J., & Laub, J. H. (1993). Structural variations in juvenile court processing: Inequality, the underclass, and social control. *Law & Society Review*, *27*, 285-311.

Scarpitti, F. R., & Stephenson, R. M. (1971). Juvenile court dispositions: Factors in the decision-making process. *Crime and Delinquency*, *17*, 142-151.

Schur, E. M. (1971). *Labeling deviant behaviors*. New York: Harper & Row.

Schur, E. M. (1973). *Radical non-intervention: Rethinking the delinquency problem*. Englewood Cliffs, NJ: Prentice Hall.

Schwartz, I. (1989). Juvenile crime: The youth corrections agenda for the 1990s. In C. Russell & C. McCauley (Eds.), *Crime and justice in the 1990s* (pp. 45-50). Nashville, TN: Vanderbilt Institute for Public Policy Studies, Vanderbilt University.

Stapleton, V. D., Aday, D., & Ito, J. A. (1982). An empirical typology of American metropolitan juvenile courts. *American Journal of Sociology*, *88*, 549-564.

Teplin, L. A. (2001). *Assessing alcohol, drug, and mental disorders in juvenile detainees* (OJJDP Fact Sheet No. 2). Washington, DC: U.S. Department of Justice.

Thomas, C. W., & Cage, R. J. (1977). The effect of social characteristics on juvenile court dispositions. *The Sociological Quarterly*, *18*, 237-252.

Thomson, R. J., & Zingraff, M. T. (1981). Detecting sentencing disparity: Some problems and evidence. *American Journal of Sociology*, *86*, 869-880.

Thornberry, T. (1973). Race, socioeconomic status and sentencing in the juvenile justice system. *The Journal of Criminal Law and Criminology*, *64*, 90-98.

Thornberry, T. P., & Christenson, R. L. (1984). Juvenile justice decision-making as a longitudinal process. *Social Forces*, *63*, 433-444.

U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention. (1990). *Juvenile court statistics, 1988*. Washington, DC: Government Printing Office.

U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention. (1999). *Minorities in the juvenile justice system* (National Report Series: Juvenile Justice Bulletin). Washington, DC: Government Printing Office.

U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention. (2000). *Juvenile court statistics, 1997*. Washington, DC: Government Printing Office.

Van Ness, S. R. (1984). Rape as an instrumental violence: A study of youth offenders. *Journal of Offender Counseling, Services, & Rehabilitation*, *9*, 161-170.

Warren, R. C., Jackson, A. M., Nugaris, J., & Farley, G. K. (1973). Differential attitudes of Black and White patients toward treatment in a child guidance clinic. *American Journal of Orthopsychiatry*, *43*, 384-393.

THE EFFECT OF MORAL RECONATION THERAPY ON THE RECIDIVISM OF YOUTHFUL OFFENDERS

A Randomized Experiment

TODD A. ARMSTRONG
Arizona State University West

A randomized experimental design was used to test the effect of moral reconation therapy on the recidivism rates of youth offenders. The sample consisted of youth offenders who were incarcerated in a county jail in Maryland ($N = 256$). A survival analysis compared the risk of recidivism of the treatment and control groups. The risk of recidivism for the treatment group was not significantly different from the risk of recidivism for the control group. A supplemental analysis compared a high implementation treatment group with the randomized control group. Individuals in the high implementation treatment group received a minimum of 30 days of treatment. Consistent with earlier results, group differences were not statistically significant.

Keywords: corrections; treatment; recidivism; cognitive behavioral

Accumulating meta-analytical evidence indicates treatment can lead to a substantial reduction in recidivism. Lösel (1995) reviewed 13 meta-analyses of rehabilitation programs published between 1985 and 1995. The meta-analyses included in Lösel's review assessed a variety of treatment modalities. Studies included in these meta-analyses used samples of both juvenile and adult offend-

AUTHOR'S NOTE: *This project was supported by Grant No. DLE-97-362, awarded by the Bureau of Justice Assistance, Office of Justice Programs, U.S. Department of Justice. The Assistant Attorney General, Office of Justice Programs coordinates the activities of the following program offices and bureaus: Bureau of Justice*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 30 No. 6, December 2003 668-687
DOI: 10.1177/0093854803256452

ers. According to Lösel, the overall mean effect size for correctional treatment programs was .l0.[1] Gendreau, Smith, and Goggin (2000) noted "that this would mean that the recidivism rates for the treatment and control groups would be 45% and 55% respectively" (p. 18). Meta-analyses subsequent to Lösel's review have continued to demonstrate that treatment can reduce recidivism (Andrews, Dowden, & Gendreau, 1999; Redondo, Sanchez-Meca, & Garrido, 1999).

Meta-analyses have found that the involvement of the researcher in program implementation is an important predictor of the effectiveness of treatment. Lipsey (1999a) stated, "Specifically, these earlier investigations found that the extent to which the researcher was involved in the design, delivery, and supervision of treatment was one of the strongest correlates of the size of the effect on recidivism, with greater researcher involvement associated with larger effects" (p. 615). Meta-analytical evidence demonstrating the relationship between researcher involvement and program effect suggests the portability of programs needs to be carefully considered. If program effects do not generalize from carefully controlled trials in which the researcher is heavily involved to "real-world" settings, then the policy implications of carefully controlled trials are substantially diminished.

This study tests the portability of the moral reconation therapy (MRT) program. Such tests are particularly important in light of MRT's widespread popularity. A Web page describing the program notes that

> MRT is a theoretical approach that has been shown to reduce recidivism . . . from 25% to 60% in more than 30 states and in Ontario, Canada and Puerto Rico. MRT is also used system-wide in the states of Washington and Oklahoma, and in Oregon's Washington County. (Better People, 2000, p. 1).

Assistance, Bureau of Justice Statistics, National Institute of Justice, Office of Juvenile Justice and Delinquency Prevention, and the Office of Victims of Crime. Points of view or opinions contained within this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice. *Correspondence concerning this article should be addressed to Todd A. Armstrong, Administration of Justice Department, Arizona State University West, P.O. Box 37100, Phoenix, AZ 85069-7100; e-mail: toddarmstrong@asu.edu.*

Published evaluations of MRT based on two trials conducted by the authors have found that exposure to the MRT program reduces recidivism (Little & Robinson, 1989; Little, Robinson, & Burnette, 1990; Little, Robinson, & Burnette, 1991; Little, Robinson, & Burnette, 1993). However, meta-analytical evidence finding a relationship between evaluator involvement and effect size indicates that the ability of others to effectively implement the MRT program should be carefully considered before strong conclusions regarding the efficacy of this program can be drawn.

THE MRT PROGRAM

MRT is based on a simplified personality theory that combines elements from Erikson and Loevinger's ego development, Maslow's hierarchy of needs, Kohlberg and Piaget's moral development theories, and the work of Carl Jung (Little & Robinson, 1988). This personality theory is presented to clients through lecture, individual discussion, treatment manuals, and workbooks. According to the authors, "The personality theory proposes that people form their personalities through a progressive accumulation of beliefs, attitudes, and habits that layer themselves over the 'inner self,' the essential essence of the person" (Little & Robinson, 1988, p. 139). This perspective posits that criminal and delinquent behaviors are "defense mechanisms" that are a reaction to tension that is a product of conflict between the personality and the inner self.

The primary goal of MRT is the moral development of the treatment client. The therapy identifies nine stages of moral development and explains that these stages exist in a continuum. These stages are paralleled by a series of treatment steps. It is anticipated that as treatment results in moral development and individuals progress through the treatment steps, they will begin to act in a manner consistent with more sophisticated levels of moral reasoning.

There appears to be a reasonable amount of confluence between MRT and the characteristics of effective treatment that have been identified by meta-analyses. A number of meta-analyses have found cognitive behavioral treatments are one of the most effective treatment types (Andrews et al., 1999; Antonowicz & Ross, 1994; Garrett,

1985; Izzo & Ross, 1990; Lipsey, 1999b; Redondo et al., 1999). MRT contains both cognitive and behavioral elements. Little and Robinson (1988) explained that MRT "strives to reeducate clients socially, morally and behaviorally and to instill appropriate goals, motivation and values" (p. 136). In addition to a cognitive-behavioral focus, MRT attempts to develop ability to delay gratification and encourages clients to end relationships with delinquent peers. Andrews et al. (1999) found that program elements focused on self-control skills and reducing antisocial peer associations were significantly related to the magnitude of treatment effects.

Little and Robinson (1988) described some of the specific MRT activities targeted at facilitating moral development, increasing self-control, and reducing association with delinquent peers. To facilitate moral development, clients are given moral dilemmas. Clients first consider these dilemmas individually and then discuss them in groups. In the context of group discussion, clients are exposed to higher levels of moral reasoning. Written exercises are used to evaluate peer and familial relationships. These exercises require the development of a specific plan to terminate relationships with delinquent peers. To increase self-control, the program has time-based requirements for several steps in the treatment process. Time periods are relatively short early in the treatment, but as the treatment progresses, they are gradually lengthened.

EVIDENCE ON THE EFFICACY OF MRT

The authors of the MRT program have conducted two quasi-experimental trials. The results of these trials show that exposure to MRT results in a reduction in recidivism (Little & Robinson, 1989; Little et al., 1990, 1991, 1993). These trials were conducted in the Shelby County Corrections Center in Memphis, Tennessee, beginning in 1987. All participants in the first trial were incarcerated at some point during 1987 and/or 1988 and were released during those years. The first trial compared 70 male inmates who were exposed to MRT with a control group of 82 male inmates. Both groups were composed of drug users who had volunteered for treatment. The authors state that the control group consisted of individuals who "did not participate in

treatment because funding limited numbers of clients" (Little et al., 1993, p. 1152) but do not discuss in detail how individuals were selected for treatment from the larger pool of volunteers. The comparability of the two groups is addressed in the second of two articles on this trial. Little et al. (1993) stated that the treatment and control groups did not differ on age, race, length of sentence, or prior arrest records. Individuals assigned to the treatment group participated in an average of 31.4 group and/or individual MRT treatment sessions during incarceration. To assess the effectiveness of MRT, recidivism data were collected in April 1991. Recidivism was defined as arrest followed by a conviction for which jail or prison time was levied and served. The study found that 24.3% of the treatment group had recidivated compared with 36.6% of the control group (Little et al., 1991). These differences were not statistically significant. The impact of MRT on the 5-year rate of return to prison was also assessed (Little et al., 1993). This comparison found that 37.1% of the treatment group had been reincarcerated, whereas 54.9% of the control group had been reincarcerated. These differences were statistically significant.

The second trial of the MRT program began in 1988 (Little & Robinson, 1989; Little et al., 1990). In this trial, the sample consisted of individuals who had been incarcerated for driving while intoxicated in the Shelby County Corrections Center at any time from February 1988 to January 1989. Both treatment and control group members were selected from a group of volunteers. The treatment group was composed of 115 inmates, whereas the control group was composed of 65 inmates. Participants were required to have at least 30 days remaining in their sentence. Although the authors did not explicitly state how individuals were assigned to treatment and control groups, they explained, "When more applications were received than bed space was available, decisions for entry were based upon maintaining a racial balance . . . thus the control and treated groups were formed semirandomly" (Little et al., 1990, p. 1382). Two comparisons were made to assess the effectiveness of MRT. First, the rates of reconviction for the treatment and control groups were compared. Data for this comparison were collected in February of 1989. Data showed that 20% of the experimental group had been reconvicted compared with 27.6% of the control group. The second comparison

contrasted the rates of reincarceration for the treatment and control groups. Data for this comparison were collected in February of 1990. In this case, data showed that 13.9% of the treatment group had been reincarcerated compared with 21.5% of the control group.

While methodological considerations such as the quasi-experimental design and the failure to include relevant control variables temper the confidence that may be placed in these results, these trials find that exposure to the MRT program can lead to a substantial decrease in recidivism. However, even if one accepts these trials as evidence of the program's efficacy, it does not lead to the conclusion that MRT should be widely implemented. Lipsey (1999a) found that on average, real-world programs were only half as effective in reducing recidivism as were "demonstration" projects. This shows that once the efficacy of a rehabilitation program has been established by a demonstration project, the ability of criminal justice system personnel to effectively implement that program needs to be tested prior to the widespread implementation of the program.

Tests of the ability of criminal justice system personnel to implement the MRT program in real-world settings are sorely lacking. Studies conducted by the authors of the program have been used as the primary empirical justification for MRT's widespread implementation. Meta-analytical evidence linking investigator involvement to program success demonstrates that tests of the portability of MRT are needed. Therefore, this study explores the effect of the MRT program on the risk of recidivism when criminal justice system personnel are responsible for the implementation of the MRT program.

METHOD

In this study, MRT was implemented by criminal justice system personnel in the Montgomery County Detention Center beginning in January of 1997. MRT was implemented as the core component of the Youthful Offender Unit (YOU). The YOU housed approximately 40 offenders. It was created as a response to increases in disorderly conduct associated with youth offenders at the Montgomery County Detention Center.

PARTICIPANTS

Participants included 256 male residents of the detention center who were incarcerated between January 1, 1997, and January 1, 1998. The age of participants ranged from 15 to 22 years ($M = 20.21$, $SD = .99$). The racial distribution of the sample was 55% African American ($n = 142$), 32% Caucasian ($n = 82$), 6% Hispanic ($n = 14$), and 7% Asian ($n = 18$). The majority of the participants (98.8%, $n = 253$) had been arrested prior to the arrest that led to their current incarceration and participation in the study. Approximately 21% ($n = 54$) of participants had four or more prior arrests. Of those who had any prior arrest, 43% ($n = 110$) had a prior arrest for violence, 48% ($n = 123$) had a prior arrest for a property offense, and 32% ($n = 82$) had a prior arrest for a drug offense.

MEASURES

Variables included in this study were the number of disciplinary violations while incarcerated, age, race, the number of total prior offenses, the number of prior offenses by offense type (violent, property, drug, and other), and the length of time until recidivism. All variables were quantified using official data. Variables for the number of disciplinary violations while incarcerated, age, and race were quantified using official records from the jail. Variables for the number of total prior offenses, the number of prior offenses by offense type, and the length of time until recidivism were quantified using official records from the jail and information gathered from the Federal Bureau of Investigation's National Crime Information Center database.

