CLASSIFICATION FOR EFFECTIVE REHABILITATION

Rediscovering Psychology

D. A. ANDREWS
Carleton University
JAMES BONTA
Ottawa-Carleton Detention Centre
R. D. HOGE
Carleton University

Four principles of classification for effective rehabilitation are reviewed: risk, need, responsivity, and professional override. Many examples of Case × Treatment interactions are presented to illustrate the principles.

Principles of classification for rehabilitation describe how particular classes of offenders may be linked with particular classes of discretionary service so that *effectiveness* of service is enhanced. "Effectiveness" has to do with achieving reductions in recidivism, "classes of offenders" refers to preservice differentiations based on the person and circumstances of offenders, and "discretionary service" refers to direct correctional service such as supervision, counseling, training, and treatment. The purpose of this article is to review four principles of classification for rehabilitation within the context of basic research and theory in the psychology of criminal conduct.

Risk of recidivism, criminogenic need, and the responsivity of offenders to different service options are the characteristics of offend-

AUTHORS' NOTE: *Request for reprints should be sent to D. A. Andrews, Department of Psychology, Carleton University, Ottawa, Canada K1S 5B6.*

CRIMINAL JUSTICE AND BEHAVIOR, Vol. 17 No. 1, March 1990 19-52

ers that may determine level, targets, and type of rehabilitative effort. The fourth principle has to do with the responsibility of professionals to step beyond routine application of risk, need, and responsivity when circumstances so warrant. These principles may be summarized as follows:

1. *Risk*: Higher levels of service are reserved for higher risk cases. In brief, intensive service is reserved for higher risk cases because they respond better to intensive service than to less intensive service, while lower risk cases do as well or better with minimal as opposed to more intensive service.
2. *Need*: Targets of service are matched with the criminogenic needs of offenders. Such needs are case characteristics that, when influenced, are associated with changes in the chance of recidivism. If reduction in the chances of recidivism is an ultimate goal, the more effective services are those that set reduced criminogenic need as intermediate target of service.
3. *Responsivity*: Styles and modes of service are matched to the learning styles and abilities of offenders. A professional offers a type of service that is matched not only to criminogenic need but to those attributes and circumstances of cases that render cases likely to profit from that particular type of service.
4. *Professional override*: Having considered risk, need, and responsivity, decisions are made as appropriate under present conditions.

Before reviewing these principles in detail, we will plant them firmly within the psychology of criminal conduct. If the analysis of rehabilitation does not occur within a favorable intellectual framework, the antirehabilitation rhetoric of mainstream criminology will contine to retard progress.

PSYCHOLOGY OF CRIMINAL CONDUCT

The psychology of criminal conduct provides a stimulating and facilitative home for the analysis and development of rehabilitation. Unfortunately for the development of rehabilitation, the psychology of criminal conduct has been discounted for years within major sectors of mainstream criminology (Andrews & Wormith, in press). This is obvious when the objectives and assumptions of psychology and

mainstream criminology are compared. Comparisons can be made with regard to focus (choice of dependent variable), choice of independent variables, and the roles afforded clinical service and systematic empirical research. First, the psychology of criminal conduct is concerned with understanding variation in the criminal conduct of individuals (e.g., recidivism), while mainstream criminology is preoccupied with interpreting aggregated crime rates, law and order, and with overcoming structured inequality in the distribution of societal wealth and power (for examples, see Barlow, 1986; Canada, 1987; Cohen, 1985; Maclean, 1986; Martinson, 1976; Schur, 1973; Taylor, Walton, & Young, 1973; Vold & Bernard, 1986). Second, psychology seeks knowledge of the full range of biological, personal, interpersonal, circumstantial, and structural/cultural covariates of criminal conduct, while the thrust of mainstream criminological theory and criticism—from beginning works (see Hirschi & Hindelang, 1977) and on through current scholarship (see Andrews & Wormith, in press)—has been to deny and discount the importance of risk and need factors at the individual level. Third, the psychology of crime is open to and has experience with the clinical tradition in general psychology, while vast sections of mainstream criminology believe that clinical service is ineffective, perhaps "evil," and certainly not as powerful or as dignified as punishment.[1] Finally, psychology endorses knowledge construction wherein an unsparing identification of potential errors in the establishment of covariates leads to the actual exploration of the effects of errors of both conceptualization and measurement. In contrast—and with the antidifferentiation and antitreatment biases that exist within criminology—failures to establish covariation are accepted as unthreatened evidence that we know nothing about individual criminal conduct. At the same time, research revealing covariation is subjected to intense criticism of a variety in which it is asserted that all potential threats to validity indicate that we know very little (Andrews & Wormith, in press; Gottfredson, 1979).

We do not mean to imply that psychology has been free of tension. From the perspective of professional training, psychology has been at best neutral if not adverse to the notion of training clinical and social psychologists in the criminal justice area. Promising university programs are appearing (Ogloff, 1988), but many factors within general

psychology have inhibited the interest of clinicians and researchers in rehabilitation. For example, a human propensity for antisocial conduct is central to psychoanalytic theory, and yet many psychodynamic theorists are more interested in neurotic misery than conduct disorder. It is as if psychoanalysts found so obvious the early familial, personality, and situational sources of antisocial behavior—parents as poor models, trainers, and objects of identification; low levels of guilt; early and generalized misconduct; poor impulse control—that they focused on the more esoteric and "deep" intrapsychic factors. Conversely, humanists reacted against the idea that antisocial behavior was natural in the absence of controls, suggesting instead that such behavior reflected the operation of controls and other inhibitors of natural goodness. Thus, humanistic therapists too focused on freeing human potential rather than predicting or influencing antisocial propensity.

Certain antiassessment themes are also apparent in the clinical tradition. Our psychodynamic colleagues ask, "What is this risk/need stuff anyway, why not rely on DSM diagnoses like we do?" The humanists ask, "Why bother with assessment when we know that all clients will profit from openness, warmth, and understanding?" Even within the field of general psychotherapy research, the potential of assessments of risk are often misunderstood. Until very recently, reviews of the predictors of outcome within therapy samples frequently left readers with the false impression that only low-risk cases "profited" from service. As we shall soon see (in Table 3), an answer to the question of "who profits from service" requires comparisons not *within* but *between* treatment groups at each level of risk.

With the above caveats in mind, a psychological "understanding" of criminal conduct is crucial to effective correctional programming. The major sources of variability in recidivism reside within the conditions established by the specifics of a sentence, and there we may identify and gain some influence over "factors that make a difference" (Andrews, 1982a). These important factors include preservice characteristics of offenders, specifics of the process and content of services planned and delivered, intermediate changes that occur in the person and circumstances of individual offenders, and their interactions on recidivism (Andrews, 1980, 1983; Andrews & Kiessling, 1980; Glaser, 1974; Hoge & Andrews, 1986; Palmer, 1974; Warren, 1969).

A psychological understanding involves knowledge of the covariates of criminal conduct and of the moderators and mediators of that covariation. The covariates may be correlates of a criminal past (uncovered through cross-sectional research), static predictors of a criminal future (uncovered through uniwave longitudinal research), dynamic predictors of a criminal future (established by studies of change in multiwave longitudinal research), and functional variables (identified through approximations of experimental ideals). Thus the risk and need principles draw upon psychology for knowledge of predictors and of dynamic predictors, respectively. The responsivity principle depends upon knowledge of how different modes or styles of behavioral influence may be differentially effective with different types of people. The override principle recognizes that empirically based decision making is, by definition, founded on existing evidence and that special conditions having to do with settings, offenders, workers, and managers may render that extant data base less relevant than it is under routine conditions. Here the professional moves beyond routine practice for specified reasons.

Interest in "differential treatment" or "matching" has grown in the last decade. This is true in human service generally (Beutler, 1979) and in corrections particularly as revealed by many detailed expositions and state-of-the-art reviews. Noteworthy are edited books on the topic (e.g., Gottfredson & Tonry, 1987), special issues of journals (*Crime & Delinquency*, July 1986; *Criminal Justice and Behavior*, March 1988), and handbooks (e.g., Clements, 1986). The remainder of the present article builds on those earlier reviews and finds *effective* service to be a function of risk, need, and responsivity. It is unfortunate that state-of-the-art reviews (e.g., Farrington & Tarling, 1985; Posey, 1988; Sechrest, 1987) continue to imply that risk assessments are somehow limited to "management" concerns and only indirectly relevant to "treatment."

RISK PRINCIPLE

Two aspects of the risk principle — prediction and matching — require discussion. The assessment of risk and the prediction of recidivism

have received most of the attention from researchers, while matching has only recently been given due consideration.

PREDICTION ISSUES

Prediction involves assessment of risk factors. Risk factors refer to personal attributes and circumstances that are assessable prior to service and are predictive of future criminal behavior. Our specification of factors assessable prior to sevice indicates that we focus here on uniwave prediction (that is, Time 1 predictor scores in relation to assessments of subsequent criminal behavior).

Contrary to antidifferentiation rhetoric in mainstream criminology, general offender samples may be sorted into risk categories with significant levels of predictive criterion validity. This has been evident from the earliest days of systematic research on the issue (see Briggs & Wirt, for a review of pioneering British and U.S. studies). For example, Burgess (1928) scored 3,000 paroled men on 21 variables and found scores to be highly related to outcome (e.g., in the best-risk category only 1.5% violated parole, compared with 76% in the poorest-risk category). Similarly, Glueck and Glueck (1930) found seven variables producing highly efficient predictions. More recently, as inspection of the first two columns of Table 1 reveals, a number of classification systems have shown predictive criterion validity, some in relation to both inprogram and postprogram indicators of antisocial behavior. The articles cited in Table 1 will introduce the classification systems to readers unfamiliar with them, but we will not review each system here. Rather, inspection of Table 2 reveals the levels of predictive accuracy that are now routinely achieved when the base rates of recidivism are in the 30% to 60% range. Facts concerning the predictability of recidivism have existed in psychology for decades. Sophisticated meta-analyses have recently served to confirm the conclusions of early investigators (e.g., Loeber & Dishion, 1983).