PROCEDURE

To test MRT offenders between the ages of 15 and 22, participants were randomized into treatment and control groups. Participants assigned to the treatment group were housed in the YOU. Participants assigned to the control group were housed in the general population. Thus, age and residence in the jail were the only eligibility criteria. Initial randomization included all youthful offenders residing in the jail who met the age requirement. Subsequent to this initial randomization, new inmates who were younger than 22 were randomized as

they entered jail. The randomization of new inmates occurred after intake but prior to permanent residential assignment. Subsequent to randomization, individuals were asked to participate in the study. Participation was requested after randomization to avoid biasing the treatment group in favor of a treatment effect.

Study participants who were randomized into the treatment condition were housed in the YOU and were exposed to the MRT program. On average, YOU residents were exposed to three sessions of MRT per week. These sessions lasted from 1 to 1½ hours. Staff members who had been trained in the MRT program taught weekly sessions. Those teaching weekly sessions included correctional counselors and corrections officers. In addition to weekly sessions, YOU residents were also indirectly exposed to the MRT program. The behavior of YOU residents was often described in the context of the level of moral reasoning that it reflected. Individuals were encouraged to display behavior that was reflective of higher stages of moral reasoning and discouraged to display behavior that was reflective of lower stages of moral reasoning.

RESULTS

Table 1 shows the correspondence between randomization and residential assignment. Of the 256 participants included in the experiment, 129 were randomized into the treatment group and 127 were randomized into the control group. Of the 129 randomized into the treatment group, 19 were never housed in the YOU and consequently were never exposed to treatment. This group of 19 resulted from the refusal of treatment ($n = 4$), exclusion due to an inability to speak English ($n = 4$), or the release of participants prior to transfer into the YOU ($n = 11$). Additional exceptions to the randomization protocol occurred in the control group. Of the 127 participants randomized into the control group, 25 were exposed to treatment in the YOU. This group of 25 included 21 participants who were residents of the dorm prior to its conversion into the YOU and 4 participants who were placed in the YOU due to concern that they may be victimized if placed in the general population.

TABLE 1: Fidelity of Placement

	Placement		
Assignment	*YOU*	*General Population*	*Total*
Randomized into			
YOU	110	19[a]	129
General population	25[b]	102	127
Total	135	121	256

NOTE: YOU = the Youthful Offender Unit.
a. Of the 19 participants, 4 were excluded due to an inability to speak English, 4 refused treatment, and 11 were released prior to transfer.
b. The 25 participants included 21 individuals who resided in the dorm that was converted into the YOU and 4 individuals who were placed in the YOU due to safety concerns.

To assess the effect of exceptions to the randomization protocol, two sets of analyses were estimated. The first compared the 129 participants who were randomized into the treatment group with the 127 participants who were randomized into the control group. This comparison is problematic in that it potentially attenuates treatment effects. In this comparison, 19 participants included in the treatment group never received treatment and 25 participants in the control group were actually in the YOU. If treatment is effective, the exposure of participants randomized into the control group to treatment and the lack of exposure to treatment for participants randomized into the treatment group will attenuate between group differences in recidivism. To address this possibility, a second set of analyses was conducted. This second set of analyses excluded those participants who were exceptions to randomization. These analyses contrasted the 110 participants who were randomized into the treatment group and exposed to the MRT program while residing in the YOU with the 102 participants who were randomized into the control group and housed in the general population. This comparison maximizes treatment-control group differences in exposure to MRT. The results of the two sets of analyses were not different. Therefore, only the results for models comparing participants who were treated in a manner consistent with randomization are reported. Hereafter, the 110 participants who were randomized into the treatment group and exposed to the MRT program while residing in the YOU are referred to simply as the treatment

TABLE 2: The Initial Equivalence of the Treatment and Control Groups

	*Treatment (*n = *110)*			*Control (*n = *102)*		
Variable	M	SD	n	M	SD	n
Race (%)						
African American[a]	.67	.46	74	.48	.50	49
Caucasian[a]	.22	.42	24	.41	.49	42
Hispanic	.06	.25	7	.07	.21	7
Asian	.05	.20	5	.04	.19	4
Age (years)	21.52	.98	110	21.39	.99	102
Prior arrests for						
Violence	.51	.69	45	.55	.80	41
Property offenses	.77	1.23	51	.71	.92	48
Drug offenses	.49	.87	33	.47	.78	34
Other offenses	.73	1.23	43	.63	.92	41

a. Groups were significantly different at $p < .05$.

group. Similarly, the 102 participants who were randomized into the control group and housed in the general population are referred to as the control group.

Analysis began with an assessment of the comparability of the treatment and control groups. A one-way analysis of variance was used to compare the age, race, prior violent arrests, prior property arrests, prior drug arrests, and other prior arrests of the groups. Means and standard deviations are presented in Table 2. Significant group differences were found for the percentages of African Americans and Caucasians. Tests for group differences using a single categorical indicator of race indicated that overall racial differences between the treatment and control groups were not significant.

Survival analysis was used to examine treatment and control group differences in the risk of recidivism. The survival function is an estimate of the probability of survival (not recidivating) to time t. The hazard function is an estimate of the conditional probability of recidivism occurring in any specified time interval $(t, t + dt)$ given survival to time t. The effect of treatment was first assessed with the nonparametric life table estimator of the risk of recidivism. This estimator allows the comparison of the treatment group's conditional probability of recidivism to the control group's conditional probabil-

ity of recidivism. The life table estimator of the conditional probability of recidivism during any time interval, given survival to the beginning of that interval, is

$$\hat{q} = \frac{d}{n - m/2}$$

where d is the number of individuals who recidivated during the interval, n is the number of participants entering the interval, and m is the number of participants censored during the interval.

Subsequently, a proportional hazards model was used to examine relationships in a multivariate context. A number of different nonparametric and parametric forms are available for survival analysis. A proportional hazards model was selected as this model makes the least restrictive assumptions about the distribution of the outcome measure (Schmidt & Witte, 1988).[2] Cox (1972) explained that this model assumes that the hazard function of all individuals differs only by a factor or proportionality. The hazard function of the proportional hazards model assumes the following form

$$\lambda(t, z) = \lambda 0(t) e^{z\beta},$$

where z is a vector of participant characteristics. The proportional hazards model assumes that the ratio of the hazard functions of any pair of participants will remain constant in time. If λ_0 is constant for all t, the proportional hazards model reduces to an exponential failure time model.

Table 3 presents data on the number of recidivists, exposure time, and survival time by group. Final recidivism data were collected on April 28, 1999. Exposure time is the length of time from first release until the end of data collection. Survival time is the length of time from first release until failure or the end of data collection, whichever came first.

Figure 1 shows the life table survival curves for the treatment and control groups. Data were grouped in 30-day intervals. The treatment group appears to be slightly less likely to have recidivated for the first 450 days of exposure, after which the control group is less likely to

TABLE 3: The Number of Recidivists, Exposure Time, Survival Time, and Treatment Length for Treatment and Control Groups

	Treatment Group (n = 110)	*Control Group (n = 102)*
Number of recidivists	71 (64.54%)	66 (64.71%)
Exposure time (days)		
Mean	563.41	616.98
Median	568	632
Survival time (days)		
Mean	307.64	295.60
Median	258	228

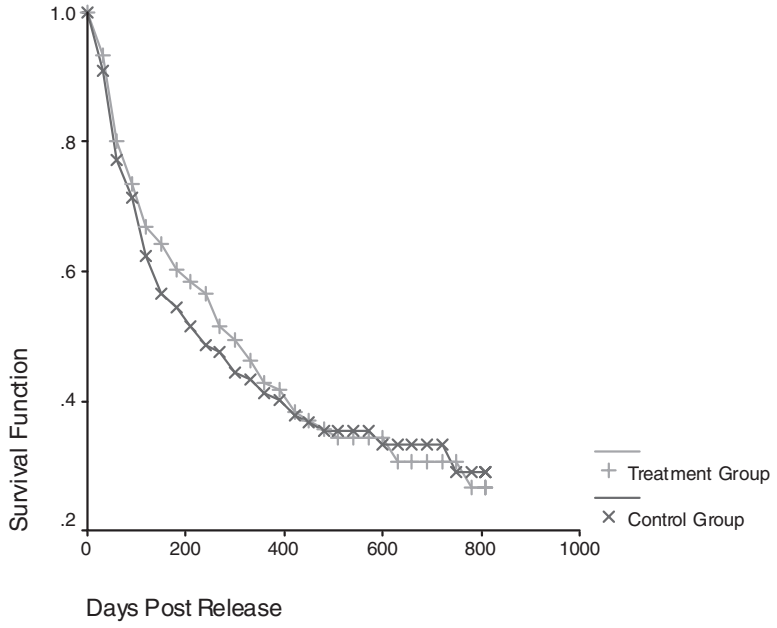

Figure 1: Probability of Postrelease Survival in Days for Treatment and Control Groups

have recidivated. Differences in the cumulative likelihood of recidivism are not significant as measured by the Wilcoxon test statistic (.42, $p = .52$).

TABLE 4: Results From the Proportional Hazards Model Testing the Effect of MRT on Recidivism

Variable	*Coefficient*	p *Value*
MRT treatment	.13	.47
Disciplinary violations	.09	.26
Prior offenses	.07	.06
Age	–.12	.19
Race (White or non-White)	–.12	.53

NOTE: MRT = moral reconation therapy.

In the next stage of analyses, a proportional hazards model was estimated. This model included control variables. The variables included were the number of disciplinary violations while incarcerated; age on April 28, 1999; a binary indicator of ethnicity (coded 1 if the participants were Caucasian and 0 if they were African American, Hispanic American, or Asian American); and the number of total prior offenses. The results of this model are presented in Table 4. None of the variables in the model, including the indicator of treatment, were significantly related to the risk of recidivism. The model does not offer a significant improvement over a naive model that assumes the explanatory variables contain no reliable information about participants' risk of recidivism.[3]

A second set of analyses comparing a high implementation treatment group with the control group explored the possibility that the inclusion in the treatment group of a number of individuals who were only briefly exposed to treatment attenuated treatment effects. Although mean and median treatment lengths—77 days and 61 days, respectively—indicate that the treatment group as a whole received a substantial amount of treatment, 39% (n = 43) of those who were included in the treatment group received less than 30 days of treatment. The second set of analyses compared a high implementation group composed of individuals who received a minimum of 30 days of treatment (the high implementation treatment group) with those who were randomized into the control group and did not receive treatment.

Table 5 presents information on the number of recidivists, exposure time, and survival time by group. Life table survival curves and a proportional hazards model were estimated to test the effect of membership in the high implementation control group. Figure 2 presents the

TABLE 5: Number of Recidivists, Exposure Time, and Survival Time of High Implementation Treatment and Control Groups

High Implementation	*Treatment Group (n = 65)*	*Control Group (n = 102)*
Number of recidivists	37 (56.92%)	66 (64.71%)
Exposure time (days)		
Mean	496.11	616.98
Median	495	632
Survival time (days)		
Mean	311.75	295.60
Median	268	228

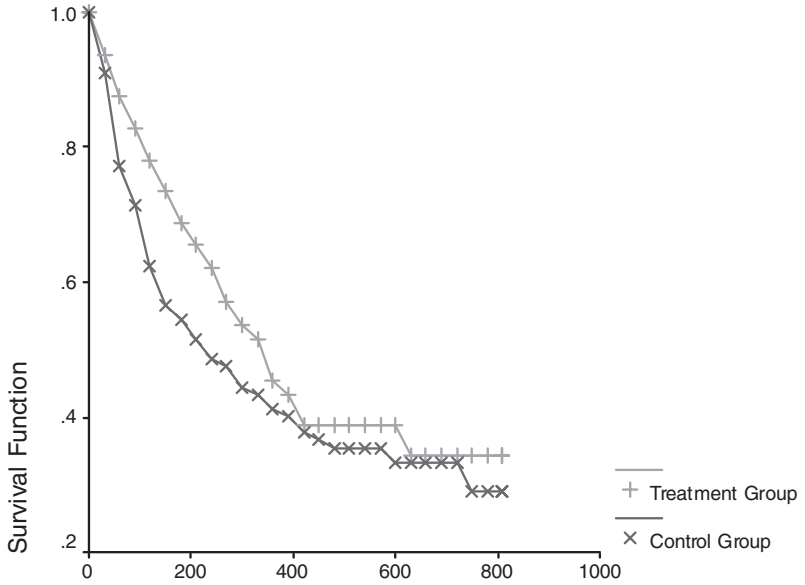

Figure 2: Probability of Postrelease Survival in Days for High Implementation Treatment and Control Groups

life table survival curves for the high implementation treatment and control groups. Data were grouped in 30-day intervals. The treatment group appears to be slightly less likely to recidivate. However, differences in the cumulative likelihood of recidivism are not significant, as measured by the Wilcoxon test statistic (2.40, $p = .12$).

TABLE 6: Results From the Proportional Hazards Model Comparing the High Implementation MRT Treatment Group and the Control Group

Variable	*Coefficient*	p *Value*
MRT treatment	.35	.11
Disciplinary violations	.12	.31
Prior offenses	.07	.12
Age	–.16	.13
Race (White or non-White)	–.27	.25

NOTE: MRT = moral reconation therapy.

The results of the proportional hazards model for the high implementation and control comparison are included in Table 6. Consistent with the results of the life table survival curves, membership in the high implementation treatment group is not significantly associated with the risk of recidivism. Once again, the model does not offer a significant improvement over a naive model that assumes the explanatory variables contain no reliable information about participants' risk of recidivism.[4]

A final set of analyses explored the possibility that the lack of difference between the treatment and control groups was due to the racial imbalance of these groups. The comparison of the treatment and control groups presented in Table 2 found that the treatment group had significantly more African Americans, whereas the control group had significantly more Caucasians. If these groups differ in their likelihood of recidivism, this difference may obscure any potential treatment effect. Therefore, separate treatment and control group comparisons were estimated for each racial group.

Racially specific survival models indicated that the lack of an apparent treatment effect was not due to the racial imbalance in the treatment and control groups. Survival models comparing African Americans in the treatment group with African Americans in the control group found that risk of recidivism for these two groups did not differ significantly ($p = .14$). Survival models comparing Caucasians in the treatment and control groups also found no statistically significant difference in the risk of recidivism ($p = .32$). Racially specific survival times also indicate that the lack of a treatment effect was not the product of racial imbalance in the treatment and control groups. The

mean survival time for the African American treatment group (318 days) was longer than that of the Caucasian treatment group (254 days). Therefore, the larger percentage of African Americans in the treatment group was associated with longer survival times and a bias toward rather than away from a treatment effect.

To summarize, in this trial, exposure to the MRT program was not associated with significant decreases in the risk of recidivism. Analyses comparing treatment and control groups found no difference in the risk of recidivism. A supplemental analysis compared a high implementation treatment group with the randomized control group. Although directional differences favored the high implementation treatment group, these differences were not significant. It is worthwhile to note that the importance of these directional differences is diminished by the possibility that they may be explained by dropout of high-risk participants from the treatment group. A second supplemental analysis tested the possibility that the lack of a treatment effect was due to racial imbalance in the treatment and control groups. This analysis also found no evidence of a treatment effect.

DISCUSSION

The results of meta-analyses and those of this study demonstrate that the ability of criminal justice system personnel to implement a treatment program needs to be carefully considered before that program can be labeled effective and implemented on a large scale (Andrews et al., 1999; Lipsey, 1999a). The importance of portability is clearly demonstrated by Lipsey (1999a), who found that real-world programs were only half as effective in reducing recidivism as were demonstration projects. The results presented here support Lipsey's conclusions and show that policy makers need to be mindful of the distinction between demonstration projects and real-world programs. This distinction is particularly important in the case of the MRT program. Demonstration projects by the program's authors provide the only empirical support for MRT's widespread implementation. The assessment of the real-world implementation of the MRT program presented in this work suggests that the program lacks portability and that its widespread implementation may have been premature.

Whereas the results of the current test do indicate that MRT lacks portability, these results are conditioned by methodological considerations. The possibility that sample characteristics may have played a part in MRT's apparent lack of portability should be considered. The trials of MRT conducted by the program's authors were based on samples of felony drug offenders and a sample of individuals convicted of driving under the influence (Little & Robinson, 1989; Little et al., 1991). The average age of the sample of drug offenders was 24.5 (SD = 6.5), and the average age of the sample of offenders driving under the influence was 36.6 (SD = 9.5). Independent sample t tests comparing the mean age of these two groups to that of the treatment group used in the current analysis (M = 20.21, SD = .99) show that the group used in this analysis was significantly younger than the groups used by the program's authors.[5] It is possible that this age difference explains the lack of a treatment effect. MRT may contain elements and processes that presuppose cognitive abilities not present in samples of younger offenders.

It is also possible implementation issues explain MRT's apparent lack of portability. Individuals in the treatment group may not have received enough treatment to cause moral development and subsequent behavioral change. This possibility was addressed by the analysis comparing the control group to a high implementation treatment group. The results of this analysis indicate that a lack of sufficient exposure to treatment does not explain the lack of a treatment effect. It is acknowledged, however, that definitive conclusions regarding the amount of exposure to treatment would require measures that directly quantify how much treatment each individual experienced.

A final implementation issue stems from questions regarding the quality of implementation rather than the quantity of implementation. The correctional officers and correctional counselors that implemented the MRT program were trained in the program by Correctional Counseling Inc. To sufficiently inculcate treatment staff members in the cognitive-behavioral principles on which MRT is based, the skills necessary to facilitate the moral development of inmates, and extant research on the MRT program, Correctional Counseling Inc. (2000) offers what is described as an "intensive 32-hour training program." However, given the complexity of a process such as moral development, it seems reasonable to suggest that criminal justice sys-

tem personnel may not be able to effectively implement an intervention targeted at a complex process such as moral development without an investment in training that is much more substantial than the 32 hours offered by Correctional Counseling Inc.

Meta-analyses have found that an emphasis on antisocial cognition and/or interpersonal skills deficits is associated with an increase in program effectiveness (Gendreau et al., 2000; Lipsey, 1999b). Programs that address antisocial cognition or interpersonal skills usually focus on specific cognitive skills (i.e., beliefs about violence or the generation of solutions to a problem situation) rather than a general capacity such as moral development. The effectiveness of programs focusing on specific cognitive skills may be tied to the ability of criminal justice system personnel to successfully implement these interventions. Those responsible for the delivery of program content in real-world settings may have a greater capacity for teaching concrete skills such as negotiation, interpersonal skills, assertiveness, and communication than they have for leading inmates on a journey of moral development.

Future efforts should continue to test the portability of rehabilitation programs. The need for studies that directly assess the characteristics that make a program portable is pronounced. In addition to meeting this need, future efforts may also address the methodological limitations of this work. These limitations include the attenuation of treatment and control group differences and a lack of fidelity to randomization. This analysis addressed the attenuation of treatment and control group differences caused by a lack of fidelity to randomization by conducting two sets of initial analyses. The first compared the 129 participants randomized into the treatment group with the 127 participants randomized into the control group. The second compared the 110 participants that were randomized into the treatment group and received treatment with the 102 participants that were randomized into the control group and were not exposed to treatment. As the results of these analyses did not differ, only those based on the second comparison were presented. Subsequent studies can avoid multiple comparisons by maximizing treatment and control group differences and improving fidelity to randomization. Treatment and control group differences may be maximized by focusing on individuals with at least 30 days remaining on their sentence, and fidelity to randomization

may be improved by asking individuals to volunteer for the experiment prior to assignment to treatment and control groups.

Criminal justice system resources are limited. When devoted to treatment, these resources should be spent on programs whose efficacy has been demonstrated in real-world settings. This work finds that the MRT program lacks portability. Although it is important to note that this is but one trial of the MRT program, it is also important to note that this trial casts doubt on the wisdom of this program's widespread implementation.

NOTES

1. Meta-analyses use a variety of techniques to compute and aggregate effect sizes. An effect size quantifies the difference between the control and treatment groups on an outcome of interest. In general, positive effect sizes are associated with the treatment group performing "better" (i.e., having less recidivism) on an outcome measure.

2. For other between-groups comparisons based on criminal justice system data that use the proportional hazards model, see Hepburn and Albonetti (1994) and Lattimore, Linster, and MacDonald (1997).

3. For the naive model, the –2 log likelihood was 1,302.22 compared with 1,296.23 for the model with explanatory variables.

4. For the naive model, the –2 log likelihood was 952.83 compared with 944.96 for the model with explanatory variables.

5. The *t* value for the comparison of the sample used in this article with the sample of felony drug offenders was –4.74, *df* = 178. The *t* value for the comparison of the sample used in this article with the sample of offenders driving under the influence was –14.99, *df* = 223.

REFERENCES

Andrews, D. A., Dowden, C., & Gendreau, P. (1999). *Clinically relevant and psychologically informed approaches to reduced re-offending: A meta-analytic study of human service, risk, need, responsivity and other concerns in justice contexts*. Unpublished manuscript.

Antonowicz, D. H., & Ross, R. R. (1994). Essential components of successful rehabilitation programs for offenders. *International Journal of Offender Therapy and Comparative Criminology*, *38*, 97-104.

Better People. (2000). *About moral reconation therapy.* Retrieved October 11, 2000, from http://betterpeople.org/mrt.html

Correctional Counseling, Inc. (2000). *About moral reconation therapy.* Retrieved October 11, 2000, from www.http://ccimirt.com

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society*, *34*, 187-220.

Garrett, C. J. (1985). Effects of residential treatment on adjudicated delinquents: A meta-analysis. *Journal of Research in Crime and Delinquency*, *22*, 287-308.

Gendreau, P., Smith, P., & Goggin, C. (2000). Treatment programs in corrections. In J. Winterdyk (Ed.), *Corrections in Canada: Social reaction to crime* (pp. 238-263). Englewood Cliffs, NJ: Prentice Hall.

Hepburn, J. R., & Albonetti, C. A. (1994). Recidivism among drug offenders: A survival analysis of the effects of offender characteristics, type of offense, and two types of intervention. *Journal of Quantitative Criminology*, *10*, 159-179.

Izzo, R. L., & Ross, R. R. (1990). Meta-analysis of rehabilitation programs for juvenile delinquents: A brief report. *Criminal Justice and Behavior*, *17*, 134-142.

Lattimore, P. K., Linster, R. L., & MacDonald, J. M. (1997). Risk of death among serious young offenders. *Journal of Research in Crime and Delinquency*, *34*, 187-209.

Lipsey, M. W. (1999a). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy and Law*, *6*, 611-641.

Lipsey, M. W. (1999b). Can intervention rehabilitate serious delinquents? *Annals of the American Academy of Political and Social Science*, *564*, 142-167.

Little, G. L., & Robinson, K. D. (1988). Moral reconation therapy: A systematic step-by-step treatment system for treatment resistant clients. *Psychological Reports*, *62*, 135-151.

Little, G. L., & Robinson, K. D. (1989). Treating drunk drivers with moral reconation therapy: A one-year recidivism report. *Psychological Reports*, *64*, 960-962.

Little, G. L., Robinson, K. D., & Burnette, K. D. (1990). Treating drunk drivers with moral reconation therapy: A two-year recidivism report. *Psychological Reports*, *66*, 1379-1387.

Little, G. L., Robinson, K. D., & Burnette, K. D. (1991). Treating drug offenders with moral reconation therapy: A three-year recidivism report. *Psychological Reports*, *69*, 1151-1154.

Little, G. L., Robinson, K. D., & Burnette, K. D. (1993). Cognitive behavioral treatment of felony drug offenders: A five-year recidivism report. *Psychological Reports*, *73*, 1089-1090.

Lösel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works in reducing reoffending* (pp. 79-111). New York: John Wiley.

Redondo, S., Sanchez-Meca, L., & Garrido, V. (1999). The influence of treatment programmes on the recidivism of juvenile and adult offenders: A European meta-analytic review. *Psychology, Crime and Law*, *5*, 251-278.

Schmidt, P., & Witte, A. D. (1988). *Predicting recidivism using survival models*. New York/ Berlin: Springer-Verlag.

THE PREDICTIVE VALIDITY OF RISK ASSESSMENT WITH VIOLENT YOUNG OFFENDERS

A 1-Year Examination of Criminal Outcome

ROSALIND E. H. CATCHPOLE
HEATHER M. GRETTON
Youth Forensic Psychiatric Services, Burnaby, British Columbia

This study is an examination of the following three instruments for their ability to predict recidivism in violent young offenders: the Structured Assessment of Violence Risk in Youth, the Youth Level of Service/Case Management Inventory, and the Psychopathy Checklist: Youth Version. The instruments were retrospectively coded from file information on 74 young violent offenders. The authors followed them up for 1 year, examining criminal charges and convictions. They examined the predictive accuracy of each instrument using areas under the curve (AUCs). For general reoffending, AUCs ranged from .74 to .78. For violent reoffending, the AUCs were all .73. Results indicated a moderate to strong relationship between each of the instruments and both general and violent reoffending. The implications of risk assessment for intervention and follow-up are discussed.

Keywords: psychopathy; adolescence; risk assessment; violence; criminality; prediction

Adolescence is a period marked by an increased likelihood of involvement in antisocial behavior (e.g., Moffitt, 1993). Youth violence is a subset of youth antisocial behavior that is of particular concern. In 2000, youths ages 12 to 17 were responsible for 16% of all Canadian violent crimes (Statistics Canada, 2001), and in 1997, they were responsible for approximately one quarter of American violent crimes (U.S. Department of Justice, 1999). In England and Wales,

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 30 No. 6, December 2003 688-708
DOI: 10.1177/0093854803256455

offenders younger than 18 were suspected of committing 18% of assaults and 39% of robberies (Barberet, 2001). Although rates of youth violence have remained quite stable during the past 5 years, by 1998, the number of youths charged with violent crimes in Canada was 77% higher than it had been 10 years earlier (Statistics Canada, 1999). In the United States, rates of violent crime peaked in 1994 and then declined. However, by 1998, rates of violent crime committed by youths had increased 33% from 1989 (U.S. Department of Justice, 2001). Economic analyses have found that serious young offenders used enormous amounts of resources (e.g., incarceration and probation) related to their criminal behavior. According to one U.S. study, the potential economic benefit of preventing a single high-risk youth from becoming a "career criminal" ranges from $1.3 to $1.5 million (Cohen, 1998). These findings point to the need from both an economic and a social perspective to find methods to predict and reduce youth criminality and violence.

PREDICTION OF RECIDIVISM

Predicting those individuals who are likely to continue to be involved in violent crime is an important goal for researchers and clinicians alike. Although there is a growing body of literature on risk assessment with adults (e.g., Douglas & Webster, 1999; Quinsey, Harris, Rice, & Cormier, 1998; Steadman et al., 2000; Webster, Douglas, Eaves, & Hart, 1997), there is a relative gap between our understanding of correlates of youth violence and the application of these

---

AUTHORS' NOTE: *Preparation of this article was supported by the Ministry for Children and Family Development (MCFD) of British Columbia. The views of the authors do not necessarily reflect those of the MCFD. The authors would like to thank the staff of Prince George and Barnaby Youth Forensic Psychiatric Services (YFPS) and Boulder Bay Youth Secure Custody Centre for facilitating this research. We thank the research assistants and volunteers at YFPS for their practical assistance with the project. Portions of this article were presented at the American Psychology and Law Society annual conference in Austin, Texas, in March 2002.* *Correspondence concerning this article should be addressed to Rosalind E. H. Catchpole or to Heather M. Gretton, Youth Forensic Psychiatric Services, 3405 Willingdon Avenue, Burnaby, BC, Canada V5G 3H4; phone: 604-660-5788; fax: 604-660-1109; e-mail: rcatchpo@sfu.ca or heather.gretton@gems7.gov.bc.ca.*

research findings to assessment and intervention practices with adolescents.

A large body of empirical research has identified particular risk factors for violence in youth. These include factors in the family (such as poor family adjustment and abuse), factors in the community (e.g., community disorganisation), peer factors (e.g., peer delinquency), and factors within the youths (such as prior violence, an early age of onset of violence, hyperactivity, poor academic achievement, and personality characteristics; for reviews, see Farrington & Loeber, 2000; Herrenkohl et al., 2000; Tolan & Gorman-Smith, 1998).