Future Efforts

Three directions are indicated in regard to preservice risk assessments. The first involves education and training of professionals in the

TABLE 1: Examples of Studies of the Predictive Validities of Risk and Need Classification Systems

System	*Study*	*Subjects*	*Outcome Criteria* *In*	*Post*	*Matching*
MMPI	Megargee & Bohn (1979)	Inmate	+	+	
	Edinger (1979)	Inmate	+		
	Edinger & Auerbach (1978)	Inmate	+		
	Walters et al. (1986)	Military	+		
	Bohn (1980)	Inmate	+		+
	Wright (1986)	Inmate	+		+
	Hanson et al. (1983)	Inmate	+		
	Motiuk et al. (1986)	Inmate	+	–	
	Louscher et al. (1983)	Inmate	–		
	Van Voorhis (1988)	Inmate	–		
I-Level	Barkwell (1980)	Juvenile			+R
	Palmer (1974)	Juvenile			+R
	Jesness (1975)	Inmates			+
	Van Voorhis (1988)	Inmate	–		
CL	Brill (1978)	Juvenile			+
	Leschied & Thomas (1985)	Juvenile			+
	Leschied et al. (1985)				+
	Van Voorhis (1988)	Inmate	–		
Quay	Quay & Love (1977)	Juvenile		+	
	Quay (1984)	Inmate	+	+	+
	Levinson (1988)	a, Inmate	+		+
		b, Inmate	+	+	+
LSI	Andrews (1982b)	P & P	+	+	
	Andrews et al. (1986)	P & P			+R
	Bonta & Motiuk (1985)	Inmate	+	+	
	Bonta & Motiuk (1987)	Inmate	+	+	+
	Bonta (in press)	Inmate	+	+	
	Motiuk et al. (1986)	Inmate	+	+	
CMC	Baird et al. (1979)	P & P		+	+
	Wright et al. (1984)	P & P		+	
	Lerner et al. (1986)	P & P			+
		P & P			+

R = Random assignment
P & P = Probation and Parole
+ Positive findings
– Negative findings

TABLE 2: Examples of the Predictive Criterion Validity of Intake Risk Assessments with the LSI

	LSI Risk Level at Intake				
	Low	*Moderate*	*High*	*Very High*	*RIOC*
Recidivism Rates of Adult Probationers (Andrews & Robinson, 1984)					
Sample 1: Ottawa	.11	.23	.49	.74	53%
(F-U: 814 days)	(211)	(105)	(210)	(35)	(561)
Sample 2: Ottawa	.10	.16	.37	.75	43%
(F-U: 28 mos)	(70)	(56)	(62)	(4)	(142)
Sample 3: B-ville	.05	.27	.41	.75	56%
(F-U: 18 mos)	(20)	(11)	(22)	(4)	(57)
Reincarceration Rates of Incacerates (PSR LSI Study)					
Ottawa Courts	.00	.06	.38	.45	91%
(F-U: 707 days)	(20)	(18)	(53)	(33)	(124)
Recidivism Rates of Young Offenders (YLSI; Scott, 1985)					
Ottawa Probation		.05	.53		83%
(F-U: 6-18 mos)		(41)	(43)		(84)
Reincarceration Rates (Bonta & Motiuk, 1987)					
Sample 1: Ottawa		.27	.52		38%
(F-U: 365 days)		(37)	(69)		(106)
Sample 2: Ottawa		.10	.59		76%
(F-U: 365 days)		(39)	(205)		(244)

RIOC: Relative Improvement Over Chance (Loeber & Dishion, 1983) is a measure of predictive accuracy that is somewhat less sensitive to base rate and selection ratios than are alternative summary measures of predictive accuracy.
F-U: Duration of follow-up

facts about prediction in corrections. It is unacceptable that many professionals and students in criminal justice are unaware of basic research on prediction. Reflecting the content of mainstream criminological textbooks, many professionals appear to know many "reasons" why prediction is impossible and yet are unaware of the actual data. In our consultation and training experiences, we are amazed to discover the many professionals and senior students who have never actually seen simple contingency tables that document how risk levels relate to recidivism rates. Training should compare directly the pseu-

doscientific knowledge destruction techniques so widespread in criminology textbooks (see Andrews & Wormith, in press) with the more intellectually serious business of reviewing threats to validity for purposes of knowledge construction. Training must also consider the value of systematic assessment as a prerequisite to clinical decision making. Statistically based predictions have been found to be as good or better than clinical prediction (Glaser, 1987; Meehl, 1954), yet systematic risk assessment is resisted or, following short-term acceptance, slowly reverts to unsystematic assessment.

The second effort involves refinement of extant instruments such as the Wisconsin Risk-Need Assessment (Baird, Heinz, & Bemus, 1979), Recidivism Prediction (Nuffield, 1982) and Level of Supervision Inventory (LSI; Andrews, 1982b) in the areas of user convenience, contextual appropriateness, efficiency, ethicality, "different versions for different folks," and differential predictability of different types of offenses (prediction of violence remains a high priority).

The third effort is ultimately more important than the first two. Here, the task is to explore the upper limits of predictive accuracy. Into the late 1970s it was widely believed that the .30/.40 "sound barrier" could not be broken (Monahan, 1981). Now definite improvements in predictive accuracy are possible when we move beyond mere fine-tuning of current practice. Current limits on uniwave prediction may be traced to inappropriate sampling of the predictor domain through errors of conceptualization and inadequate sampling of both the predictor and criterion domains through errors of measurement (Andrews, Wormith, & Kiessling, 1985; Glaser, 1987).

With reference to errors of conceptualization, the predictability of recidivism may be increased through the inclusion of the personal, interpersonal, and circumstantial variables suggested by psychodynamic (e.g., Glueck & Glueck, 1950), control (e.g., Hirschi, 1969), and social learning (e.g., Akers, 1973; Andrews, 1980; Ross & Fabiano, 1985) theories and the exclusion of those variables having to do with social origins that are the heart of anomie and subcultural theories. The empirical fact is that class of origin and constructs such as alienation and strain cannot bear the weight they have been assigned in anomie and subcultural theory (Andrews et al., 1985; Hirschi, 1969; Johnson, 1979; Tittle, Vilimez, & Smith, 1978; Zamble & Porporino,

1988). Rather, the upper limits of the predictive criterion validity of risk assessments reside in the five key indicators of antisocial propensity: behavior, feelings, cognitions, personality, and associates. A history of antisocial behavior is best assessed broadly in terms of a variety of antisocial acts, evident in a variety of settings, committed from a young age and extending into adolescence (Loeber, 1982). Look too for antisocial activity resulting in official processing and even continuing while official penalties are being served (Andrews et al., 1985; Loeber, 1982; Loeber & Dishion, 1983; see content of risk scales cited in Table 1). Also indicated are assessments of antisocial personality, feelings, cognitions, and associates that do not rely heavily upon the already sampled facts of a self-reported or official criminal history (Andrews et al., 1985, 1986; Motiuk, Bonta, & Andrews, 1986). Other candidates for assessment include unstable employment record/poor academic achievement and, among young people, conflicts with parents and exposure to poor parenting (Andrews, 1983; Loeber & Dishion, 1983).

With reference to errors of measurement, at least two actions are indicated: multimethod and multidomain assessments of predictor variables, and multimethod assessments of the criterion variable over extended follow-up periods (Andrews, 1983; Andrews et al., 1985). Psychology also demands explorations of construct validity in which various methods of risk assessment are explored in studies yielding convergent, divergent, and predictive validity estimates (Andrews, 1983; Andrews, Kiessling, Mickus, & Robinson, 1986; Motiuk et al., 1986).

THE MATCHING ISSUE

On the basis of their cross-sectional research in the 1940s, Glueck and Glueck (1950) were among the first to suggest that the purpose of systematic risk assessments was to identify those high-risk cases that required intensive service. Inspection of column three of Table 1 reveals that risk assessments conducted with some instruments have been found to interact with service variables. Table 3 presents some concrete illustrations of the interaction of risk and level of service across various measures of risk, types of service, measures of out-

come, and types of subjects. The findings of every study in Table 3 indicate that the more intensive service option was either unrelated to outcome with low-risk cases or associated with significantly poorer outcome than was less intensive service. Conversely, every study in Table 3 reveals that the effects of service were positive and significant among higher risk cases, albeit not always among the highest risk cases. This pattern was evident among juveniles exposed to a companionship program (O'Donnell, Lydgate, & Fo, 1971), crisis-oriented family counseling (Byles & Maurice, 1982), and a child welfare project (Andrews, Robinson, & Balla, 1986). It was also found among probationers exposed to relatively intensive supervision in Wisconsin (Baird et al., 1979) and Ontario (Andrews & Kiessling, 1980; Andrews, Kiessling, Robinson, & Mickus, 1986). Interestingly, the Andrews and Friesen (1987) example reveals the differential effects of self-management efforts by lower and higher risk probationers. Table 3 reveals similar results with regard to who profits from family therapy (Andrews, Hoge, Robinson, & Hoge, 1986), problem-solving training in preschool (Spivack & Shure in Rappaport, 1977), and services for opiate addicts (Woody et al., 1984) and discharged mental patients (Kirk, 1976). Even at the aggregate level in the area of crime prevention, helicopter patrols were more effective in high-crime areas than low-crime areas (Kirchner et al., 1980).

The findings in Table 3 were selected for purposes of illustration and should not imply that interactions are always found. The validity of risk principle may depend upon the particular assessments of risk and outcome employed and almost certainly depends upon the targets and types of service employed (Andrews, Kiessling, Robinson, & Mickus, 1986; Andrews, Robinson, & Balla, 1986; Beutler, 1979). For example, Andrews and Kiessling (1980) reported that high levels of an inappropriate service (nondirective "active listening") was associated with *increased* recidivism among high-risk probationers.

In summary, the findings in Table 2 illustrate the predictive validity of preservice risk assessments, while the findings in Table 3 illustrate that predictability of outcome may be enhanced through consideration of preservice risk, service, and their interaction. Turning to criminogenic need and multiwave studies, we find that the predictability of

TABLE 3: Examples of Risk × Service Interactions in Corrections, Child Welfare, Family Service, Preschool and Other Settings

Intake Risk Level	*Level of Service*				*p <*
	Less		*More*		
Arrest Rates of Juveniles (O'Donnell et al., 1971)					
High	.78	(23)	.56	(50)	*
Low	.16	(195)	.22	(285)	*
Arrest Rates of Juveniles (Byles & Maurice, 1982)					
High	.92	(12)	1.00	(18)	ns
Moderate	.92	(25)	.57	(42)	*
Low	.43	(114)	.57	(94)	ns
Out-of-Home Placement Rates in Child Welfare (Andrews et al., 1986)					
High	.72	(32)	.38	(42)	*
Low	.20	(30)	.36	(42)	ns
Reoffending Rates of Wisconsin Probationers (Baird et al., 1979)					
High	.37	(113)	.18	(113)	*
Moderate	.18	(71)	.13	(71)	ns
Low	.03	(58)	.10	(58)	ns
Recidivism Rates of Adult Probationers (Andrews & Kiessling, 1980)					
High	.58	(23)	.31	(34)	*
Low	.12	(62)	.17	(58)	ns
Mean Probation Negative Outcome Scores (Andrews et al., 1986)					
High	1.58	(12)	1.11	(54)	*
Low	.25	(28)	.65	(98)	*
Recidivism Rates of Adult Probationers (Andrews & Friesen, 1987)					
High	.92	(12)	.25	(4)	*
Low	.12	(8)	.29	(17)	ns
Favorable Outcome Scores in Family Service (Andrews, Hoge, Robinson, & Hollett, 1986)					
High	3.00	(6)	2.29	(14)	ns
Moderate	1.57	(7)	3.00	(32)	*
Low	3.40	(10)	3.41	(41)	ns
Maladjustment Rates: Problem Solving Training (Spivak in Rappaport, 1977)					
High	.75	(56)	.40	(72)	*
Low	.14	(50)	.10	(41)	ns

TABLE 3 Continued

Intake Risk Level	*Level of Service*				*p <*
	Less		*More*		
Illegality Outcome Scores of Opiate Addicts (Woody et al., 1984)					
High	108	(11)	91	(21)	*
Moderate	73	(16)	71	(28)	ns
Low	52	(12)	60	(22)	ns
Readmission Rates of Discharged Mental Patients (Kirk, 1976)					
High	.50		.29		*
Low	.30		.33		ns
Burglary Rates of Neighborhoods (Kirchner et al., 1980)					
High	2.83		1.22		nr
Low	1.65		2.14		nr

NOTE: This method of illustrating Case x Treatment interactions is preferable to "matched" versus "mismatched" comparisons because the latter comparisons are too easily confounded with risk levels.
(N in parentheses)
ns: not significant
nr: not reported
*$p < .05$

recidivism increases still further when actual changes in the person and circumstances of offenders are monitored.

NEED PRINCIPLE

Criminogenic needs are a subset of risk factors. They are dynamic attributes of offenders and their circumstances that, when changed, are associated with changes in the chances of recidivism. Clearly not all "needs" are criminogenic. Criminogenic needs are ones in which assessments of change (or retests) possess a level of predictive criterion validity that is incremental to the criterion validity of pretests. In brief, in multiwave longitudinal studies assessments of dynamic variables are conducted at Times 1 and 2. Then, the scores at Time 2 (or the differences between scores at Times 1 and 2) are examined in relation to subsequent recidivism, with the contribution of Time 1

TABLE 4: Two Examples of the Predictive Criterion Validity of Retest Scores with Intake Scores Controlled

(A) Three Year Postprobation Recidivism Rates by Six Month Retest Procriminal Attitude Scores and Intake Risk (Andrews & Wormith, 1984)

Risk Level At Intake	*Retest Risk Level* Low	Moderate	High	Overall
High	07 (1/14)	43 (6/14)	40 (4/10)	29 (11/38)
Moderate	10 (2/20)	37 (9/24)	57 (8/14)	33 (19/58)
Low	10 (4/38)	20 (3/15)	67 (2/3)	16 (9/56)
Overall	10 (7/72)	34 (18/53)	52 (14/27)	19 (39/152)

(B) Recidivism Rates by Risk Level at Last Available of Quarterly LSI Reassessments and Intake LSI Risk Level (Andrews & Robinson, 1984)

Risk Level At Intake	*Retest Risk Level* Low	Moderate	High	Very High	Overall
Very High	—	—	50	100	75 (4)
High	—	00	27	100	41 (22)
Moderate	00	33	40	—	27 (11)
Low	00	00	00	100	05 (20)
Overall	00	10	32	100	28
	(19)	(10)	(19)	(9)	(57)

NOTE: In first row of "A", 38 cases scored high risk at intake on the attitude measure, and 11 (29%) were reconvicted over the follow-up period. However, those initially high scorers whose risk levels had dropped to low levels at six months had a recidivism rate of only 7% (1/14).

scores statistically controlled. Table 4 provides two illustrations of findings from multiwave longitudinal investigations. Inspection reveals that the six-month retest scores of probationers on a measure of antisocial attitudes were much more predictive of recidivism than were attitudes as assessed at intake. The other illustration reveals similar results with retests on the LSI.

Still more convincing evidence of criminogenic need arises from controlled program evaluations in which deliberately induced changes on intermediate targets (that is, suspected criminogenic need factors) are examined in relation to recidivism. Here we look for functional links among variation in service, changes on intermediate targets, and recidivism. The findings are most impressive in cases in which the

level of covariation between treatment and recidivism is reduced significantly when statistical controls are introduced for changes on assessments of intermediate targets. This pattern of results would affirm the functional validity of our selection of intermediate treatment goals. This is the type of informtion that is required for the systematic criterion validation of assessments of targets (Andrews, 1982a, 1983; Andrews & Kiessling, 1980; Hoge & Andrews, 1986; Kazdin, 1985; Mash, 1985; Nelson & Hayes, 1979).[2]

While the vast amount of research on criminogenic need has been cross-sectional or uniwave longitudinal, social psychological theory is highly suggestive regarding criminogenic need. For example, most theories of criminal conduct support the criminogenic significance of procriminal cognitions (or personal attitudes, values, and thinking styles favorable to violation of the law). Procriminal sentiments are basic to psychodynamic and social control perspectives (weak superego, disbelief in the validity of the law), differential association theory (definitions favorable or unfavorable to violations of the legal code), subcultural theory (internalization of norms), labeling theory (criminal identifications), anomie theory (internalized prohibitions), conflict theory (the value context of human choice), and, of course, integrative social learning theories.

Moreover, statistical modeling of cross-sectional data has consistently confirmed that deviant attitudes are among the variables most strongly correlated with criminal behavior and that most effectively serve as the mediators of the effects of other correlates of delinquency in domains such as family, school, and peer associations (Matsueda, 1988). The only serious threats to procriminal sentiments as need factors are assessments of antisocial associates and antisocial personality.

Multiwave longitudinal studies suggest that the criminogenic status afforded procriminal sentiments in theory and cross-sectional research is warranted. Inspection of Table 5 reveals that reassessments of antisocial attitudes, antisocial associates, antisocial personality, trouble at school/home, and drug abuse were incrementally predictive of recidivism over and above the predictability yielded by intake assessments. Conversely, the predictive validity of assessments of the personal distress of probationers (anxiety, alienation, low self-esteem) was very low both at intake and upon reassessment. In fact, additional

TABLE 5: Dynamic Predictive Criterion Validity Estimates for Various Potential Criminogenic Need Factors in Two Samples of Probationers (Andrews & Wormith, 1984)

Construct Assessed	*Sample 1*	*Sample 2*
Antisocial Attitudes	.56 (108/194)	.36 (039/111)
Antisocial Associates	na	.17 (015/088)
Antisocial Personality	.18 (024/124)	.24 (051/212)
Trouble at School/Home	.39 (036/096)	.36 (046/126)
Drug Abuse	na	1.41 (031/022)
Alienation	ns	.13 (010/077)
Conventional Success Orientation	ns	ns
Empathy	ns	ns
Personal Distress (High Anxiety/ Low Self-Esteem)	ns	ns

NOTE:. The estimates are the increase in R square obtained by introducing six-month retest scores, expressed as the proportion of the R square yielded by intake scores alone. Thus, for example, in Sample 1, there was a 56% gain in the predictability of recidivism through consideration of retest attitude scores (and the overall R square with both intake and retest scores entered as predictors was .302 [.108 + .194]). Drug abuse estimates are from Voss (1982).
na: not available
ns: nonsignificant

evidence (not presented in Table 5) suggests that increased self-esteem, unless accompanied by anticriminal gains, may be criminogenic in offender samples (Andrews, 1983; Wormith, 1984).

Considerable effort has been applied to developing reliable and objective assessments of the needs of various offender groups (Clements, 1986; Duffee & Duffee, 1981). Correctional professionals may well be interested in addressing the multiple needs of offenders, including areas of need that are irrelevant to recidivism but that do suggest the importance of a caring and supportive environment. From the perspective of rehabilitation, however, "needs" have all too often been assessed independently of recidivism. Both the Wisconsin Risk-Need Assessment (Baird et al., 1979) and the LSI (Andrews, 1982b) systems use a variety of information ranging from stable risk factors (e.g., criminal history) to dynamic indices (e.g., employment). No dynamic predictive validity estimates in regard to criminal behavior have been reported for the Wisconsin system, and only two small-scale studies of the validity of assessments of change have been conducted with the

LSI (Andrews & Robinson, 1984). The results were promising in that reassessment of LSI risk-need scores were more strongly correlated with recidivism than were LSI intake scores (Table 4).

Psychometric approaches to the assessment of needs/targets remain high priority issues in rehabilitation in particular and in human service in general. One especially interesting issue has to do with the possibility that different types of offenders may present different "needs" (Lukin, 1981). For example, while increasing self-esteem may be an inappropriate target for antisocial offenders, it may be an appropriate one for neurotic/depressed offenders. In addition, structured need assessments may be best supplemented by detailed interviews that focus on the specific circumstances surrounding prior criminal acts. Finally, research on criminogenic need may assess and reassess potential needs beyond those targeted within particular programs and, better still, if time-varying covariates (targeted and nontargeted) continue to be assessed over extended follow-ups (Andrews, 1983).

RESPONSIVITY PRINCIPLE

Having established risk and criminogenic need, the third classification issue has to do with the responsivity of offenders to different styles and modes of service. Here the relevant research and theory is of two types—the literature on the effectiveness of service within general offender samples, and that having to do with how characteristics of offenders may interact with style and mode of service. We begin with the general effectiveness literature.

GENERAL LITERATURE

Sentenced offenders are not undergraduates attending a university clinic, business people seeking a weekend of psychorecreation, or mental health clients. Thus the literature on the effectiveness of correctional treatment in particular is important. It is composed of four sets of studies (although debates continue regarding which studies belong in which set). These sets are: (1) studies with methodological problems so serious that no conclusions could be drawn, (2) reason-

ably well-controlled studies in which treatment effects were not established, (3) reasonably well-controlled studies in which one treatment (whether labeled "experimental" or "control") was found to be significantly more effective than another, and (4) reasonably well-controlled studies in which the relative effectiveness of treatments was dependent upon type of client (or setting or worker).

Reviewers of the effectiveness literature have consistently found that at least 40% and up to 80% of the better-controlled studies reported significant treatment effects. For example, positive effects of treatment were found in 75% (3/4) of the better-controlled studies reviewed by Kirby (1954). The corresponding figures were 60% (13/22) in Bailey (1966), 50% (9/18) in Logan (1972), 58% (19/33) in Andrews (1974), 86% (81/95) in Gendreau and Ross (1979), 43% (10/23) in Hollett (1984), 50% (25/50) in Ross and Fabiano (1985), and approximately 50% in both Martinson (1974) and Lab and Whitehead (1988), depending upon the particular studies surveyed. Notably, these rates do not include studies in which the treatments designated "experimental" were less effective than those designated "control." If these studies were interpreted as examples of the superiority of the "control" as a serious service option, the rates for treatment would increase even above the values just reviewed. For example, Craft, Stephenson, and Granger (1966), Murphy (1972), and Klein (1971) found that relationship-oriented and cohesion-building group approaches were significantly less effective than "firm-but-fair" approaches. While a "miss" from the perspective of nondirective group therapy, the findings represent a "hit" from the perspective of the value of firm-but-fair approaches to treatment.

Drawing upon many reviews (Andrews, 1974, 1979, 1980, 1982; Cullen & Gendreau, 1989; Gendreau & Ross, 1979, 1987; Hollett, 1984; Ross & Fabiano, 1985), some things are known about the characteristics of effective programs in general correctional samples. Indeed, the conclusions are sufficiently strong to inform professionals in rehabilitation and to lead to policy statements that actively encourage rehabilitative effort and evaluation of that effort.

Paraphrasing Andrews and Kiessling (1980, pp. 462-463), effective rehabilitative efforts involve workers who are interpersonally warm, tolerant, and flexible, yet sensitive to conventional rules and proce-

dures. These workers make use of the authority inherent in their position without engaging in interpersonal domination (firm but fair); demonstrate in vivid ways their own anticriminal/prosocial attitudes, values, and beliefs; and enthusiastically engage the offender in the process of increasing rewards for noncriminal activity. The worker exposes and makes attractive concrete alternatives to procriminal attitudes and behavior. The worker does not depend upon the presumed benefits of a warm relationship with the offender, nor does he or she assume that the offenders will self-discover these alternatives. The alternatives are demonstrated through words and action, and explorations of the alternatives are encouraged through modeling, reinforcement, and concrete guidance.