One individual factor related to risk for future violence was psychopathy. Psychopathy is a personality construct that has consistently been linked with violence and antisocial behavior in adults (e.g., Grann, Långström, Tengström, & Kullgren, 1999; Harris, Rice, & Quinsey, 1993; Hemphill, Hare, & Wong, 1998; Rice & Harris, 1995; Salekin, Rogers, & Sewell, 1996; Steadman et al., 2000). Because of its strong predictive power, psychopathy, as measured by the Hare Psychopathy Checklist–Revised (Hare, 1991), has been included as a key component in most adult risk assessment instruments.

Psychopathy is a constellation of traits that includes both an affective-interpersonal dimension (e.g., glibness or superficial charm and callousness or lack of empathy) and a behavioral dimension (e.g., need for stimulation or proneness to boredom, poor anger control, and criminal versatility; Hare, 1991). The youth literature suggests that psychopathy as measured by the Psychopathy Checklist–Youth Version (PCL:YV; Forth, Kosson, & Hare, in press) may play a role in understanding youth violence and antisocial behavior (e.g., Brandt, Kennedy, Patrick, & Curtin, 1997; Gretton, McBride, Hare, O'Shaughnessy, & Kumka, 2001; Gretton, Hare, & Catchpole, 2003).

Some recently developed instruments have used multiple known risk factors and correlates of youth antisocial behavior to guide predictions of future risk for criminality and violence. Two such youth risk assessment instruments are the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2002) and the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2002). Given its strong predictive validity for recidivism and violence, psychopathic traits were incorporated as a risk item in the individual-clinical section of version 1 of the SAVRY.

This study was a preliminary examination of the ability of two novel youth risk assessment instruments and the construct of psychopathy to predict general and violent recidivism in a sample of adolescent offenders. As part of a larger treatment outcome study, retrospective file-based risk assessments and prospective examinations of criminal charges and convictions were performed for 74 violent youths across two sites where violent offender treatment programs (VOTPs) were initiated.

METHOD

PARTICIPANTS

A total of 74 participants were obtained from two sites located in British Columbia, Canada. At the first site, Boulder Bay Youth Secure Custody Centre (BBYSCC), participants were 33 male adolescents ages 15 to 19 years who were incarcerated at the BBYSCC between 1998 and 1999. The BBYSCC was a closed custody center located in a wilderness setting near Vancouver. Through a collaborative effort between the BBYSCC and Youth Forensic Psychiatric Services, a treatment program for violent offenders was developed for youths convicted of violent offenses. Of these youths, 17 attended the VOTP at the BBYSCC, whereas the comparison group of 16 violent offenders were randomly selected from a list of violent offenders who had been incarcerated at the BBYSCC for similar periods as treated youths.

The second site was Prince George Youth Forensic Psychiatric Services (PGYFPS), an outpatient-based forensic psychiatric assessment and treatment facility. Participants from the PGYFPS site were 41 youths (30 male and 11 female) convicted of a violent offense, 20 of whom participated in the VOTP at PGYFPS and 21 of whom served as a comparison group.

Across both sites, adolescents were selected by clinicians to attend the VOTP. Youths who were not able to attend the VOTP for a variety of reasons (incorrect sentence length, lack of appropriate placement, and so forth) formed the comparison group.

Demographic information. A total of 55.4% ($n = 41$) of the youths were White, 29.7% ($n = 22$) were Aboriginal, 8.1% ($n = 6$) were Asian, and 5.4% ($n = 4$) were of other ethnic backgrounds (one youth's ethnicity was unknown). Chi-square analyses revealed no differences in ethnicity between treated and comparison youths ($p = .60$) but did find differences in ethnicity across sites, $\chi^2(3, N = 73) = 15.07$, $p = .002$, with visual inspection of the data suggesting differences in the number of Aboriginal youths across sites.

Overall, for youth on which information was available, 53% had engaged in daily drug or alcohol use at some point in their lives ($n = 32$). Chi-square analyses revealed no differences in frequency of drug or alcohol use across sites or treatment versus comparison groups (ps $> .5$). Youths spent an average of 8.3 years ($SD = 6.2$) with both biological parents. Independent-samples t tests revealed no differences across sites or across treatment versus comparison groups (ps $> .25$). A total of 68.9% of youths for which information was available ($n = 31$) had witnessed violence between caregivers (chi-square analyses revealed no differences across sites or treatment versus comparison groups, ps $> .25$).

The mean number of conduct disorder symptoms, as defined by the *DSM-IV* (American Psychiatric Association, 1994), was 6.5 ($SD = 2.8$) out of a maximum of 15. Independent sample t tests revealed no differences in number of conduct disorder symptoms across sites or treatment versus comparison groups (ps $> .5$). Only 4.1% of youth ($n = 3$) did not have a pre-index history of violent offending. Youths were on average 16 years of age at index offense (range = 12.4 to 18.3; $SD = 1.3$). Independent sample t tests revealed no differences in age at index offense across treatment versus comparison groups ($p > .5$) and revealed only a very slight trend toward Prince George youth being younger at age of index offense ($p = .16$). Types of index offenses committed include robbery or armed robbery (25.7%; $n = 19$), assault (14.9%; $n = 11$), aggravated assault or assault causing bodily harm (36.5%; $n = 27$), attempted murder or manslaughter (2.7%; $n = 2$), other violent offenses (13.5%; $n = 10$), and other nonviolent offenses (32.4%; $n = 24$). Percentages do not add to 100% due to the possibility of multiple index offenses. Chi-square analyses were conducted to look at the entire set of offense types and revealed no significant dif-

ferences across treatment versus comparison groups (*p*s > .15). In addition, the data were visually inspected, revealing very similar percentages of youths engaging in specific types of offenses across sites and treatment conditions. Overall, the individual, familial, and offending characteristics of the youths were very similar across sites and across treatment versus comparison groups at preassessment.

DATA SOURCES

Background. This study was initiated after the adolescents had been discharged from their respective facilities. Therefore, we were unable to directly interview the participants to obtain information for the risk assessment instruments. Instead, we relied on the extensive information collected in the forensic files.

We recorded background information from clinical and forensic files, including (when available) predisposition reports, social histories, police and victim statements, psychological and/or psychiatric reports, and summaries documenting the youths' behavior and progress during any of their assessments, treatments (if applicable), and in the case of the BBYSCC, incarcerations. We also coded the risk instruments (SAVRY and YLS/CMI) and psychopathy (PCL:YV) from file information following the youths' discharge from the BBYSCC or PGYFPS. All raters had thorough training in the specific instruments and had achieved a reliability of at least .80 on 10 training files.

Follow-up information. We obtained criminal record information using British Columbia Corrections files allowing for a 12-month follow-up for each youth in the study. We used information from this source to tabulate the number of criminal offenses (charges and convictions), types of offenses, months spent in custody, months on probation, and months free in the community for each participant. For the BBYSCC, the follow-up period was calculated as 12 months from the date of discharge from the BBYSCC. For PGYFPS, the follow-up period was defined in two ways. For youths who attended treatment, the follow-up period was defined as 12 months from the end of treatment. For youths who did not attend treatment (the comparison group for the treatment outcome study), the follow-up period was defined as

12 months from the date of discharge from assessment plus 8 months. This lapse from date of discharge to assessment was to account for effects such as maturation, thus allowing a more direct comparison with treated youth who spent an average of 8 months in treatment following their assessment.

OUTCOME VARIABLES

Outcome variables included offenses (as indicated by charges and convictions on British Columbia Corrections) that occurred during the follow-up period. We coded violent offenses (including assaults, robbery, intimidation, and attempted murder), and we also created a general recidivism category that included any reoffending (including both violent and nonviolent) during the follow-up period.

INSTRUMENTS

SAVRY. The SAVRY (Borum et al., 2002) is a tool designed to assist clinicians in evaluating risk for violence in an adolescent population. It was developed to fill a need in the adolescent forensic field for a clinical risk assessment instrument. Its structure is based on the HCR-20 (Webster et al., 1997), a 20-item adult instrument that examines historical, clinical, and risk management variables associated with violent behavior. The SAVRY includes 10 historical risk factors, 6 social-contextual risk factors, 8 individual-clinical risk factors (including psychopathy), and 6 protective factors, all derived from literature on correlates of violence. The historical, social, and clinical factors are coded on a 3-point scale. The protective factors are scored on a dichotomous (present or absent) scale. For risk assessment, the SAVRY can be used in two ways. Items can be summed to give a total score and/or a clinical judgment can be applied, taking into account the total score, the protective factors, and any other relevant information to give an overall rating of low, moderate, or high risk for violence. In this study, we used the clinical judgment ratings of risk for long-term violence for categorical analyses and the total score for analyses requiring a continuous variable. The SAVRY was coded based on file information for 66 of the 74 youths (in eight cases, there was insufficient file information to code the SAVRY).

Very few data were available on the psychometric properties of the SAVRY. In a preliminary study, Bartel and Forth (2000) found that when item sums are used, the mean score for the SAVRY was 21.2 in a sample of incarcerated young male offenders. They also found positive associations between the SAVRY total score and the number of violent acts, nonviolent acts, and violent versatility (r = .20 to .33), and negative relationships between the SAVRY and age at onset of violent and nonviolent acts (r = –.40). Furthermore, a correlation of .83 with the YLS/CMI and a .73 correlation with the PCL:YV in a sample of incarcerated male youth offenders give preliminary evidence for concurrent validity.

In this study, reliabilities were calculated for both the SAVRY total score and the SAVRY clinical judgment based on a subsample of 21 participants (roughly one quarter of the total sample size). For the total score, the single-rater intraclass correlation (ICC) was 0.81. For the clinical judgment (a 3-point categorical rating), the single-rater ICC was 0.77.

YLS/CMI. The YLS/CMI (Hoge & Andrews, 2002) is a 42-item inventory that assesses the following eight categories of criminogenic factors: prior and current offenses and dispositions, family circumstances and parenting, education and employment, peer relations, substance use, leisure and recreation, personality and behavior, and attitudes and orientation. The YLS/CMI was developed from the Level of Supervision Inventory (Andrews, 1982) based on the assumption that decisions about young offenders must be based on relevant assessments of risk and need characteristics. It was founded on the General Personality and Social Psychological Model of Criminal Conduct (Andrews & Bonta, 1990) that attributes the causes of youth criminality to both characteristics and circumstances of the youths. It further suggests that interventions with high-risk youths can be effective in reducing recidivism if they target criminogenic needs of youths.

Each item on the YLS/CMI is coded as either present or absent, and the present items are then summed to give a total score. Total scores can range from 0 to 42. Cutoff scores, based on total scores, give four categories of risk for continued criminal activity (labeled *summary of risk and needs factors*): low, moderate, high, and very high. Other sec-

tions of the YLS/CMI (not coded here) assess other considerations such as appropriate contact level, case management considerations, and so forth and apply clinical judgment to the assessment of risk. The YLS/CMI was coded from available file information for 67 of the 74 youths (in seven cases, there was insufficient information to make a final rating).

Jack (2000) found that the YLS/CMI correlated with PCL:YV scores ($r = .70$), with past violent convictions ($r = .30$), and with past other convictions ($r = .20$). Overall, she found a significant relationship between the YLS/CMI and general recidivism, although not specifically with violent recidivism. Jung and Rawana (1999) found that the YLS/CMI (or the MRNAF as it was alternately called) was predictive of general recidivism during a 6-month period in both male and female native and non-native youth offenders.

In this study, reliabilities were calculated for the YLS/CMI total score only because cutoff scores used for categorical analyses are based on total scores, and not on a clinical judgment. For the total score, the single-rater ICC ($N = 21$) was 0.80.

PCL:YV. The PCL:YV (Forth et al., in press) is a 20-item clinical rating tool that assesses youths on several behavioral and personality characteristics associated with psychopathy. The PCL:YV closely parallels the adult version of the PCL, the Hare Psychopathy Checklist–Revised (Hare, 1991), with some items having been slightly modified for use with adolescents. Each item is scored on a 3-point scale. The total score can range from 0 to 40, with higher scores reflecting a greater number of psychopathic traits. Mean scores in forensic youth populations range from 23 to 25 (Brandt et al., 1997). For categorical analyses, PCL:YV scores were divided into the following three groups: low (0 to 19.9), moderate (20 to 29.9), and high (30 to 40). The PCL:YV was coded for 67 of the 74 youths from file information (in seven cases, there was insufficient information to make a final rating).

The Hare Psychopathy Checklist–Revised has been found to have strong psychometric properties (Salekin et al., 1996), and studies of the PCL:YV show equally promising results. The PCL:YV has high internal consistency (Cronbach's αs of .85 to .90) and high interrater reliability (ICCs of .82 to .94; Brandt et al., 1997; Gretton et al., 2001). Validity data show that the PCL:YV has a good predictor of recidi-

vism in a sample of adolescent sex offenders (Gretton et al., 2001) and in general offenders (Brandt et al., 1997).

In this study, the reliability was calculated for the PCL:YV total score based on a subsample of 21 participants. The single-rater ICC for the PCL:YV total score was 0.90.

RESULTS

POSTRELEASE OFFENSES

Overall for general reoffending, 58% of youths ($n = 43$) committed at least one post-release offense during the one-year follow-up period. The mean number of offenses was 2.7. The base rate for violent reoffending was 23% ($n = 17$). The average number of violent offenses was .50. Types of violent offenses committed during the follow-up period included assault (4.1%; $n = 3$), aggravated assault or assault causing bodily harm (9.5%; $n = 7$), robbery (9.5%; $n = 7$), intimidation (14.9%; $n = 11$), unlawful confinement or forcible seizure (2.7%; $n = 2$), harassment or stalking (1.4%; $n = 1$), and use or possession of a weapon (1.4%; $n = 1$). Percentages do not add up to 23% due to the possibility of an individual being charged with multiple types of follow-up offenses.