Hollett (1984) reviewed 23 controlled outcome studies in which impact on recidivism could be examined as a function of the presence of authority, anticriminal modeling, and concrete problem solving in treatment. She found that 73% (8/11) of the studies that examined structured programs reported a significant impact upon recidivism, as opposed to only 17% (2/12) of the studies that explored less structured programming. Similarly, Ross and Fabiano (1985) found that 94% (15/16) of the studies that explored structured approaches to cognitive change reported positive effects relative to 29% (10/34) of noncognitive programs. Like Gendreau and Ross (1987), we too think that recent, more sophisticated meta-analyses support our overall conclusions.[3]

The basic elements of effective service, therefore, appear to be highly general in their applicability, and, we think, their success depends primarily upon the risk and need principles having been simultaneously implemented (Andrews, 1979, 1980; Andrews & Kiessling, 1980). The effectiveness of certain styles and modes of service, however, appear to depend upon certain case characteristics. These specific responsivity considerations are presented below in the form of working clinical hypotheses that are subject to routine evaluation.

SPECIFIC RESPONSIVITY CONSIDERATIONS

From the earliest days of the "talking cure," Freud (1953) warned psychodynamic therapists that their highly verbal, evocative, relationship-

dependent and insight-oriented therapy was inappropriate for cases with poor verbal ability and/or with cases displaying narcissistic and/or psychotic disorders. He stressed that some degree of experienced discomfort and an ability to enter into an emotional relationship with the therapist were crucial to success. Freud added that without immediate social support for both treatment and personal change, the chances of successful treatment were miniscule. He went so far as to admit that once his therapeutic reputation was established, he accepted only cases personally and socially committed to service gains.

Interestingly, one of the earliest controlled tests of the responsivity principle in corrections (and human service generally) reflects just such psychoanalytic considerations. In the PICO project (Grant, 1965), inmates were judged either "amenable" or "nonamenable." Relative to nonamenables, amenable offenders were verbally skilled, mature, anxious, and motivated to participate in sessions with a caseworker. That is, classification was based upon the individual's apparent ability to respond to psychodynamic casework. Assigned to psychodynamic casework or to a routine casework control condition, treated amenables had lower recidivism rates than nontreated amenables (see Table 6). More important, there were no differences in the recidivism of "untreated" amenables and nonamenables. That is, amenability was not a risk factor but a responsivity factor.

In the Camp Elliott Study (Grant, 1965), residential "closed living groups" were constructed so that military inmates would be encouraged to work out here-and-now interpersonal problems that presumably were rooted in early family experiences and that contributed to antisocial conduct. The idea was that interaction with therapists and supervisors (supportive parental figures) would promote the resolution of underlying problems and hence reduce recidivism. In the evaluation, two variables were crucial. One was the client factor: High-maturity inmates were relatively perceptive, anxious, and reflective relative to lower maturity inmates. The other factor was treatment: Three teams of unit supervisors were differentiated according to their interpersonal skills and maturity. The units supervised by the more skilled supervisors were described as "therapeutic communities," while units with the least-skilled supervisors were described as "more traditionally military" and "authoritarian" in structure and process. As

TABLE 6: Examples of the Responsivity Principle

PICO: Mean Follow-Up Months Incarcerated (Grant, 1965)			
	Psychodynamic Casework		
	No	*Yes*	
Amenable	4.8	2.1	*
Nonamenable	4.8	5.5	ns
Camp Elliott: Estimated Success Rates (Grant, 1965)			
	Level of Structure		
	Low	*High*	
High Maturity	.72	.60	*
Low Maturity	.46	.60	*
Recidivism Rates of High-Risk Probationers (Andrews & Kiessling, 1980)			
	Supervision by Citizen Volunteers		
	No	*Yes*	
High Empathy	.80	.00	*
Low Empathy	.48	.42	ns
Mean Number of New Offenses (Leschied, in Reitsma-Street, 1984)			
	Level of Structure		
	Low	*High*	
High Conceptual Level	nr	nr	nr
Low Conceptual Level	1.54	.47	*
Mean Estimated Residualized Depression Score (Simons et al., 1985)			
	Type of Treatment		
	Drug	*Cognitive*	
High Self-Control	+3.8	−3.0	*
Low Self-Control	−3.2	+6.0	*
Positive Behavior Change Rate (Sarason & Ganzer, 1973)			
	Stress Level (TV Feedback)		
	Low	*High*	
High Anxiety	.74	.07	*
Low Anxiety	nr	nr	ns
Mean Procriminal Change Scores (Andrews, 1980)			
	Level of Interpersonal Interaction		
	Lower	*Higher*	
High Anxiety	−2.4	1.0	*
Low Anxiety	2.9	−4.7	*

(continued)

TABLE 6 Continued

Prerevocation Warrant Rates of Parolees (Lerner et al., 1986)			
	CMS		
	No	*Yes*	
High Risk	.23 (296)	.15 (235)	*
Moderate Risk	.17 (740)	.11 (608)	*
Low Risk	.07 (339)	.06 (333)	ns

nr: not reported
ns: not significant
$p < .05$

predicted, among high-maturity inmates, interaction with unit supervisors who were ranked highest on interpersonal skills and maturity tended to produce positive attitude change and success upon discharge. However, the success of low-maturity inmates was clearly greatest among those who interacted with the least-skilled supervisors (see Table 6). Among other interpretations offered, Grant (1965) suggested that appeals for self-reflection and interpersonal interaction without a set of clear rules actively promoted acting-out on the part of low-maturity inmates.

Alternatively, in our opinion, it was only under high-structure conditions that low-maturity inmates had the opportunity to hear and learn anticriminal messages. PICO and Camp Elliott anticipated the failure of other interpersonally focused and relationship-oriented programs with low-empathy, immature, and frankly antisocial cases (Andrews & Kiessling, 1980, see Table 6; Craft et al., 1966; Goodman, 1972; Klein, 1971; Murphy, 1972). Similarly, many findings based on the Conceptual Level System (Reitsma-Street & Leschied, 1988) and the I-Level System (Harris, 1988; Jesness, 1988) suggest that delinquents functioning at low conceptual/maturity levels respond best to structured programs.

It should be noted that reports on I-Level results are still considered by some reviewers to be too "selective" (Sechrest, 1987), and the Conceptual Level findings in the correctional area may be examples of risk rather than responsivity. For example, it is now well-established that low conceptual level is a risk factor (Reitsma-Street & Leschied,

1988), and "low" as opposed to "high" structure may well be the equivalent of "less" versus "more" exposure to relevant service. The work conducted in educational settings, however, is strong (Reitsma-Street, 1984), and it is likely that Conceptual Level is both a risk factor and a responsivity factor. Similarly, possible problems with the selectivity of I-Level research findings aside, the I-Level recommendations regarding differential treatment have high face validity and considerable clinical appeal (Harris, 1988).

Cognitive Interpersonal Skill Hypothesis

This working clinical hypothesis is relatively strong but certainly requires additional study across a variety of measures of empathy/cognitive maturity (Van Voorhis, 1988) and program structure (Reitsma-Street, 1984). In brief, styles and modes of service that are interpersonally and verbally demanding and that depend upon self-regulation, self-reflection, and interpersonal sensitivity (that is, low-structure styles of service) should be actively avoided with all but those offenders who present relatively high levels of interpersonal and conceptual functioning. On the whole, if one is uncertain about the interpersonal/conceptual level of particular cases, the more directive and structured style of service should be chosen because the evidence regarding the relative effectiveness of "high structure" for low-functioning cases is stronger than the evidence regarding the importance of "low structure" for higher functioning cases. Interestingly, recent research on the treatment of depressives by Simons, Lustman, Wetzel, and Murphy (1985) has suggested the general importance of matching according to level of cognitive skills. For example, patients who entered therapy with relatively high levels of self-management skills responded positively to cognitive therapy, while the other patients responded best to drug therapy (see Table 6).

Anxiety Hypothesis

The "anxiety" component of Freud's ideas on differential treatment will emerge again under the motivation hypothesis. Here, the anxiety

hypothesis refers to the interaction of anxiety (in particular, interpersonal anxiety) and the level of interpersonal and intrapersonal confrontation involved in treatment. In particular, the clinical hypothesis is that highly anxious cases respond poorly to stressful interpersonal confrontation, while less anxious cases may well profit from such programs. Sarason and Ganzer (1973) found that highly anxious delinquents responded very well to modeling and yet responded very poorly when the stressor of televised feedback was introduced into social skills training (see Table 6). Similarly, anxious inmates have been found to deteriorate in group programs that involved very intensive interpersonal exchange (Andrews, 1980, see Table 6). Recognizing that many competing correlates of success exist (including interpersonal and cognitive maturity as well as risk levels), the effectiveness of guided group interaction programs in several replications also appears to have been limited to those who were able to handle intense interpersonal exchanges (Stephenson & Scarpitti, 1974).

Antisocial Personality Hypothesis

There is now no question regarding the predictive criterion validity of various approaches to the assessment of antisocial personality (Andrews & Wormith, in press). Equally so, antisocial personality types are presumed to be relatively unresponsive to rehabilitative efforts. Our working hypothesis is that the effectiveness of the treatment of "psychopaths" is a combined function of their high risk levels (intensive controls and service are indicated), multiple criminogenic needs (impulsivity, procriminal sentiments, and isolation from anticriminal others should be targeted), and scores on responsivity factors (low empathy, low anxiety) that indicate high levels of structure.

Sensation-Seeking Hypothesis

A general restlessness and a propensity for risky activities and excitement-seeking behavior are reasonably well-established correlates of delinquency (Glueck & Glueck, 1950; Hagan, Gillis, & Simpson, 1985). Thus styles of service that include novel and exciting opportunities and events are indicated.

Motivation Hypothesis

It is generally accepted that narcissistic and antisocial cases are not highly motivated to participate in treatment. Not only is experienced distress rather low, but the "symptoms" and the "acting out" are often highly rewarding. Thus, with high-risk but weakly motivated cases, it is particularly important that treatment is readily accessible and of the outreach type. With these cases the total environmental surround should be supportive of participation in programming and, perhaps, mandated by the court. There is some evidence that legal contingencies embedded within a therapeutic context may be helpful (Gendreau & Ross, 1987).

Social Support Hypothesis

The interpersonal environment of many offenders is not likely to be highly supportive of anticriminal change. For example, most studies of the circumstances of offenders have found that the presence of antisocial associates is a major correlate of delinquency (Glueck & Glueck, 1950). Thus effective programs will either isolate cases from those environments (Klein, 1971) or actively neutralize the procriminal pressures of criminal associates through the structure of the program (Andrews, 1980; Stephenson & Scarpitti, 1974).

Case Management Strategies

The full classification system of the Wisconsin Risk Assessment Scale includes consideration of risk, need, and what are called case management strategies. Lerner, Arling, and Baird (1986) have described a study by the Texas Board of Pardons in which parolees classified on the Wisconsin system were assigned to one of two parole supervision conditions: regular supervision or Client Management Classification (CMC). CMC attempts to match the level and type of intervention to risk level. Six-month follow-ups showed CMC intervention to be the most effective condition in reducing recidivism among the medium- and high-risk offenders, with no impact upon the low-risk group (see Table 6).

Gender/Race/Ethnicity Hypotheses

Being male and nonwhite are considered to be risk factors in North America. However, gender, ethnicity, and age may also be important responsivity factors (e.g., Carver & Owen, 1984). Thus the future will see examinations of the effectiveness of rehabilitation programs specifically matched to age, gender, and race/ethnicity.

Mentally Disordered Offenders

The chronic offender with histories of both psychiatric and criminal justice involvement may respond best to a monitored but low-stress sheltered workshop environment (Fairweather, Sanders, Cressler, & Maynard, 1969).