ASSESSMENT INSTRUMENTS

Risk classification. The percentage of individuals classified as low, moderate, high, and very high risk for the SAVRY and the YLS/CMI is presented as follows. On the SAVRY, 25.8% of youths ($n = 17$) were identified as low risk, 43.9% as moderate risk ($n = 29$), and 30.3% as high risk ($n = 20$). On the YLS/CMI, 3.0% of youths ($n = 2$) were identified as low risk, 28.4% ($n = 19$) as moderate risk, 64.2% ($n = 43$) as high risk, and 4.5% ($n = 3$) as very high risk.

Across the two risk assessment instruments there was considerable variability in the number of youth categorized at each risk level. Overall, between 3% (YLS/CMI) and 25% (SAVRY) were classified as low risk. A substantial range from 30% (SAVRY) to 69% (YLS/CMI) was identified as high or very high risk by the instruments.

Psychopathy classification. According to the PCL:YV, 31.3% of youths ($n = 21$) were low on psychopathy (0 to 19.9), 41.8% ($n = 28$) were moderate on psychopathy (20 to 29.9), and 26.9% ($n = 18$) were high on psychopathy (30 to 40). The mean PCL:YV score for all youths was 23.8 (range = 8.0 to 35.6, $SD = 6.9$).

Correlations. Bivariate correlations revealed strong relationships between the YLS/CMI summary of risk and needs, the SAVRY clinical risk judgment, and the PCL:YV total score. The correlation between the SAVRY clinical risk judgment and the YLS/CMI summary of risk and needs was .64 ($p < .01$), and between the SAVRY risk judgment and the PCL:YV, the total score was .68 ($p < .01$).[1] The correlation between the YLS/CMI summary of risk and needs and the PCL:YV total score was .75 ($p < .01$).

DATA ANALYSIS

Receiver operating characteristic (ROC) analyses. ROC analyses evaluate an instrument's ability to predict an event (in this case, reoffending). ROC analysis plots sensitivity versus specificity (technically, one minus specificity) and computes an area under the curve (AUC) for a continuous predictor (in this case, SAVRY, YLS/CMI, and PCL:YV). An AUC is defined as the probability that a randomly selected individual who recidivated will score higher on the specified risk assessment measure than will a randomly selected individual who did not recidivate (Mossman & Somoza, 1991). An AUC of 1 indicates a perfect measurement, whereas an AUC of .5 indicates no improvement over chance. ROC is less affected by base rates than are other measures and is an index of the "trade-off" between sensitivity and specificity of an instrument. (See Douglas, Ogloff, Nicholls, & Grant, 1999, for a comprehensive explanation of ROC analyses.) We compute AUCs for the PCL:YV, the YLS/CMI, and the SAVRY for both general and violent reoffending.

Survival analyses. Survival curve analyses estimate the time (in months) it takes for youths to reoffend and the rate of occurrence of that event (the survival function). That is, they determine the propor-

tion of youths who have not reoffended at each month of the follow-up period ("survival"). Survival analyses are used to obtain survival curves for categories of risk and psychopathy (low, moderate, and high). Survival curves give a picture of offending patterns over time to see not only whether certain risk groups recidivate in greater proportions than other risk groups but whether they reoffend more quickly. Log rank tests are used for group and pairwise comparisons.

PREDICTION OF RECIDIVISM

ROC analyses. ROC AUCs are presented as follows. Because AUCs are designed to measure the relationship between a continuous predictor variable and a dichotomous outcome variable, we used total scores for all the instruments for this particular analysis. Because psychopathy is an item on the SAVRY, we computed AUCs for the SAVRY total score both including and excluding the psychopathy item.

In this study, for general recidivism, AUCs were .74 for the SAVRY total score (.73 for the SAVRY total score excluding the psychopathy item), .74 for the YLS/CMI total score, and .78 for the PCL:YV total score (ps < .01), indicating a fairly strong predictive relationship between risk classification and general recidivism. For violent recidivism, AUCs were .73 for the SAVRY total score (.71 for the SAVRY total score excluding the psychopathy item), .73 for the YLS/CMI total score, and .73 for the PCL:YV total score (ps < .01), also indicating moderately strong ability of the instruments in predicting the occurrence of violence.

Survival. Log rank tests revealed a significant difference between PCL:YV groups and general reoffending. The high PCL:YV group was significantly different from both the low PCL:YV group, $\chi^2(1, N = 39) = 18.64$, $p < .001$, and the moderate PCL:YV group, $\chi^2 (1, N = 46) = 9.61$, $p = .002$. The low and moderate groups did not differ significantly.

A survival curve depicting violent offenses by each group on the PCL:YV is presented in Figure 1. For violent reoffending, only the PCL:YV high group was significantly different from the PCL:YV low

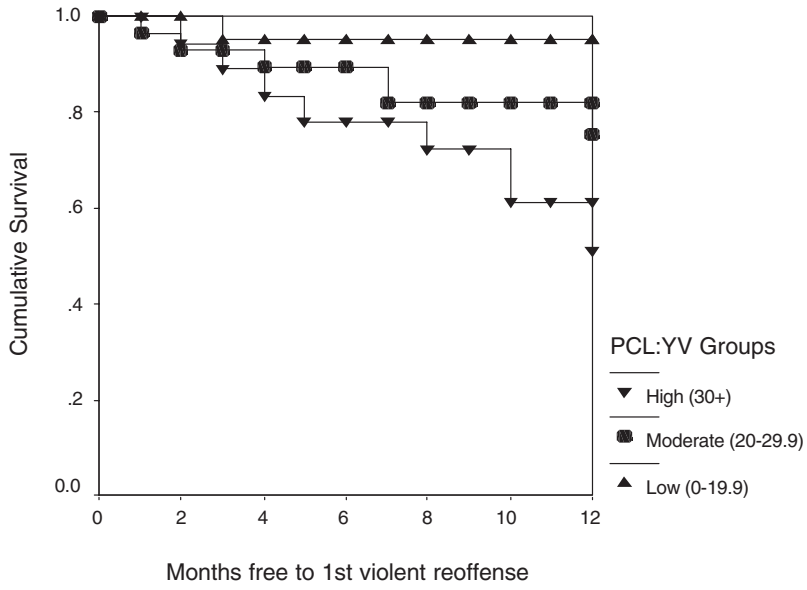

Figure 1: Survival to First Violent Reoffense by the Psychopathy Checklist–Youth Version Group

group, χ^2 (1, N = 39) = 8.25, p = .004. The other groups were not significantly different from each other.

Log rank tests revealed a significant difference between YLS/CMI risk level and general reoffending. Because of the small number of individuals in both the low- and very high–risk groups (2 and 3, respectively), the low- and moderate-risk group was collapsed and the high- and very high–risk group was collapsed for these analyses. The low- and moderate-risk group was significantly different from the high- and very high–risk groups, χ^2 (1, N = 67) = 9.73, p = .002.

A survival curve depicting violent offenses by each level of risk on the YLS/CMI is presented in Figure 2. For violent reoffending, the low- and moderate-risk group was also significantly different from the high- and very high–risk group, χ^2 (1, N = 67) = 7.53, p = .006.

Log rank tests revealed a significant difference between SAVRY risk level and general reoffending. The high risk group was significantly different from the low risk group, χ^2 (1, N = 37) = 5.95, p = .01.

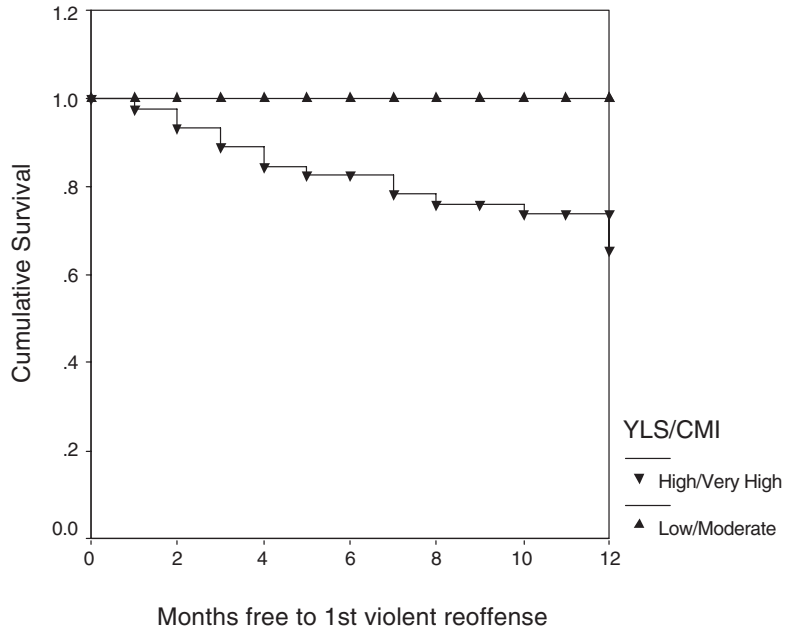

Figure 2: Survival to First Violent Reoffense by the Youth Level of Service/Case Management Inventory Risk Level

The high-risk group was also significantly different from the moderate-risk group, χ^2 (1, N = 49) = 5.33, p = .02. The low- and moderate-risk group did not significantly differ with respect to general reoffending.

A survival curve depicting violent offenses by each level of risk on the SAVRY is presented in Figure 3. For violent reoffending, the high-risk group and the low-risk group differed significantly, χ^2 (1, N = 37) = 6.00, p = .01. The high-risk group also differed significantly from the moderate-risk group, χ^2 (1, N = 49) = 3.92, p = .05. The low and moderate groups did not differ significantly from each other.

DISCUSSION

This study provides evidence for the predictive ability of two risk assessment instruments (YLS/CMI and SAVRY) and a personality

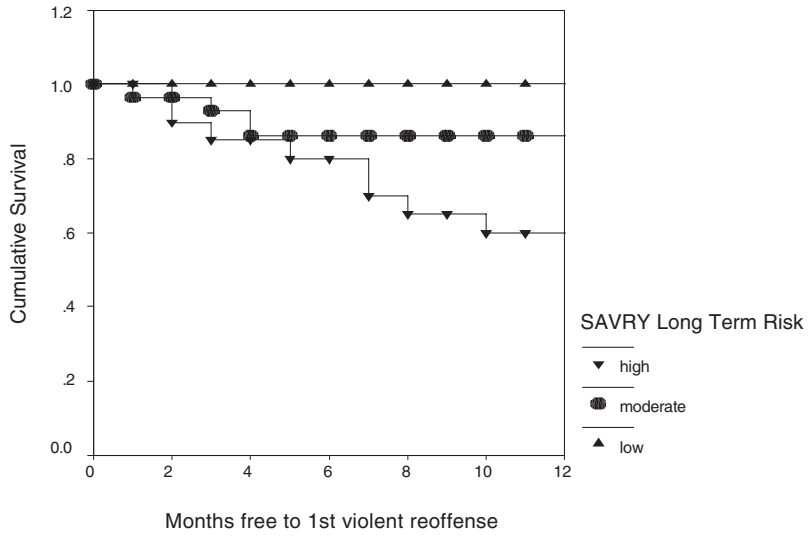

Figure 3: Survival to First Violent Reoffense by the Structured Assessment of Violence Risk in Youth Risk Level

construct (psychopathy, as measured by the PCL:YV) to predict general and violent recidivism.

GENERAL REOFFENDING

All of the instruments were significantly related to general offending, as evidenced by the ROC analyses. There was a strong relationship between each of the instruments and general offending, with ROC AUCs ranging from .74 to .78. Recall that an AUC is defined as the likelihood that a youth who recidivated will score higher on a specified risk assessment measure than will a youth who did not recidivate (Mossman & Somoza, 1991). Thus, in the case of the SAVRY, YLS/CMI, and PCL:YV, AUCs of .74 to .78 indicate that there is a 74% to 78% chance that an individual who recidivated will score higher on the measure than will a nonrecidivist. Survival analyses also confirmed that the instruments were able to differentiate those who were more likely to reoffend and more likely to do so more quickly.

VIOLENT REOFFENDING

The two risk measures (the SAVRY and the YLS/CMI) as well as the personality construct of psychopathy (as measured by the PCL:YV) were similarly and significantly related to violent recidivism, all with ROC AUCs of .73. Survival analyses revealed that for all the instruments, low-risk (or low psychopathy) and high-risk (or high psychopathy) youths could be meaningfully differentiated in terms of their violent reoffending patterns during the 1-year follow-up period.

PREDICTION OF RECIDIVISM

Overall, this study provides evidence for the ability of three instruments to differentiate risk for recidivism in a sample of violent youth offenders. It is not surprising that the construct of psychopathy was strongly related to both general and violent recidivism, which is similar to findings in adult samples (e.g., Hemphill et al., 1998). The two risk assessment instruments (SAVRY and YLS/CMI) were also significantly related to risk for general and violent recidivism. Notably, there were differences between the two instruments in categorizing risk. The YLS/CMI tended to identify very few youths as low risk, and more youths were identified as higher risk on the YLS/CMI than on the SAVRY. Given the different primary purposes of the two instruments, these differences in categorization may be expected. The YLS/CMI is designed to predict general criminality (Hoge & Andrews, 2002), and the SAVRY is designed to predict future violent offending (Borum et al., 2002). Despite this difference in threshold for defining low, medium, and high risk, both instruments were similarly predictive of both general and violent recidivism.

ISSUES RELATED TO YOUTH RISK ASSESSMENT

Several factors potentially affect the accuracy of youth risk assessments and warrant further investigation. Developmentally, adolescence is a time of change and maturation. The influence of dynamic factors related to development (such as responsibility and social perspective taking) is not well understood. Youth risk assessment should take into account the developing nature of adolescent functioning and

corresponding changes that relate to risk. In practice, this suggests the need for up-to-date assessments that take into account current functioning and criminogenic needs so dynamic factors related to risk can be addressed and accuracy of risk estimates optimized. In this study, risk assessment instruments were coded retrospectively based on extensive file information collected during the youths' admissions. Previous research has found that reliable and valid assessments can be made from extensive file information (e.g., Grann, Långström, Tengström, & Stålenheim, 1998; Gretton et al., 2001). In our study, we found that reliable assessments could be made based on file review and that the current instruments showed strong relationships to violent outcome. As such, this study represents an important initial step. However, there is a need for more prospective research on risk assessment conducted at the time of assessment or intervention so that dynamic factors and their contribution to violent risk and outcome can be examined more thoroughly.

There is a lack of available outcome data in the youth risk assessment literature, in particular, violent outcome data. This lack of data limits our understanding of risk assessments, including reliability, predictive ability, and generalizability across settings and youth populations. Although this study provides some early data toward developing literature on adolescent risk assessment, more outcome data need to be generated.

Given the stability of aggression over time and issues of long-term risk among highly aggressive youths (Huesmann & Moise, 1999), the time frame over which an instrument predicts recidivism is an important consideration when evaluating its efficacy (Steadman, 2000). This study does not address longer term risk assessment beyond 1 year, an important issue in determining how youths will fare as they enter adulthood and an important area for future research. The ability to make specific predictions concerning seriousness and types of violent outcome is another important goal for future research. The ability to predict the likelihood of serious violent offending is an ongoing issue in the adult risk assessment literature (e.g., Steadman, 2000) and is one that has not yet been addressed with adolescents.

The use of risk assessments for the purpose of prediction and issues surrounding the accuracy of prediction remain an important consideration. Incorrect classification of youths can have negative implica-

tions for both the youths and the broader community. In this study, none of the instruments were perfect in their predictions; the instruments made errors of both over- and underprediction. Clearly, underpredicting violence (failing to determine that a person will be violent) is a serious error, as it results in undetected violent individuals who will cause harm to others. The consequences of overpredicting violence depend on the purpose of the assessment and its use. Overprediction that is used to justify retributive justice (such as increased incarceration time or stricter probation conditions) outside of the context of public safety and youth rehabilitation is an error that interferes with the rights and freedoms of the youths. Given the limitations in understanding and in accuracy of youth risk assessment, we caution against the use of youth risk assessment procedures for the purpose of prediction without consideration of intervention.

RISK ASSESSMENT AS A GUIDE FOR INTERVENTION

The purpose of the Young Offender's Act in Canada (replaced in 2003 by the Youth Criminal Justice Act) is to maintain the safety of the public and to provide rehabilitative intervention to youths by addressing the relevant needs and circumstances of youths (Justice Department of Canada, 2000). Under this framework, risk assessment and intervention are complementary processes that identify risk factors for the youths (risk assessment) and address these issues (intervention).

This study found that the assessment instruments meaningfully differentiated risk for violent recidivism among previously violent youth offenders. Only 1 youth out of the 17 (5.9%) defined as low risk on the SAVRY violently reoffended in the 1-year follow-up, whereas 8 out of 20 youths (40%) identified as high risk on the SAVRY violently reoffended in the 1-year follow-up. Using the YLS/CMI, none of the 21 youths identified as low or moderate risk violently reoffended in the 1-year follow-up. In comparison, 30% (14 of 46) of the youths identified as high or very high risk on the YLS/CMI violently reoffended in the follow-up. Examining violent criminal outcome using the PCL:YV, only 1 out of the 21 youths identified as low on psychopathy (4.8%) violently recidivated, whereas 8 out of 18 youths identified as high on psychopathy (44.4%) violently recidivated dur-

ing the follow-up period. Using criminal records as an indicator of violent outcome, each of the instruments differentiated risk among a group of previously violent offenders.

These findings indicate that the current instruments meaningfully differentiated risk for ongoing violence even among youth previously convicted for violent offenses. When we review these findings from a risk/needs perspective, they suggest that even in the case of violent young offenders, some individuals clearly are in need of a more intensive intervention than are others (for discussions of this issue, see Flannery & Huff, 1999; Ward & Dockerill, 1999). It may be argued that youths identified by risk assessment instruments as lower risk may be appropriate for a community-based "management" intervention that addresses areas of specific concern related to risk for those youths. The risk-need-responsivity principle (Andrews & Bonta, 1990) implies that the most intensive treatment should be reserved for high-risk offenders with strong criminogenic needs and that relapse prevention procedures should be used to enhance self-management skills and to maximize generalization to noninstitutional environments. This study provides evidence for the use of adolescent risk assessment measures for designing and administrating intervention programs specific to the criminogenic needs of the youths, with the primary goal being risk reduction.

NOTE

1. Psychopathy is included as an item on the Structured Assessment of Violence Risk in Youth (SAVRY). Because the clinical judgment of risk on the SAVRY includes evaluating the entire instrument, it was not possible for SAVRY clinical judgment raters to be blind to the Psychopathy Checklist–Youth Version (PCL:YV) score, although all other items on the SAVRY were coded before the PCL:YV score was known.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andrews, D. A. (1982). *The Level of Service Inventory (LSI)*. Toronto, ON, Canada: Ministry of Correctional Services.

Andrews, D. A., & Bonta, J. (1990). Classification for effective rehabilitation. *Criminal Justice and Behavior, 17*, 19-52.

Barberet, R. (2001). Youth crime in Western Europe: Will the old world imitate the new? In S. O. White (Ed.), *Handbook of youth and justice* (pp. 207-222). New York: Kluwer Academic.

Bartel, P., & Forth, A. (2000, March). *The development and use of the Structured Assessment of Violence Risk in Youth*. Paper presented at the American Psychology and Law Society Biennial Conference, New Orleans, Louisiana.

Borum, R., Bartel, P., & Forth, A. (2002). *Manual for the Structured Assessment for Violence Risk in Youth (SAVRY). Consultation version*. Tampa: Florida Mental Health Institute, University of South Florida.

Brandt, J. R., Kennedy, W. A., Patrick, C. J., & Curtin, J. J. (1997). Assessment of psychopathy in a population of incarcerated adolescent offenders. *Psychological Assessment*, *9*, 429-435.

Cohen, M. A. (1998). The monetary value of saving a high risk youth. *Journal of Quantitative Criminology*, *14*, 5-33.

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 Violence Risk Assessment Scheme and the Psychopathy Checklist: Screening version. *Journal of Consulting and Clinical Psychology*, *67*, 917-930.

Douglas, K. S., & Webster, C. (1999). The HCR-20 violence risk assessment scheme: Concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior, 26*, 3-19.

Farrington, D. P., & Loeber, R. (2000). Epidemiology of juvenile violence. *Child and Adolescent Psychiatric Clinics of North America*, *9*, 733-748.

Flannery, D. M., & Huff, C. R. (1999). Implications for prevention, intervention, and social policy with violent youth. In D. M. Flannery & C. R. Huff (Eds.), *Youth violence: Prevention, intervention and social policy* (pp. 293-306). Washington, DC: American Psychiatric Press.

Forth, A. E., Kosson, D. S., & Hare, R. D. (in press). *The Psychopathy Checklist: Youth version (PCL:YV)*. Toronto, Ontario: Multi-Health Systems.

Grann, M., Långström, L., Tengström, A., & Kullgren, G. (1999). Psychopathy (PCL-R) predicts violent recidivism among criminal offenders with personality disorders in Sweden. *Law and Human Behavior, 23*, 205-217.

Grann, M., Långström, N., Tengström, A., & Stålenheim, E. G. (1998). Reliability of file-based retrospective ratings of psychopathy with the PCL-R. *Journal of Personality Assessment*, *70*, 416-426.

Gretton, H. M., Hare, R. D., & Catchpole, R. E. H. (2003). *Psychopathy and offending from adolescence to adulthood: A ten year follow up.* Manuscript under review.

Gretton, H. M., McBride, M., Hare, R. D., O'Shaughnessy, R., & Kumka, G. (2001). Psychopathy and recidivism in adolescent sex offenders. *Criminal Justice and Behavior, 28*, 427-449.

Hare, R. D. (1991). *The Hare Psychopathy Checklist–Revised (PCL-R)*. Toronto, Ontario: Multi-Health Systems.

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.

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3*, 139-170.

Herrenkohl, T. I., Maguin, E., Hill, K. G., Hawkins, J. D., Abbott, R. D., & Catalano, R. F. (2000). Developmental risk factors for youth violence. *Journal of Adolescent Health*, *26*, 176-186.

Hoge, R., & Andrews, D. (2002). *The Youth Level of Service/Case Management Inventory.* Toronto, ON, Canada: Multi-Health Systems.

Huesmann, L. R., & Moise, J. F. (1999). Stability and continuity of aggression from early childhood to young adulthood. In D. J. Flannery & C. R. Huff (Eds.), *Youth violence: Prevention, intervention and social policy* (pp. 73-95). Washington, DC: American Psychiatric Press.

Jack, L. A. (2000). *Psychopathy, risk/need factors, and psychiatric symptoms in high risk youth: Relationships between variables and their link to recidivism*. Unpublished doctoral dissertation, Simon Fraser University, British Columbia, Canada.

Jung, S., & Rawana, E. P. (1999). Risk and need assessment of juvenile offenders. *Criminal Justice and Behavior*, *26*, 69-89.

Justice Department of Canada. (2000, August). *Young Offender's Act*. Retrieved March 23, 2001, from http://canada.justice.gc.ca/en/laws/Y-1/text.html

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674-701.

Mossman, D., & Somoza, E. (1991). ROC curves, test accuracy, and the description of diagnostic tests. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *3*, 330-333.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.

Rice, M. E., & Harris, G. T. (1995). Psychopathy, schizophrenia, alcohol abuse, and violent recidivism. *International Journal of Law and Psychiatry*, *18*, 333-342.

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the Psychopathy Checklist and Psychopathy Checklist—Revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice*, *3*, 203-215.

Statistics Canada. (1999). *Change in rates of youth violent crime from 1977-1998*. Retrieved August 1, 2001, from http://www.statcan.ca/english/indepth/85-002/feature/jurl1999012021sda.htm

Statistics Canada. (2001). *Rates of youth and adult criminal behaviour by type*. Retrieved August 1, 2001, from http://www.statcan.ca/english/Pgdb/State/Justice/legal14.htm

Steadman, H. J. (2000). From dangerousness to risk assessment of community violence: Taking stock at the turn of the century. *Journal of the American Academy of Psychiatry and Law*, *28*, 265-271.

Steadman, H. J., Silver, E., Monahan, J., Appelbaum, P. S., Clark-Robbins, P., & Mulvey, W. P., et al. (2000). A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior*, *24*, 83-100.

Tolan, P. H., & Gorman-Smith, D. (1998). Development of serious and violent offending careers. In R. Loeber & D. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 68-85). Thousand Oaks, CA: Sage.

U.S. Department of Justice. (1999). *Office of Juvenile Justice and Delinquency Prevention statistical briefing book*. Retrieved March 18, 2002, from http://www.ojjdp.ncjrs.org/ojstatbb/html/qa136.html

U.S. Department of Justice. (2001, August). *Person offences in juvenile court, 1989-1998* (Office of Juvenile Justice and Delinquency Prevention Fact Sheet No. 32). Retrieved March 18, 2002, from http://www.ncjrs.org/pdffiles1/ojjdp/fs200132.pdf

Ward, A., & Dockerill, J. (1999). The predictive accuracy of the violent offender treatment program risk assessment scale. *Criminal Justice and Behavior*, *26*, 125-141.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. (1997). Assessing risk of violence to others. In C. D. Webster, D. Christopher, & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 251-277). New York: Guilford.

MORE EVIDENCE FOR THE VALIDITY OF THE SELF-APPRAISAL QUESTIONNAIRE FOR PREDICTING VIOLENT AND NONVIOLENT RECIDIVISM

A 5-Year Follow-Up Study

WAGDY LOZA
Kingston Penitentiary
Carleton University
AMEL LOZA-FANOUS
Carleton University

The goal of this research was to examine the effectiveness of the Self-Appraisal Questionnaire (SAQ) in predicting release outcome during a 5-year period. The SAQ is a quantitative risk-need instrument consisting of 72 items that constitute eight subscales. The SAQ was administered to 305 federally sentenced Canadian male offenders prior to their release into the community. They were followed up for 60 months at 4-month intervals. Outcome criteria measures were violent recidivism, commission of a new offense, general recidivism (returning to any form of custody), and any failure (a composite measure recording failure on any of the following variables: negative parole reports, violation of parole conditions, incurring new charges, or a new conviction). Results demonstrated that the SAQ has adequate predictive validity during the 5-year period.

Keywords: predictive validity over 5 years; offender recidivism; Self-Appraisal Questionnaire.

The practice of using objective actuarial measures designed specifically for predicting which offenders are likely to recidivate is a fairly new phenomenon to correctional systems. Prior to the development of these measures, professionals relied on clinical judgment and

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 30 No. 6, December 2003 709-721
DOI: 10.1177/0093854803256448

measures that were not specifically designed for predicting recidivism when estimating risk for reoffending. This practice still exists as evidenced by a recent survey (Boothby & Clements, 2000) that indicated measures that are specifically designed for predicting recidivism (e.g., Level of Service Inventory Revised, Andrews & Bonta, 1995; Violence Risk Appraisal Guide, Harris, Rice, & Quinsey, 1993) are each used by less than 1% of the forensic or correctional psychologists in the United States and that 87% of these psychologists are still using the Minnesota Multiphasic Personality Inventory (MMPI). The high number of psychologists still using the MMPI, as opposed to using actuarial measures specifically designed for offender populations, is quite puzzling when one considers that none of the original clinical scales of the MMPI were specifically designed for use with offenders or the prediction of violent behavior (Quinsey, 1979). In fact, using some of the original MMPI scales (i.e., the Psychopathic Deviate and the Manic subscales) or other MMPI scales developed from the MMPI (i.e., the Over Controlled Hostility and the Megargee's MMPI typologies) to distinguish between violent and nonviolent individuals has not been found to be reliable or valid for such use (Gynther, Altman, & Warbin, 1973; Lothstein & Jones, 1978; McCreary, 1976) or useful in the prediction of recidivism (Holland & Holt, 1975; Mendelzys, 1979).