A number of clinical hypotheses have been listed as specific responsivity considerations. However, we continue to feel that risk and need considerations are crucial and that the major aspect of responsivity is that of choosing treatment approaches found to be effective with offender samples in general. In other words, specific responsivity factors are of relatively minor significance when the more general guidelines regarding effective styles of treatment are followed; use of authority, anticriminal modeling, and problem-solving efforts that focus on criminogenic need are generally effective when offered in an interpersonally facilitative and enthusiastic manner.

THE OVERRIDE PRINCIPLE

Final decisions regarding rehabilitative service are a joint function of risk, need, and responsivity considerations. The power and specificity of these considerations will increase with the quality of research and theory in psychology, just as research on risk, need, and responsivity will enhance the overall quality of psychological knowledge. However, rehabilitation professionals will always be called upon to step beyond extant knowledge in their decision making. Systematic monitoring and follow-up of the consequences of these overrides will surely be one source of new insights.

CONCLUSIONS

Risk, need, and responsivity considerations in the psychology of criminal conduct may better reflect current knowledge and opinion regarding discretionary services for purposes of rehabilitation than does the discounting of knowledge so characteristic of major portions of mainstream criminology. Risk, need, and responsivity considerations provide reasonable guides to service and research in rehabilitation. Of course, we agree with Sechrest (1987) that better theory and more high-quality research is required. More to the point, work is required that builds on the base of extant knowledge and professional opinion within the psychology of criminal conduct, and that base is not something to lament.

Lamentable are perspectives that confuse rehabilitation with punishment, humanitarian reform, and "nice" or "tyrannical" behavior. Also lamentable is that rehabilitation is not viewed as a professional area of practice, replete with a growing body of core psychological knowledge and opinion with which practitioners and managers should be familiar before "innovative" programs are introduced (witness the failure of so many diversion projects, as described by Gendreau and Ross [1987]). Unfortunately, we do not know much about the dissemination, adoption, and maintenance of effective programs (Backer, Liberman, & Kuehnel, 1986; Gendreau & Andrews, 1979; Gendreau & Ross, 1987). Recognizing that the action resides in consideration of preservice characteristics of offenders, the specifics of service planned and delivered, and intermediate change in the person and circumstances of offenders, we now need to work on creating broad setting and program conditions that support the efforts of rehabilitation professionals (Andrews, 1980; Kiessling & Andrews, 1980). The creation of these settings may enhance not only the delivery of service but also the effectiveness of services that have so often been offered under less than supportive conditions.

NOTES

1. Many criminologists seem to *know* that the causes of crime are buried deep in political economy, culture, and social structure, just as they *know* that intervention is mere tinkering. Thus

reviewers, who uniformly found that at least 40% of the evaluation studies uncovered positive effects, reach the following types of conclusions regarding the effectiveness of rehabilitation: "The results are far from encouraging" (Lab & Whitehead, 1988). "Lamentably, . . . we do not know very much" (Sechrest, 1987, p. 317). "There is not now . . . any basis for any policy or recommendations regarding rehabilitation" (Sechrest, White, & Brown, 1979, p. 34). "Nothing [or almost nothing] works" (Martinson, 1974). "Much of what is now being done about crime may be so wrong that the net effect . . . is to increase rather than decrease crime" (Logan, 1972, p. 381). "There has been no apparent progress in the actual demonstration of the validity of various type of correctional treatment" (Bailey, 1966, p. 157).

2. In brief, (1) reassessments of need should be shown to be responsive to theoretically relevant intervention; that is, change scores are found to link with service, and the magnitude of the service-change correlation we call a "dynamic dependent validity" estimate; (2) reassessment scores (or change scores) should be shown to be incrementally predictive of recidivism relative to intake scores; this we call "dynamic predictive criterion validity"; and (3) controls for change scores should be shown to reduce the magnitude of the service-recidivism link; this we now call evidence of "functional validity". These validities of change scores are crucial, but need assessments are also, of course, subject to judgments regarding reliability, content, and construct validity as well as user convenience and utility.

3. Subsequent to writing this article, we conducted our own meta-analysis of the correctional treatment literature (Andrews et al., 1989). Applications of the principles of risk, need, and responsivity revealed that appropriate correctional treatment was significantly more effective than criminal sanctioning without the provision of rehabilitative service and more effective than service inconsistent with the three principles.

REFERENCES

Akers, R. L. (1973). *Deviant behavior: A social learning approach*. Belmont, CA: Wadsworth.

Andrews, D. A. (1974, February). *Outcome evaluations of group counselling in corrections*. Paper presented at the Symposium on Corrections, Ontario Psychological Association Meetings, Ottawa, Canada.

Andrews, D. A. (1979). *The dimensions of correctional counselling and supervision process in probation and parole*. Toronto: Ontario Ministry of Correctional Services.

Andrews, D. A. (1980). Some experimental investigations of the principles of differential association through deliberate manipulations of the structure of service systems. *American Sociological Review, 45*, 448-462.

Andrews, D. A. (1982a). *The supervision of offenders: Identifying and gaining control over the factors that make a difference*. Program Branch User Report. Ottawa: Solicitor General of Canada.

Andrews, D. A. (1982b). *The Level of Supervision Inventory (LSI): The first follow-up*. Toronto: Ontario Ministry of Correctional Services.

Andrews, D. A. (1983). Assessment of outcome in correctional samples. In M. J. Lambert, E. R. Christensen, & S. S. DeJulio (Eds.), *Assessment of psychotherapy outcome*. New York: Wiley.

Andrews, D. A., & Friesen, W. (1987). Assessments of anticriminal plans and the prediction of criminal futures: A research note. *Criminal Justice and Behavior, 14*, 33-37.

Andrews, D. A., Hoge, R. D., Robinson, D., & Hollett, J. (1986). *The family service assessment and evaluation project: Final report, phase 1*. Ottawa: Family Service Canada.

Andrews, D. A., & Kiessling, J. J. (1980). Program structure and effective correctional practices: A summary of the CaVIC research. In R. R. Ross & P. Gendreau (Eds.), *Effective correctional treatment*. Toronto: Butterworths.

Andrews, D. A., Kiessling, J. J., Mickus, S., & Robinson, D. (1986). The construct validity of interview-based risk assessment in corrections. *Canadian Journal of Behavioural Science, 18*, 460-470.

Andrews, D. A., Kiessling, J. J., Robinson, D., & Mickus, S. (1986). The risk principle of case classification: An outcome evaluation with young adult probationers. *Canadian Journal of Criminology, 28*, 377-396.

Andrews, D. A., & Robinson, D. (1984). *The Level of Supervision Inventory: Second report*. A report to Research Sevices (Toronto) of the Ontario Ministry of Correctional Services.

Andrews, D. A., Robinson, D., & Balla, M. (1986). Risk principle of case classification and the prevention of residential placements: An outcome evaluation of Share the Parenting. *Journal of Consulting and Clinical Psychology, 54*, 203-207.

Andrews, D. A., & Wormith, J. S. (1984). *Criminal sentiments and criminal behaviour*. Programs Branch User Report. Ottawa: Solicitor General of Canada.

Andrews, D. A. & Wormith, J. S. (in press). Personality and crime: Knowledge destruction and construction in criminology. *Justice Quarterly*.

Andrews, D. A., Wormith, J. S., & Kiessling, J. J. (1985). *Self-reported criminal propensity and criminal behavior: Threats to the validity of assessment and personality*. Programs Branch User Report. Ottawa: Solicitor General of Canada.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1989). *Does correctional treatment work? A clinically-relevant and psychologically-informed meta-analysis*. Paper presented at the Research Seminar of National Associations Active in Criminal Justice, Ottawa, Canada.

Backer, T. E., Liberman, R. P., & Kuehnel, T. G. (1986). Dissemination and adoption of innovative psychosocial interventions. *Journal of Consulting and Clinical Psychology, 54*, 111-118.

Bailey, W. C. (1966). Correctional outcome: An evaluation of 100 reports. *Journal of Criminal Law, Criminology and Police Science, 57*, 153-160.

Baird, S. C., Heinz, R. C., & Bemus, B. J. (1979). *Project report 14: A two year follow-up*. Wisconsin: Department of Health and Social Services, Case Classification/Staff Deployment Project, Bureau of Community Corrections.

Barkwell, L. (1980). Differential probation treatment of delinquents. In R. R. Ross & P. Gendreau (Eds.), *Effective correctional treatment*. Toronto: Butterworths.

Barlow, H. D. (1986). *Introduction to criminology*. Boston: Little, Brown.

Beutler, L. E. (1979). Toward specific psychological therapies for specific conditions. *Journal of Consulting and Clinical Psychology, 47*, 882-897.

Bohn, M. J. (1980). Inmate classification and the reduction of institutional violence. *Corrections Today, 42*, 48-49.

Bonta, J. (in press). Native inmates: Institutional response, risk, and need. *Canadian Journal of Criminology*.

Bonta, J., & Motiuk, L. L. (1985). Utilization of an interview-based classification instrument: A study of correctional halfway houses. *Criminal Justice and Behavior, 12*, 333-352.

Bonta, J., & Motiuk, L. L. (1987). The diversion of incarcerated offenders to correctional halfway houses. *Journal of Research in Crime and Delinquency, 24*, 302-323.

Briggs, P. F., & Wirt, R. D. (1965). Prediction. In H. C. Quay (Ed.), *Juvenile delinquency: Research and theory*. Princeton, NJ: Van Nostrand.

Brill, R. (1978). Implications of the conceptual level matching model for treatment of delinquents. *Journal of Research in Crime and Delinquency, 15*, 229-246.

Burgess, E. W. (1928). Factors determining success or failure on parole. In A. A. Bruce, A. J. Harno, E. W. Burgess, & J. Landesco (Eds.), *The workings of the indeterminate-sentence law and the parole system in Illinois.* Springfield, IL: State Board of Parole.

Byles, J. A., & Maurice, A. (1982). The juvenile services project. *Canadian Journal of Criminology, 24*, 155-165.

Canada. (1987). *Sentencing reform: A Canadian approach (Report of the Canadian Sentencing Commission)*. Ottawa: Supply and Services Canada.

Carver, V., & Owen, T. (1984). *Amethyst Women's Addiction Centre: Final report.* A report submitted to Health Services and Promotion Branch (Toronto), Health and Welfare Canada.

Clements, C. B. (1986). *Offender needs assessment.* College Park, MD: American Correctional Association.

Cohen, S. (1985). *Visions of social control.* Cambridge, UK: Polity Press.

Craft, M., Stephenson, G., & Granger, C. (1966). A control trial of authoritarian and self-governing regimes with adolescent psychopaths. *American Journal of Orthopsychiatry, 34*, 543-554.

Cullen, F. T., & Gendreau, P. (1988). The effectiveness of correctional rehabilitation. In L. Goodstein & D. L. MacKenzie (Eds.), *The American prison: Issues in research policy*. New York: Plenum.

Duffee, D. E., & Duffee, B. W. (1981). Studying the needs of offenders in prerelease centers. *Journal of Research in Crime and Delinquency, 18*, 232-253.

Edinger, J. D. (1979). Cross-validation of the Megargee MMPI typology for prisoners. *Journal of Consulting and Clinical Psychology, 47*, 234-242.