Currently, correctional professionals have several measurements to assist them when predicting violent and nonviolent recidivism. These include the Level of Service Inventory–Revised (Andrews & Bonta, 1995), the General Statistical Information on Recidivism (Nuffield, 1982), the Violence Risk Appraisal Guide (Harris, Rice, & Quinsey, 1993), and the Self-Appraisal Questionnaire (SAQ; Loza, 1996). Although the Psychopathy Checklist–Revised (Hare, 1991) was not initially developed for the prediction of recidivism, there is evidence

AUTHORS' NOTE: *We wish to express our appreciation to Lynn Ann Spincer and Toni Hemmati for their help with the data collection. Also, many thanks to Anita Cumbilton and Franca Cortoni for editing the drafts. We also appreciate the thorough review and helpful suggestions of anonymous reviewers and the editorial assistant. The opinions expressed are those of the authors and do not necessarily represent the official views of any institution.* *Correspondence concerning this article should be addressed to Wagdy Loza, 4 Strathaven Pl., Kingston, Ontario, K7M 6S6; phone: 613-634-0227; e-mail: wml@post.queensu.ca.*

that it is useful in making these risk predictions (Hare, 1985; Harpur, Hakstian, & Hare, 1988; Rice & Harris, 1995; Serin, Peters, & Barbaree, 1990; Wong, 1987). The volume of recent publications demonstrates that the contribution and development of actuarial measures for predicting recidivism is gaining momentum.

To continue using these actuarial measures, it behooves us for ethical and professional reasons (i.e., to conform to the professional standards for developing psychological measures) to vigorously examine their reliability and validity, particularly their predictive validity. In fact, predictive validity is the most valuable tool for evaluating measures developed for predicting future behavior. Therefore, it is necessary to empirically assess the predictive validity of existing instruments over time with the same or different populations and different prediction parameters (Monahan, 1981).

The goal of this research was to examine the effectiveness of the SAQ in predicting release outcomes during a 5-year period. This goal is consistent with our plan to continue to validate the SAQ over time. Unlike the other measures, the SAQ is a self-report risk-need measure. It was designed to be multifaceted, covering the predominant predictive content areas that have been demonstrated to be reliable and valid for the assessment and prediction of postrelease offending. Other practical advantages for using the SAQ are that the SAQ is a self-report measure and as such, maximal objectivity is assured and the SAQ is more convenient and economical to use than are traditional methods of risk assessment. The instructions are simple and can be given by paraprofessionals. Offenders provide "yes" or "no" responses to the items, and it usually takes approximately 15 to 20 minutes to complete. The SAQ may also be administered in a group session. Interpreting the results is straightforward, requiring minimal professional time and minimal training. There is no need for prior extensive experience or special skills for credibly and reliably interpreting its results; as such, it is the most economical of the available tools. Other advantages are that more than 50% of SAQ items tap into dynamic factors; many of the offender's responses could be used as part of an individualized cognitive treatment plan; the SAQ is designed to help in the prediction of both violent and nonviolent recidivism; endorsement of some of the SAQ's statements could aid clinicians in the diagnosis of antisocial personality disorder, a history of

conduct disorder, and substance abuse; and the SAQ is scored without reliance on the results of other tools or sources.

The reliability and the construct and concurrent validity of the SAQ have been previously demonstrated (Loza, Dhaliwal, Kroner, & Loza-Fanous, 2000). The predictive validity of the SAQ during a 2-year period has also been demonstrated (Loza & Loza-Fanous, 2000). Furthermore, the SAQ was found to be an equivalent predictor of general and violent recidivism compared with three other risk prediction instruments (Kroner & Loza, 2001). Similarly, the SAQ demonstrated that it is at least as effective as four other well-established measures for predicting postrelease outcome (Loza & Loza-Fanous, 2001). This study was designed to be identical to the Loza and Loza-Fanous (2000) study that demonstrated the predictive validity of the SAQ during a 2-year period. However, the follow-up period for this study was extended to 5 years, and the "parole violations" criterion in the previous study was replaced with "committing new offenses." We felt that this criterion is more reflective of release failures than is the parole violation criterion. Based on our 2-year follow-up study, our hypothesis was that the SAQ would be valid for the 5-year period.

METHOD

PARTICIPANTS

A total of 599 male offenders who were incarcerated in various institutions in the Ontario region of the Canadian federal correctional system and who had completed the SAQ formed the sample pool for this study. Of the participants, 90% were selected randomly; they volunteered to complete the SAQ. The rest of the sample completed the SAQ during the process of completing assessments for release purposes. Of the 599 offenders who had completed the SAQ, 305 were released and followed up. The average age of the released participants was 30.69 (SD = 9.28, range = 17 to 68 years). Their current sentence length ranged from 2 years to life (M = 6.60 years, SD = 5.06). In calculating the means for sentence length, the life sentence for first-degree and second-degree murders was entered as 25 and 15 years, respectively. A total of 75% of the participants had been convicted

(past or current) for violent offenses (i.e., an offense against a person such as murder, kidnapping, rape, and assault). The nonviolent offenders were convicted of nonviolent crimes, such as property offenses (robbery without violence, break enter and theft, driving while impaired, and mischief). Whereas 89.23% ($n = 273$) of the participants were granted early release on parole, 10.77% ($n = 32$) were released at the end of their sentence. The average time lapsed between the completion of the SAQ and participants' releases was approximately 18 months. The racial composition of the participants consisted of Caucasians (79.40%, $n = 242$), African Canadians (7.97%, $n = 24$), and North American Natives, Asians, or those of "other" origin (12.63%, $n = 39$).

MEASURES

The SAQ is an empirically developed self-report instrument that is based on theoretical principles and assumptions regarding criminal behavior and is composed of 72 true or false items. The first six of the SAQ's eight subscales are used for prediction of recidivism. The Criminal Tendencies subscale taps into antisocial attitudes, beliefs, behaviors, and feelings. The Antisocial Personality Problems subscale covers characteristics similar to those used to diagnose antisocial personality disorder, which has been the psychiatric diagnosis traditionally used to predict recidivism. The importance of antisocial personality disorder in the prediction of recidivism has been repeatedly demonstrated (Andrews & Bonta, 1998). The Conduct Problems subscale assesses childhood behavioral problems. The Criminal History subscale assesses past criminality. The Alcohol and/or Drug Abuse subscale assesses substance abuse problems. Finally, the Antisocial Associates subscale taps into the offenders' antisocial associates and their influence on offenders' involvement in criminal activities. The other two subscales of the SAQ are the Anger and the Validity subscales. The Anger subscale could be used for assigning offenders to treatment programs dealing with anger. This scale consists of five items. These items are not included in the total score of the SAQ due to the controversial relationship between anger and recidivism (Loza & Loza-Fanous, 1999a, 1999b). Thus, the number of items used in the prediction of recidivism is 67 out of the total number

of the 72 SAQ items. The Validity subscale can be used to validate offenders' truthfulness in responding to the SAQ's items. Items included in this subscale are also included in the predictive subscales. Previous research (Loza et al., 2000) reported the psychometric properties of the SAQ. Briefly, the 1-week test-retest reliability coefficient was .95 for the total scale and ranged from .95 to .69 on the subscales. The total scale and subscale correlations ranged from .52 to .87. Subscale and subscale correlations, with the exception of the correlation between Criminal Tendencies and Alcohol and Drugs subscales, ranged from .25 to .58. The coefficient alphas for all SAQ subscales ranged from .42 to .87. Item-subscale correlations ranged from .19 to .76. The SAQ's subscales were correlated with other measures, which had demonstrated validity for assessing similar constructs. These correlations ranged from .28 to .65.

CRITERION MEASURES FOR POSTRELEASE FAILURES

Criteria used for this research consisted of failure in one or more of the following categories: (a) committed new offense as documented by official records; (b) recidivism (defined by returning to any form of custody, including county jails, for any reason such as parole violations, suspensions, or revocations of conditional release); (c) violent offense (i.e., offense against another person such as murder, rape, assault, or serious threats of violence), which was pulled out and specifically considered as its own category to determine the SAQ's ability to predict not only general failure but also offenses of a more violent nature; (d) any failure (includes the commission of any of the previous offenses as well as having been convicted for new charges, obtained negative parole reports from the parole supervisor, and committed any parole violation, e.g., consumption of alcohol or drugs).

PROCEDURE

Participants were sampled from several federal correctional institutions in the Ontario region. All completed SAQs were included in the data analysis. To determine the robustness of the SAQ, the Validity subscale was not used to validate the offenders' truthfulness in responding to the SAQs. Predictor variables were offenders' total and

subscale SAQ scores. Criterion measures for postrelease outcome were success or failure on each of the follow-up criteria. Follow-up data for postrelease outcome were collected every 4 months for 60 months (i.e., 15 periods) from the date offenders were released. The data were gathered from the Canadian Police Information Centre, a national criminal database maintained by the Royal Canadian Mounted Police. The follow-up periods were terminated (a) at the date offenders were "unsuccessful" on any of the postrelease criteria, (b) at the point where the offenders had reached 5 years from their release date, or (c) at the date of reviewing offender files for this study if the offenders had not completed 60 months of parole supervision. A total of 40% of the offenders failed on the first year of their release, and five, two, and one offenders failed on the second, third, and fourth years of their releases, respectively. All participants who completed the SAQ were included in this study (i.e., the Validity subscale was not used and no one was excluded as a result of potentially invalid responses).

RESULTS

The predictive validity of the SAQ was examined by first correlating the SAQ total and subscale scores with postrelease outcome criterion measures. The SAQ total score and all SAQ subscales correlated significantly with the criterion measures (see Table 1).

The SAQ's accuracy in predicting postrelease outcome was assessed by investigating the risk ratio for participants' failure on release. Using total SAQ scores, offenders were grouped, with approximately 33% (n = approximately 100) in each group, into low- (SAQ total score = 2 to 19), medium- (20 to 30), and high- (31 to 58) risk groups. These groups were compared on postrelease criterion measures using the low-risk group as a reference group for the medium- and high-risk groups. Table 2 shows that the risk ratios for failure in the medium-risk group ranged from 3.19 to 3.99 times higher than did those offenders in the low-risk group. Furthermore, the risk ratios for offenders failing in the high-risk group ranged from 4.88 to 10.69 times higher than did those in the low-risk group. Table 2

TABLE 1: Correlation Matrix of the Self-Appraisal Questionnaire (SAQ) Scales and Postrelease Criterion Measures (*n* = 305)

SAQ Subscales	*New Offense (*n = *63)*	*Recidivism (*n = *144)*	*Violent Acts (*n = *54)*	*Any Failure (*n = *167)*
Criminal Tendencies	.30***	.30***	.29***	.28***
Antisocial Personality Problems	.28****	.35****	.19***	.36****
Conduct Problems	.29****	.34***	.25***	.36****
Criminal History	.29****	.36***	.21****	.40****
Alcohol or Drugs	.22***	.32***	.26****	.37****
Associates	.12*	.16**	.20****	.20****
SAQ total	. 37****	.43****	.34****	.45****
Percentage of failure	28	47	18	55

NOTE: *n* = number of participants who were unsuccessful on outcome measures
*$p < .05$. **$p < .01$. ***$p < .0001$. ****$p < .0001$.

also shows that all differences between low- and medium-risk groups and low- and high-risk groups were statistically significant.

The probability of the SAQ correctly classifying offenders as either at risk or not at risk against the postrelease criterion measures was also examined to assess the predictive validity of the SAQ. Sensitivity and specificity methods were used to investigate conditional outcome probabilities. Sensitivity is the probability of the SAQ correctly identifying offenders at risk among those who actually failed on any of the postrelease outcome criterion measures. Specificity is the probability of the SAQ correctly identifying low-risk offenders among those who did not actually fail on the postrelease outcome criterion measures. Overall, the SAQ demonstrated sensitivity in the range of 68% to 85% and specificity in the range of 56% to 70%. The highest percentage of sensitivity was for the commission of violent acts, and the highest percentage of specificity was for any failure on release.

The next step in examining the predictive validity of the SAQ was through investigation of the SAQ's relative improvement of predictions over chance (RIOC). The RIOC is a measure to evaluate the predictive efficiency of a measure. It takes into account both the base rate (overall recidivism rates) and the selection ratio (the number of individuals identified as "unsuccessful," i.e., recidivated, according to the predictive measure). Loeber and Dishion (1983) and Mossman (1994) explained in detail this procedure. The SAQ showed strong RIOC on

TABLE 2 SAQ's Subscales & Total Number of Past Offenses, History of Violence, Number of Arrests, and Institutional Infractions

	Violent Offenses						*Total Number of Offenses*						*Number of Past Arrests*						*Number of Institutional Infractions*					
	Low (n = 44)		*High (n = 24)*				*Low (n = 44)*		*High (n = 42)*				*Low (n = 49)*		*High (n = 37)*				*Low (n = 46)*		*High (n = 50)*			
SAQ Subscales	M	SD	M	SD	t-*value*	r	M	SD	M	SD	t-*value*	r	M	SD	M	SD	t-*value*	r	M	SD	M	SD	t-*value*	r
Criminal tendencies	8.64	5.39	11.58	5.42	2.12*	.24*	8.09	4.11	10.93	5.82	2.60**	.24*	8.27	4.77	11.08	5.36	2.53**	.33**	8.52	4.77	10.58	5.50	1.84	.29**
Antisocial personality problems	2.21	1.46	2.46	1.35	0.69	.09	1.82	1.23	2.74	1.48	3.13***	.42***	1.90	1.23	2.77	1.53	2.79**	.31**	1.91	1.38	2.68	1.38	2.55**	.35***
Conduct problems	7.29	4.52	6.67	4.97	1.12	.18	5.98	4.11	8.95	4.60	3.16***	.25*	6.67	4.85	8.43	4.04	1.83	.32**	5.98	4.23	9.10	4.44	3.33***	.42***
Criminal history	2.95	1.43	3.29	1.88	0.77	.36**	2.00	1.38	3.38	1.38	6.7****	.56***	2.20	1.50	3.78	1.40	5.03***	.48***	2.56	1.63	3.25	1.61	1.95*	.18
Alcohol/ drugs	3.07	2.08	4.08	2.21	1.83	.24*	2.55	2.12	4.29	1.9	4.00***	.50***	2.96	2.13	3.97	2.17	2.16*	.42***	2.76	2.12	4.13	2.07	3.02***	.15
Associates	1.74	1.04	1.88	1.12	0.49	.07	1.43	1.02	1.93	1.05	2.23*	.14	1.43	1.12	2.00	.88	2.65**	.13	1.46	1.05	1.93	1.02	2.09*	.25**
SAQ total score	25.91	11.53	31.96	12.32	1.96*	.29**	21.86	10.16	32.64	11.31	4.64***	.43***	23.43	11.93	32.03	10.26	3.59***	.46***	23.20	10.85	31.65	11.72	3.45***	.41***

*$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$.

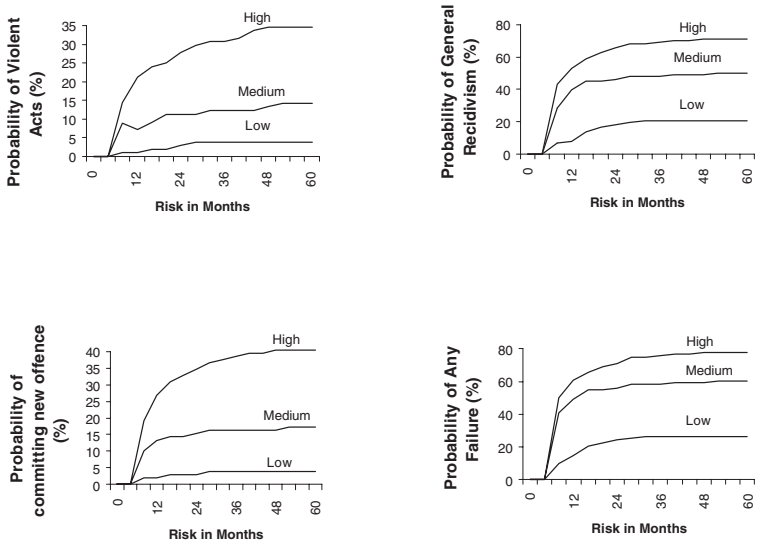

Figure 1: Probability of Failures of the Low- (SAQ = 2 to 19), Medium- (SAQ = 20 to 30), and High-Risk (SAQ = 31 to 58) Groups on Criterion Measures as a Function of Time

NOTE: SAQ = Self-Appraisal Questionnaire.

all the outcome criteria, varying from 41% to 69%. The 69% RIOC was for the commission of violent acts. Thus, the relationship between the SAQ and outcomes is well beyond chance (i.e., 50%). Therefore, 41%, for example, is 41% above a 50% rate. The RIOC statistic accounts and corrects for a chance relationship between the SAQ and the outcome criteria.

The final step in examining the SAQ was to determine the probability of failures for the low-, medium-, and high-risk groups on the postrelease criterion measures as a function of time. The log rank statistical test was used as part of the process of examining the survival analysis of each of the three groups and the comparison between them. The SAS computerized statistical package handles this procedure. Charts in Figure 1 indicate that the SAQ's high-risk group was always unsuccessful with more frequency and sooner than the medium- and low-risk groups on all postrelease criterion measures. Also, the medium-risk group was always unsuccessful with more frequency and faster than the low-risk group. All comparisons among the three risk groups presented in Figure 1 were significant at $p < .0001$.

DISCUSSION

The goal of this study was to examine the effectiveness of the SAQ in predicting release outcomes during a 5-year period. The results supported the initial hypothesis that the SAQ would be valid for making postrelease predictions during the 5-year period. The results were better compared with the findings in the 2-year follow-up study (Loza & Loza-Fanous, 2000). First, although the correlations between the SAQ total and subscale scores with outcome criterion measures were similar, in the previous study, two subscales (the Antisocial Personality Problems and the Alcohol and Drug) did not significantly correlate with the postcriterion of violent acts. Second, this study showed an improvement in both sensitivity and RIOC when compared with the 2-year follow-up study; whereas the sensitivity ranged from 55% to 72% in the previous study, in this study, it was 68% to 85%. Furthermore, whereas RIOC varied from 38.8% to 66.6% in the previous study, in this study, the RIOC varied from 41% to 69%. Third, the results of this survival analysis reflect some improvement over that of the 2-year follow-up study, particularly on the criterion of committing violent acts after release.

The improvement of the results in this study compared with those of the previous study may be due to the increase in the percentage of failure over time (base rate). This provides further evidence for the sensitivity and predictive validity of the SAQ as the SAQ was able to reflect these changes that occurred over time. In this study, the percentages of failures on the postrelease criteria were slightly higher than in the previous study. For example, whereas the percentages for failure in the previous study were 11.8% on the violent acts and 50.3% on the any-failure criteria, these percentages increased to 18% and 55%. This was expected because as time goes by, it is anticipated that more offenders would "fail," and the SAQ was able to predict this.

The results of this study lend further support to the predictive validity of the SAQ, particularly in predicting postrelease violent acts, which is the most important part of release predictions. This finding is consistent with the statement by Loza and Loza-Fanous (2000) that "it is expected that as the base rate for the violent acts upon release increase, so will the ability of the SAQ to predict violent behavior" (p. 1190).

Along with its demonstrated accuracy in predicting post-release failure, the SAQ presents advantages over most other actuarial measures (Loza et al., 2000; Loza & Loza-Fanous, 2000). Probably the most important advantage of the SAQ is that it is more convenient and economical to use than are traditional methods of risk assessment.

The SAQ still needs to be validated with different subtypes of offenders such as women, non-Caucasian populations, and sex offenders. Also, the SAQ needs to be cross-validated in different settings and examined for its predictive validity over longer periods of time. We suggest that future researchers examine these areas. Nevertheless, the results of this study are consistent with the results of the 2-year predictive study and once again provide support for the validity of the SAQ as a risk-need measure.

REFERENCES

Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto, ON, Canada: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (1998). *The psychology of criminal conduct* (2nd ed.). Cincinnati, OH: Anderson.

Boothby, J. L., & Clements, C. B. (2000). A national survey of correctional psychologists. *Criminal Justice and Behavior*, *27*, 716-732.

Gynther, M. D., Altman, A., & Warbin, R. W. (1973). A new actuarial-empirical automated MMPI interpretive program: The 4-3/3-4-code type. *Journal of Clinical Psychology*, *27*, 234-237.

Hare, R. D. (1985). Comparison of the procedures for the assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, *53*, 7-16.

Hare, R. D. (1991). *Manual for the Hare Psychopathy Checklist–Revised*. Toronto, ON, Canada: Multi-Health Systems.

Harpur, T. J., Hakstian, A. R., & Hare, R. D. (1988). Factor structure of the Psychopathy Checklist. *Journal of Consulting and Clinical Psychology*, *56*, 741-747.

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.

Holland, T. R., & Holt, N. (1975). Prisoner intellectual and personality correlates of offense severity and recidivism probability. *Journal of Clinical Psychology*, *31*, 667-672.

Kroner, D., & Loza, W. (2001). Evidence for the efficacy of self-report in predicting violent and nonviolent criminal recidivism. *Journal of Interpersonal Violence*, *16*, 168-177.

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin*, *94*, 68-99.

Lothstein, L. M., & Jones, P. (1978). Discriminating violent individuals by means of various psychological tests. *Journal of Personality Assessment*, *42*, 237-243.

Loza, W. (1996). *Self-Appraisal Questionnaire (SAQ): A tool for assessing violent and non-violent recidivism*. Unpublished manuscript.

Loza, W., Dhaliwal, G., Kroner, D., & Loza- Fanous, A. (2000). Reliability, construct, and concurrent validities of the Self-Appraisal Questionnaire: A tool for assessing violent and non-violent recidivism. *Criminal Justice and Behavior, 27*, 356-374.

Loza, W., & Loza-Fanous, A. (1999a). Anger and prediction of violent and non-violent offender's recidivism. *Journal of Interpersonal Violence*, *14*, 1014-1029.

Loza, W., & Loza-Fanous, A. (1999b). The fallacy of reducing rape and violent recidivism by treating anger. *International Journal of Offender Therapy and Comparative Criminology*, *43*, 492-502.

Loza, W., & Loza-Fanous, A. (2000). Predictive validity of the Self-Appraisal Questionnaire (SAQ): A tool for assessing violent and non-violent release failures. *Journal of Interpersonal Violence*, *15*, 1183-1191.

Loza, W., & Loza-Fanous, A. (2001). The effectiveness of the Self-Appraisal Questionnaire (SAQ) in predicting offenders' post-release outcome: A comparison study. *Criminal Justice and Behavior, 28*, 105-121.

McCreary, C. P. (1976). Trait and type differences among male and female assaultive and nonassaultive offenders. *Journal of Personality Assessment*, *40*, 617-621.

Mendelzys, N. (1979). Correlates of offense severity and recidivism probability in a Canadian sample. *Journal of Clinical Psychology, 35*, 897-907.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques*. Beverly Hills, CA: Sage.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, *62*, 783-792.

Nuffield, J. (1982). *Parole decisions-making in Canada: Research towards decision guidelines*. Ottawa: Solicitor General of Canada.

Quinsey, V. L. (1979). Assessments of the dangerousness of mental patients held in maximum security. *International Journal of Law and Psychiatry*, *2*, 389-406.

Rice, M. E., & Harris, G. T. (1995). Methodological development, violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology*, *63*, 737-748.

Serin, R. C., Peters, R. D., & Barbaree, H. E. (1990). Predictors of psychopathy and release outcome in a criminal population. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*, 419-422.

Wong, S. (1987). *The criminal and institutional behaviors of psychopaths*. Ottawa, Canada: Corrections Research and Development, Solicitor General Canada.

ERRATUM

PAUL GENDREAU
CLAIRE GOGGIN
PAULA SMITH
University of New Brunswick

A recent publication (Gendreau, Goggin, & Smith, 2002) has generated a great deal of discussion among correctional and forensic psychologists regarding the predictive supremacy of two very high-profile offender risk prediction measures: the Level of Service Inventory–Revised (LSI-R) and the Psychopathy Checklist–Revised (PCL-R).

A mistake was made, however, in the original publication with regard to one type of comparison of the predictive utilities of the two measures. It concerned the assessment of effect size heterogeneity, a procedure common to many meta-analyses. Upon recalculation, it was found that the distribution of effect sizes was not homogeneous as had been reported originally (Gendreau et al., 2002). Using procedures for the assessment of outliers recommended by Bonta, Hanson, and Law (1998) and Rosenthal (1991), several effect sizes were removed from three of the four measure by recidivism categories. Both the original and revised results are presented in Table 1 and suggest that if anything, removal of the outliers tended to improve the predictive validity of the LSI-R relative to that of the PCL-R.

For example, within the general recidivism category, the mean ϕ for the 22 LSI-R effect sizes was .35, whereas that of the 19 PCL-R effect sizes was .17. After weighting by sample size and number of effect sizes, the predictive superiority of the former was maintained (i.e., z^+ LSI-R = .32; z^+ PCL-R = .14). Consistent with the original results, the 95% confidence intervals (*CI*) about mean ϕ and z^+ for each measure again showed no overlap. The common language (*CL*) effect size indi-

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 30 No. 6, December 2003 722-724
DOI: 10.1177/0093854803256464

TABLE 1: Mean Effect Sizes for LSI-R and PCL-R by Type of Recidivism With and Without Outliers[a]

Predictor	*k*	*N*	M_{ϕ} *(SD)*	CI_{ϕ}	z^{+}	CI_{z+}	*Q*
General recidivism							
LSI-R	33	7,367	.37 (.12)	.33 to .41	.39	.36 to .41	115.30*
With outliers removed	22	2,681	.35 (.12)	.29 to .40	.32	.28 to .36	44.24*[b]
PCL-R	30	4,365	.23 (.15)	.17 to .28	.24	.21 to .27	101.71*
With outliers removed	19	1,975	.17 (.15)	.10 to .24	.14	.09 to .18	40.17*[b]
Violent recidivism							
LSI-R	16	3,297	.26 (.08)	.22 to .30	.28	.24 to .31	18.61
PCL-R	26	4,823	.21 (.09)	.17 to .25	.23	.20 to .26	52.93*
With outliers removed	24	3,853	.20 (.08)	.16 to .23	.20	.16 to .23	24.45

NOTE: LSI-R = Level of Service Inventory–Revised; PCL-R = Psychopathy Checklist–Revised; *k* = effect sizes per measure; *N* = offenders per measure; M_{ϕ} (*SD*) = mean phi and standard deviation between the measure and recidivism; CI_{ϕ} = confidence interval about mean phi; $z^{+} = \Sigma([z_{\phi}] \times [n-3] \div \Sigma[N-3])$ per measure per type of recidivism where z_{ϕ} = Fisher's transformed ϕ, *n* = sample per effect size, and *N* = sample per measure per type of maltreatment per type of criminal behavior; CI_{z+} = 95% confidence interval about z^{+}.

a. Common language effect size indicators for mean ϕ values with outliers removed; general recidivism: LSI-R versus PCL-R 82%; violent recidivism: LSI-R versus PCL-R 72%.

b. *Q* was significant but was reduced by 50% of its original value.

*$p < .05$.

cator (McGraw & Wong, 1992) revealed that 82% of the time, the LSI-R produced greater correlations with the criterion than did the PCL-R, up slightly from its original value of 78%.

In predicting violent recidivism, the respective effect sizes (i.e., ϕ, z^+) for the LSI-R were also greater than were those of the PCL-R once outliers were eliminated, with overlapping *CI*s only in the case of mean ϕ. Again, the *CL* favored the LSI-R 72% of the time compared with 66% in the original analysis.

In closing, it should be noted that some meta-analysts (i.e., Hunter & Schmidt, 1990) have made persuasive arguments against the elimination of outliers given the problems inherent in significance testing. Furthermore, in removing outliers one may exclude data from methodologically sound studies of large samples conducted by reputable researchers, as was the case in this reanalysis. Thus, for readers whose primary concern is not significance testing but estimation of effect size magnitude (i.e., point estimates with associated *CI*s, *CL* type statistics, and so forth, see Gendreau, 2002), it may be more useful to include all available effect sizes in comparing the predictive validity of measures, as in this instance.

REFERENCES

Bonta, J., Hanson, K., & Law, M. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, *123*, 123-142.

Gendreau, P. (2002). We must do a better job at cumulating knowledge. *Canadian Psychology / Psychologie canadienne*, *43*, 205-210.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? *Criminal Justice and Behavior*, *29*, 397-426.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

McGraw, K. O., & Wong, S. P. (1992). A common language effect size. *Psychological Bulletin*, *111*, 361-365.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.