Edinger, J. D., & Auerbach, S. M. (1978). Development and validation of a multidimensional multivariate model for accounting for infractions in a correctional setting. *Journal of Personality and Social Psychology, 38*, 1472-1489.

Fairweather, G. W., Sanders, D. H., Cressler, D. L., & Maynard, H. (1969). *Community life for the mentally ill: An alternative to institutional care.* Chicago: Aldine.

Farrington, D. P., & Tarling, R. (1985). *Prediction in criminology.* Albany: State University of New York Press.

Freud, S. (1953). *A general introduction to psychoanalysis.* New York: Permabooks.

Gendreau, P., & Andrews, D. A. (1979). Psychological consultation in correctional agencies: Case studies and general issues. In J. J. Platt & J. Wicks (Eds.), *The psychological consultant.* New York: Grune & Stratton.

Gendreau, P., & Ross, R. R. (1979). Effective correctional treatment: Bibliotherapy for cynics. *Crime & Delinquency, 25*, 463-489.

Gendreau, P., & Ross, R. R. (1987). Revivication of rehabilitation: Evidence from the 1980s. *Justice Quarterly, 4*, 349-408.

Glaser, D. (1974). Remedies for the key deficiency in criminal justice evaluation research. *Journal of Research in Crime and Delinquency, 10*, 144-154.

Glaser, D. (1987). Classification for risk. In D. M. Gottfredson & M. H. Tonry (Eds.), *Prediction and classification: Criminal justice decision making.* Chicago: University of Chicago Press.

Glueck, S., & Glueck, E. T. (1930). *500 criminal careers.* New York: Knopf.

Glueck, S., & Glueck, E. T. (1950). *Unravelling juvenile delinquency.* Cambridge, MA: Harvard University Press.

Gottfredson, D. M., & Tonry, M. (1987). *Prediction and classification: Criminal justice decision making*. Chicago: University of Chicago Press.

Gottfredson, M. R. (1979). Treatment destruction techniques. *Journal of Research in Crime and Deliquency, 16*, 39-54.

Goodman, G. (1972). *Companionship therapy*. San Francisco: Jossey-Bass.

Grant, J. D. (1965). Delinquency treatment in an institutional setting. In H. C. Quay (Ed.), *Juvenile delinquency: Research and theory*. Princeton, NJ: Van Nostrand.

Hagan, J., Gills, A. R., & Simpson, J. (1985). The class structure of gender and delinquency: Toward a power-control theory of common delinquent behavior. *American Journal of Sociology, 90*, 1151-1177.

Hanson, R. W., Moss, C. S., Hosford, R. E., & Johnson, M. E. (1983). Predicting inmate penitentiary adjustment: An assessment of four classificatory methods. *Criminal Justice and Behavior, 10*, 293-309.

Harris, P. W. (1988). The interpersonal maturity level classification system: I-level. *Criminal Justice and Behavior, 15*, 58-77.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.

Hirschi, T., & Hindelang, M. J. (1977). Intelligence and delinquency: A revisionist review. *American Sociological Review, 42*, 571-587.

Hoge, R. D., & Andrews, D. A. (1986). A model for conceptualizing interventions in social service agencies. *Canadian Psychology, 27*, 332-341.

Hollett, J. (1984). *Effective elements of correctional programs*. Unpublished comprehensive examination, Department of Psychology, Carleton University, Ottawa, Canada.

Jesness, C. F. (1975). Comparative effectiveness of behavior modification and transactional analysis programs for delinquents. *Journal of Consulting and Clinical Psychology, 43*, 758-779.

Jesness, C. F. (1988). The Jesness inventory classification system. *Criminal Justice and Behavior, 15*, 78-91.

Johnson, R. E. (1979). *Juvenile delinquency and its origins: An integrative theoretical approach*. Cambridge: Cambridge University Press.

Kazdin, A. E. (1985). Selection of target behaviors: The relationship of the treatment focus to clinical dysfunction. *Behavioral Assessment, 7*, 33-47.

Kiessling, J. J., & Andrews, D. A. (1980). Behaviour analysis systems in corrections: A new approach to the synthesis of correctional theory, practice, management and research. *Canadian Journal of Criminology, 22*, 412-427.

Kirby, B. C. (1954). Measuring effects of treatment of criminals and delinquents. *Sociology and Social Research, 38*, 368-374.

Kirchner, R. E., Schnelle, J. F., Domash, M., Larson, L., Carr, A., & McNees, M. P. (1980). The applicability of a helicopter patrol procedure to diverse areas: A cost-benefit analysis. *Journal of Applied Behavior Analysis, 13*, 143-148.

Kirk, S. A. (1976). Effectiveness of community services for discharged mental hospital patients. *American Journal of Orthopsychiatry, 46*, 646-659.

Klein, M. W. (1971). *Street gangs and street workers*. Englewood Cliffs, NJ: Prentice-Hall.

Lab, S. P., & Whitehead, J. T. (1988). An analysis of juvenile correctional treatment. *Crime & Delinquency, 34*, 60-85.

Lerner, K., Arling, G., & Baird, S. C. (1986). Client management classification strategies for case supervision. *Crime & Delinquency, 32*, 254-271.

Leschied, A. W., Jaffe, P. G., & Stone, G. L. (1985). Differential response of juvenile offenders to two different environments as a function of conceptual level. *Canadian Journal of Criminology, 27*, 467-476.

Leschied, A. W., & Thomas, K. E. (1985). Effective residential programming for "hard-to-serve" delinquent youth: A description of the Craigwood program. *Canadian Journal of Criminology, 27*, 161-177.

Levinson, R. B. (1988). Developments in the classification process: Quay's AIMS approach. *Criminal Justice and Behavior, 15*, 24-38.

Loeber, R. (1982). The stability of antisocial and delinquent child behavior: A review. *Child Development, 53*, 1431-1446.

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

Logan, C. H. (1972). Evaluation research in crime and delinquency: A reappraisal. *Journal of Criminal Law, Criminology and Police Science, 63*, 378-387.

Louscher, P. K., Hosford, R. E., & Moss, C. S. (1983). Predicting dangerous behavior in a penitentiary using the Megargee typology. *Criminal Justice and Behavior, 10*, 269-284.

Lukin, P. R. (1981). Recidivism and changes made by delinquents during residential treatment. *Journal of Research in Crime and Delinquency, 18*, 101-112.

Maclean, B. D. (1986). *The political economy of crime*. Scarborough: Prentice-Hall.

Martinson, R. (1974). What works?—Questions and answers about prison reform. *Public Interest, 35*, 22-54.

Martinson, R. (1976). California research at the crossroads. *Crime & Delinquency, 22*, 178-191.

Mash, E. J. (1985). Some comments on target selection. *Behavioral Assessment, 7*, 63-78.

Matsueda, R. L. (1988). The current state of differential association theory. *Crime & Delinquency, 34*, 277-306.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.

Megargee, E. I., & Bohn, M. J. (1979). *Classifying criminal offenders: A new system based on the MMPI*. Beverly Hills, CA: Sage.

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

Motiuk, L. L., Bonta, J., & Andrews, D. A. (1986). Classification in correctional halfway houses: the relative and incremental predictive validities of the Megargee-MMPI and LSI systems. *Criminal Justice and Behavior, 13*, 33-46.

Murphy, B. C. (1972). *A quantitative test of the effectiveness of an experimental treatment program for delinquent opiate addicts*. Ottawa: Information Canada.

Nelson, R. O., & Hayes, S. C. (1979). Some current dimensions of behavioral assessment. *Behavioral Assessment, 1*, 1-16.

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

O'Donnell, C. R., Lydgate, T., & Fo, W.S.O. (1971). The buddy system: Review and follow-up. *Child Behavior Therapy, 1*, 161-169.

Ogloff, J.R.P. (1988, June). *Law and psychology in Canada: The need for training and research*. Paper presented at the annual meeting of the Canadian Psychological Association, Montreal, Canada.

Palmer, T. (1974). The Youth Authority's community treatment project. *Federal Probation*, (March), 3-14.

Posey, C. D. (1988). Introduction. *Criminal Justice and Behavior, 15*, 5-7.

Quay, H. C. (1984). *Managing adult inmates: Classification for housing and program assignments*. College Park, MD: American Correctional Association.

Quay, H. C. & Love, C. T. (1977). The effects of a juvenile diversion program on rearrest. *Criminal Justice and Behavior, 4*, 377-396.

Rappaport, J. (1979). *Community psychology: Values, research and action.* New York: Holt, Rinehart & Winston.

Reitsma-Street, M. (1984). Differential treatment of young offenders: A review of the conceptual level matching model. *Canadian Journal of Criminology, 26,* 199-212.

Reitsma-Street, M., & Leschied, A. W. (1988). The conceptual level matching model in corrections. *Criminal Justice and Behavior, 15,* 92-108.

Ross, R. R., & Fabiano, E. A. (1985). *Time to think: A cognitive model of delinquency prevention and offender rehabilitation.* Johnson City, TN: Institute of Social Sciences and Arts.

Sarason, I. G., & Ganzer, V. J. (1973). Modeling and group discussions in the rehabilitation of juvenile delinquents. *Journal of Counseling Psychology, 20,* 442-449.

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

Scott, D. (1985). *The Youth Level of Service Inventory in probation.* Unpublished honours thesis, Carleton University.

Sechrest, L. (1987). Classification for treatment. In D. M. Gottfredson & M. F. Tonry (Eds.), *Prediction and classification: Criminal justice decision making.* Chicago: University of Chicago Press.

Sechrest, L., White, S. O., & Brown, E. D. (1979). *The rehabilitation of criminal offenders: Problems and prospects.* Washington, DC: National Academy of Sciences.

Simons, A. D., Lustman, P. J., Wetzel, R. D., & Murphy, G. E. (1985). Predicting response to cognitive therapy of depression: The role of learned resourcefulness. *Cognitive Therapy and Research, 9,* 79-89.

Stephenson, R. M., & Scarpitti, F. R. (1974). *Group interaction as therapy: The use of the small group in corrections.* Westport, CT: Green Wood Press.

Taylor, I., Walton, P., & Young, J. (1973). *The new criminology: For a social theory of deviance.* London: Routledge & Kegan Paul.

Tittle, C. R., Villimez, W. J., & Smith, D. A. (1978). The myth of social class and criminality: An empirical assessment of the empirical evidence. *American Sociological Review, 43,* 643-656.

Van Voorhis, P. (1988). A cross-classification of five offender typologies: Issues of construct and predictive validity. *Criminal Justice and Behavior, 15,* 109-124.

Vold, G. B., & Bernard, T. J. (1986). *Theoretical criminology* (3rd ed.). New York: Oxford University Press.

Voss, R. (1982). *Changes in alcohol and drug use as predictors of recidivism.* Unpublished honors thesis, Department of Psychology, Carleton University, Ottowa, Canada.

Walters, G. D., Scrapansky, T. A., & Marrlow, G. A. (1986). The emotionally disturbed military criminal offender: Identification, background, and institutional adjustment. *Criminal Justice and Behavior, 13,* 261-285.

Warren, M. Q. (1969). The case for differential treatment of delinquents. *Annals of the American Academy of Political and Social Science, 381,* 47-59.

Woody, G. E., McLellan, A. T., Luborsky, L., O'Brien, C. P., Blaine, J., Fox, S., Herman, I., & Beck, A. T. (1984). Severity of psychiatric symptoms as a predictor of benefits from psychotherapy: The Veteran's Administration-Penn Study. *American Journal of Psychiatry, 141,* 1172-1177.

Wormith, J. S. (1984). Attitude and behavior change of correctional clientele: A three year follow-up. *Criminology, 22,* 595-618.

Wright, K. N. (1986). An exploratory study of transactional classification. *Journal of Research in Crime and Delinquency, 23,* 326-348.

Wright, K. N., Clear, T. R., & Dickson, P. (1984). Universal applicability of probation risk-assessment instruments. *Criminology, 22*, 113-134.

Zamble, E., & Porporino, F. J. (1988). *Coping, behavior, and adaptation in prison inmates.* New York: Springer-Verlag.

A META-ANALYSIS OF THE PREDICTORS OF ADULT OFFENDER RECIDIVISM: WHAT WORKS!*

PAUL GENDREAU
TRACY LITTLE
CLAIRE GOGGIN
University of New Brunswick

Meta-analytic techniques were used to determine which predictor domains and actuarial assessment instruments were the best predictors of adult offender recidivism. One hundred and thirty-one studies produced 1,141 correlations with recidivism. The strongest predictor domains were criminogenic needs, criminal history/history of antisocial behavior, social achievement, age/gender/race, and family factors. Less robust predictors included intellectual functioning, personal distress factors, and socioeconomic status in the family of origin. Dynamic predictor domains performed at least as well as the static domains. The LSI-R was identified as the most useful actuarial measure. Recommendations for developing sound assessment practices in corrections are provided.

Verification of the risk factors most predictive of adult offender recidivism and identification of the actuarial instruments best suited to that end have major implications for corrections policymakers, practitioners, and program evaluators. The cost-effective and humane management of prisons, particularly in light of the dramatic increase in incarceration rates (Mauer, 1994), dictates that maximum security prisons be reserved for the highest risk offenders. Moreover, the design of effective offender treatment programs is highly dependent on knowledge of the predictors of recidivism (Gendreau et al., 1994).

Andrews and Bonta (1994) identify two categories of risk factors: static and dynamic. Static factors (i.e., age, previous convictions) are aspects of the offender's past that are predictive of recidivism but cannot be changed. Dynamic risk factors, or what Andrews and Bonta commonly refer to as criminogenic needs (e.g., antisocial cognitions, values, and behaviors), are mutable and thus serve as the appropriate targets for treatment (Andrews

* The research was funded by contracts No. 1514-UN/4200 and No. 9314-UN/525, Corrections Branch, Ministry Secretariat, Solicitor General of Canada to the first author. We are indebted to Drs. Eric Marchand and Ken McGraw for their advice on meta-analytic statistics and to Drs. Don Andrews and James Bonta for their comments on earlier drafts of the manuscript.

et al., 1990a). There is, however, little consensus regarding the measurement of specific offender risk factors.

PREDICTORS OF RECIDIVISM

There is no disagreement in the criminological literature about some of the predictors of adult offender recidivism, such as age, gender, past criminal history, early family factors, and criminal associates. There has been, however, considerable controversy and/or lack of interest in dynamic risk factors. There are three reasons for this. First, because of ideological concerns and the professional self-interest of significant segments of the professions of criminology and sociology, the import of individual differences (i.e., offender needs, abilities, attitudes, and personality styles) has been derided in some criminological literature (Andrews and Wormith, 1989; Rowe and Osgood, 1984; Wilson and Herrnstein, 1985).

Second, some methodologists (e.g., Jones, 1996) have expressed skepticism about dynamic risk factors because of their supposed unreliability. Unlike their static counterparts, dynamic risk factors can change over time and their measurement involves some degree of subjectivity. Since elementary psychometric theory reminds one that unreliability in measurement necessarily leads to an underestimation of validity (Cronbach, 1990), this line of reasoning implies that, collectively, dynamic variables must be relatively weak predictors of criminal behavior.

Third, criminal justice professionals have been, by and large, antipathetic to the possibility that assessment of criminogenic needs might enhance the prediction of criminal behavior (Bonta, 1996; Gendreau and Ross, 1987). The widely used Wisconsin classification system (Baird, 1981) illustrates this point. This instrument contains a useful needs component, but Bonta (1996) found just two studies that reported on the predictive validity of those items. Further, the emergence of the "new penology" (Feeley and Simon, 1992), which is concerned with managing large aggregates of offenders in a simplistic input-output, businesslike fashion, has further contributed to the lack of interest in dynamic variables.

This denial of the utility of dynamic risk factors has serious ramifications for corrections professionals who are routinely required to reclassify offenders for prison transfers, parole/probation supervision, and treatment services. Simply put, reclassification is devalued if the measurement of change has little validity.

Three specific types of predictors have also been the subject of much debate. They are social class of origin, intelligence, and personal distress. Social class of origin (i.e., parents' occupation, education), has been the bedrock variable used in support of sociological theories of crime that assert that criminal behavior is determined largely by one's social location

(Andrews and Bonta, 1994). Tittle and Meier (1990, 1991) have challenged this view, showing social class of origin (socioeconomic status, or SES) to be a very weak predictor of juvenile delinquency.

The view that criminals are less intelligent than nonoffenders has been prevalent for decades (Goddard, 1920). Over the years, a number of studies have demonstrated a correlation between intelligence and delinquency (Hirschi and Hindelang, 1977). Recently, with the publication of *The Bell Curve* (Herrnstein and Murray, 1994), arguably the strongest claim yet has been made that IQ is a particularly powerful predictor. Their conclusions have serious implications for the provision of treatment programs for offenders, since IQ, in their view, is considered to be largely immutable.

According to Andrews et al. (1990a) personal distress variables (e.g., low self-esteem, anxiety) are not risk factors and are, therefore, inappropriate targets for treatment. Their conclusions are in stark contrast to the practices of many therapists and programs that give priority to lowering offenders' anxiety level and raising their self-esteem. The genesis of this perspective is, most likely, a consequence of the training received in mental health theory and practice (e.g., psychodynamic theory, phenomenology), where treatment professionals initially gained experience before emigrating to corrections in the 1960s (Gendreau, 1996). The current widespread popularity of the recovery and self-help agendas (see Kaminer, 1992) lends further credibility to the notion that personal distress factors are suitable targets for intervention, a view which in our opinion, has been generalized to corrections, where surveys of treatment programs have found that it is not uncommon for programs to attempt to alleviate offenders' personal distress (Gendreau et al., 1990; Hoge et al., 1993).

To date, reviews of the evidence concerning the predictors of recidivism have been limited in scope and narrative in nature—except for two reviews that employed meta-analytic procedures. One meta-analysis, however, was quite preliminary (Gendreau et al., 1992), and the other was restricted to twin and adoption studies that combined juvenile and adult samples (Walters, 1992).

ACTUARIAL MEASURES FOR PREDICTING RECIDIVISM

Bonta (1996) has categorized risk assessment measures within a developmental framework. First-generation techniques are based on clinical intuition and professional judgment. There is a plethora of literature documenting the lack of validity of this approach (Meehl, 1954), even among the most highly trained clinicians and scholars (Little and Schneidman,

1959). This perspective is still commonplace among corrections professionals (Clear and Gallagher, 1985).

Second-generation assessments are actuarial in nature. They are based on standardized, objective risk prediction instruments, such as the Salient Factor Score (SFS) (Hoffman, 1983), that are based almost entirely on static criminal history items. These kinds of measures provide little direction for classification and treatment decisions because the fixed nature of the items does not allow for changes in the offender's behavior to be reflected on subsequent retesting.

Bonta's third generation consists of two types of instruments. One of them encompasses risk prediction measures that include dynamic factors (e.g., Community Risk/Needs Management scale, Motiuk, 1993; Level of Service Inventory (LSI-R), Andrews and Bonta, 1995; the Wisconsin system, Baird, 1981), which assess a wide range of criminogenic needs. The second type includes personality test scales in the antisocial personality/sociopathy/psychopathy content area. While these scales (e.g., the MMPI Pd scale, the Psychopathy Checklist (PCL-R), Hare, 1991; the Socialization scale (Soc) of the California Personality Inventory (CPI), Gough, 1957) do contain static items, the majority of items are dynamic in nature.

Reviews of the risk-measure literature have also been, with one exception (Simourd et al., 1991), narrative in nature. Their meta-analysis reported that the PCL-R and the Soc scale of the CPI were better predictors of recidivism than the MMPI Pd scale. Unfortunately, most of the studies available to the authors were postdictive.[1]

A final comment concerns the fact that the validity of various theories of criminal behavior relies, somewhat, on the prediction literature. Anomie/strain (Merton, 1957) and subcultural theories (A. Cohen, 1955; Matza, 1964) support SES and, to some extent, personal distress as strong predictors. Contemporary reformulations of differential association, social learning, and control theories (Andrews and Bonta, 1994; LeBlanc et al., 1988; Widom and Toch, 1993) center on antisocial peers, learned antisocial values, early criminogenic family factors, and personality dimensions (e.g., egocentricity). Strong biologically oriented theories base much of their credence on IQ and twin studies (see Herrnstein and Murray, 1994; Walters, 1992).

In summary, our review of the literature on predictors of recidivism for adult offenders has indicated a need for a comprehensive, quantitative

1. Brief mention should also be made of a few quantitative within-subject study prospective comparisons of several risk instruments and personality scales (Gendreau et al., 1979a, b; Gough et al., 1965; Motiuk et al., 1986; Motiuk, 1991; Serin et al., 1990). The results from these studies indicated that, in most instances, risk measures (SFS, LSI-R) were better predictors of offender recidivism than were antisocial personality scales such as the MMPI Pd.

research synthesis (i.e., meta-analysis) of the major classes of predictors of recidivism and the available prediction instruments. The potential advantages of meta-analysis over narrative reviews have been summarized in detail elsewhere (Cooper and Hedges, 1994). It has become the review method of choice in many applied areas (e.g., Lipsey and Wilson, 1993) and has recently led to advances in knowledge in the correctional field (Andrews et al., 1990b; Bonta and Gendreau, 1990; Gendreau and Andrews, 1990; Lipsey, 1992; Walters, 1992).

The questions we address in this study are as follows:

1. Which predictor domains predict recidivism, and are some more potent than others?
2. Are dynamic predictors as a group inferior to static predictors in their ability to predict recidivism?
3. Are there differences among composite measures of risk prediction instruments and measures of antisocial personality in their ability to predict recidivism?
4. Are the strongest predictors of recidivism associated with different theories of criminal behavior?
5. What guidelines are forthcoming from the meta-analysis that will assist criminal justice professionals in making more accurate assessments of criminal behavior?

METHOD AND PROCEDURE

SAMPLE OF STUDIES

A literature search for relevant studies published between January 1970 and June 1994 was conducted using the ancestry approach and library abstracting services. For a study to be included, the following criteria applied:

1. Data on the offender were collected prior to the recording of the criterion measures. A minimum follow-up period of six months was required. If a study reported more than one follow-up period, data from the longest interval were used.
2. Treatment studies that directly attempted to change offender personality or behavior were not included.
3. The criterion or outcome measure of recidivism had to be recorded when the offender was an adult (18 years or older).
4. The criterion or outcome measure had to have a no-recidivism category. Studies that used "more" versus "less" crime categorizations were not used. The criterion measures were arrest, conviction, incarceration, parole violation, or a combination thereof.
5. The study was also required to report statistical information that

could be converted, using meta-analytic formulas (Rosenthal, 1991), into the common metric or effect size of Pearson *r*.

CODING THE STUDIES

For each study the following information was recorded:

1. Coder characteristics: date, coder identity.
2. Study characteristics: published document, type of publication, funding source, multidisciplinary authorship, judgment of senior author's knowledge of the area, gender of authors, affiliation of authors, geographic location of study, decade in which study was published.
3. Study sample characteristics: age, gender, race, urban/rural, SES, risk level, crime history, psychological make-up.
4. Study methodology: extreme groups design, attrition, follow-up length, type of outcome measure, sample size, statistical value.

The accuracy of coding was assessed using the index: agreement = number of agreements ÷ (number of agreements + number of disagreements) (Yeaton and Wortman, 1993). The second author coded all studies. The first author blindly coded a random sample of 30 studies. Percentage agreement scores for the two raters ranged from 85% to 98% across coding categories. Where disagreements occurred, the coding used was based on the first author's classification.

PREDICTOR CATEGORIES

The predictors were initially sorted into 18 domains (Category I). The coding criteria are detailed in the appendix. Then, for the purposes of research synthesis, the 18 domains were collapsed into 8 all-encompassing predictor domains (Category II): (1) age/gender/race, (2) criminal history, (3) criminogenic needs, (4) family factors, (5) intellectual functioning, (6) personal distress, (7) SES, and (8) social achievement.

EFFECT SIZE CALCULATION

Pearson product-moment correlation coefficients were produced for all predictors in each study that reported a numerical relationship with the criterion. When statistics other than Pearson *r* were presented, they were converted to *r* using the appropriate statistical formulas (Rosenthal, 1991). Where a *p* value of greater than .05 was the only reported statistic, an *r* of .0 was assigned.

Next, the obtained correlations were transformed using Fisher's table. Then, according to the procedures outlined by Hedges and Olkin (1985:230–232), the statistic z^+, representing the weighted estimation of Pearson *r*, was calculated for each predictor domain by dividing the sum of

the weighted z_rs per predictor domain by dividing the sum of each predictor's sample size minus three across that domain.

In order to determine the practical utility of various predictors relative to each other, the common language (CL) effect size indicator (McGraw and Wong, 1992) was also employed. The CL measure is little affected by changes in base rates and selection ratios, which makes it ideal for prediction studies (Rice and Harris, 1995). The CL statistic converts an effect size into the probability that a predictor-criterion score sampled at random from the distribution of one predictor domain (e.g., criminogenic needs) will be greater than that sampled from another distribution (e.g., personal distress).

SIGNIFICANCE TESTING

To determine which of the predictor domains predicted criterion significantly different from zero, the mean z^+ values for each domain were multiplied by the value of $(N - 3k)^{1/2}$, where N = the number of subjects per predictor domain and k = the number of predictors per domain (Hedges and Olkin, 1985).

One-way ANOVAs and the Student–Newman–Keuls (SNK) multiple comparison test were then applied to the mean r values of those domains that significantly predicted criterion better than zero in order to assess which domains differed significantly from each other.

Mindful of the debate regarding alternatives to the use of parametric methods as tests of significance in meta-analyses, the mean z^+ values for significant predictor domains were also assessed using an analog to the ANOVA's F test, the goodness-of-fit statistic Q (Hedges and Olkin, 1985). Following that, post hoc comparisons of the differences between mean z^+ values of each pair of significant predictor domains were conducted using the z test (E. Marchand, personal communication, June 15, 1994).

Finally, one-way ANOVAs and the SNK test using Pearson r were employed to assess whether type of outcome criteria, length of follow-up, and study characteristics were related to effect size.

The CL statistic does not involve significance testing.

Unless otherwise specified, alpha was set at .05 two-tail for all significance tests.

RESULTS

We identified 131 studies as suitable for the meta-analysis. These studies generated 1,141 effect sizes with future criminal behavior.

For those variables for which at least 60% of the studies reported information on the study characteristics sampled, the results were as follows: (1) 86% of the studies were published, 58% in journals; (2) 73% of the

senior authors had published in the area previously, 51% of them were male; (3) 44% and 54% of authors were based in an academic or government agency setting, respectively; (4) the studies were evenly distributed across the decades and the majority emanated from the United States and Canada, although Canadian studies contributed the majority (63%) of effect sizes; (5) 95% of studies consisted of male or mixed samples; (6) only 5% of studies employed an extreme groups design; and (7) 83% did not suffer subject attrition of more than 10% of their sample.

PREDICTOR DOMAINS: CATEGORY I

Table 1 presents the mean effect sizes for the 18 levels of Category I in conjunction with the number of effect sizes (k) and the total number of subjects associated with each predictor domain (N). The domains are grouped as follows: static ($n = 10$), dynamic ($n = 7$), and composite measures ($n = 1$).

The following is an example of how to read Table 1. Across the 131 studies sampled, a quantitative relationship between the predictor age and recidivism was reported on 56 occasions and involved a total of 61,312 subjects. The associated mean Pearson r for age with outcome was .15 (S.D. = .12), with younger age being positively correlated with poorer outcome. Mean z^+, the weighted estimation of Pearson r for age with outcome, was .11. Application of Hedges and Olkin's (1985) method for testing the significance of the mean z^+ values confirmed age as a significant predictor of recidivism.

All predictor domains were significant predictors of recidivism. The largest mean r values were found for adult criminal history, antisocial personality, companions, and criminogenic needs. Risk scale measures, which contained information from several predictor domains, produced the highest mean r value with recidivism (.30).

The conclusions reached by the parametric (ANOVA, SNK) statistical analysis were virtually identical to those of the F-test analog (Q, Z-test comparison). We report the results of the standard parametric analysis.

A one-way ANOVA applied to the mean r values (excluding composite risk scales) indicated there was a significant difference across the predictor domains [$F(16, 1001) = 5.59$]. An SNK multiple comparison test of the mean r values is specified in Table 1. Adult criminal history and criminogenic needs produced the greatest frequency of significant differences. Each of these was significantly different from family structure, intellectual functioning, personal distress, and SES.

PREDICTOR DOMAINS: CATEGORY II

With the exception of the risk scales domain, the predictor domains

Table 1. Mean Effect Sizes for Predictor Domains: Category I

Predictor (k)	N	$M\ r$	$M\ z^+$
Static[a]			
1. Age (56)	61,312	.15(.12)[b]	.11*
2. Criminal History: Adult (164)	123,940	.18(.13)[c]	.17*
3. History of Antisocial Behavior: Preadult (119)	48,338	.13(.13)[b]	.16*
4. Family Criminality (35)	32,546	.12(.08)	.07*
5. Family Rearing Practices (31)	15,223	.15(.17)[b]	.14*
6. Family Structure (41)	24,231	.10(.08)	.09*
7. Gender (17)	62,021	.10(.07)	.06*
8. Intellectual Functioning (32)	21,369	.07(.14)	.07*
9. Race (21)	56,727	.13(.15)	.17*
10. SES (23)	13,080	.06(.11)	.05*
Dynamic[a]			
11. Antisocial Personality (63)	13,469	.18(.12)[d]	.18*
12. Companions (27)	11,962	.18(.08)[e]	.21*
13. Criminogenic Needs (67)	19,809	.18(.10)[c]	.18*
14. Interpersonal Conflict (28)	12,756	.15(.10)[b]	.12*
15. Personal Distress (66)	19,933	.05(.15)	.05*
16. Social Achievement (168)	92,662	.15(.14)[e]	.13*
17. Substance Abuse (60)	54,838	.14(.12)[b]	.10*
Composite Measures			
18. Risk Scales (123)	57,811	.30(.14)	.30*

NOTES: k = effect sizes per predictor domain; N = subjects per predictor domain; $M\ r$ = mean Pearson r (S.D.); $M\ z^+ = \Sigma\ [(z_r) \times (n - 3)] \div \Sigma\ [(n - 3)]$, where n = number of subjects per effect size.

[a] $M\ r$: $F\ (16, 1001) = 5.59$, $p < .05$.

[b] 1, 3, 5, 14, 17 vs. 15; SNK post hoc comparison, $p < .05$.

[c] 2, 13 vs. 6, 8, 10, 15; SNK post hoc comparison, $p < .05$.

[d] 11 vs. 8, 10, 15; SNK post hoc comparison, $p < .05$.

[e] 12, 16 vs. 8, 15;SNK post hoc comparison, $p < .05$.

* $p < .05$.

from Category I were collapsed into 8 groups (see Table 2). All predictor domains were significantly greater than 0. There were significant differences among the 8 predictor domains [$F(7, 1010) = 10.00$]. The SNK multiple comparison test of the mean r values revealed that the predictor domains criminal history and criminogenic needs were significantly greater than those of family factors, intellectual functioning, personal distress, and SES.

Table 2. Mean Effect Sizes for Predictor Domains: Category II

Predictor (k)	N	$M\ r$	$M\ z^+$
Static[a]			
1. Age/Gender/Race (94)	180,060	.14(.12)[b]	.11*
2. Criminal History[c] (282)	171,159	.16(.13)[d]	.16*
3. Family Factors (107)	72,000	.12(.12)[b]	.08*
4. Intellectual Functioning (32)	21,369	.07(.14)	.07*
5. SES (23)	13,080	.06(.11)	.07*
Dynamic[a]			
6. Criminogenic Need Factors[e] (246)	113,153	.17(.11)[d]	.14*
7. Personal Distress (66)	19,933	.06(.15)	.05*
8. Social Achievement (168)	92,662	.15(.14)[f]	.13*
Static versus Dynamic[g]			
9. Static (536)	457,552	.12(.14)	.11*
10. Dynamic (482)	226,664	.15(.13)	.13*

NOTES: k = effect sizes per predictor domain; N = subjects per predictor domain; $M\ r$ = mean Pearson r (S.D.); $M\ z^+ = \Sigma\ [(z_r) \times (n - 3)] \div \Sigma\ [(n - 3)]$, where n = number of subjects per effect size.

[a] $M\ r$: $F\ (7, 1010) = 10.00$, $p < .05$.

[b] 1, 3 vs. 4, 7; SNK post hoc comparison, $p < .05$.

[c] Criminal history = adult plus preadult.

[d] 2, 6 vs. 3, 4, 5, 7; SNK post hoc comparison, $p < .05$.

[e] Criminogenic need factors = antisocial personality, companions, interpersonal conflict, criminogenic needs, and substance abuse.

[f] 8 vs. 4, 5, 7; SNK post hoc comparison, $p < .05$.

[g] $M\ r$: $F\ (1, 1016) = 6.18$, $p < .05$.

* $p < .05$.

The eight predictor domains were classified into dynamic and static factors. The dynamic grouping consisted of criminogenic needs factors, personal distress, and social achievement. The mean r values for dynamic (.15) and static (.12) were significantly different [$F(1, 1016) = 6.18$].

The CL effect size indicator provided another approach to examining the relative usefulness of the eight predictor domains from Table 2 as well as the static-dynamic comparison. The CL scores, summarized in Table 3, indicate the percentage of time that one of a pair of predictors produced larger correlations with outcome.

Table 3 can be read in the following way. With regard to direction, unbracketed scores favor the horizontal axis predictor while bracketed

Table 3. Common Language Effect Size Indicators

	CH	CN	F	I	PD	SES	SA
AGR	[54]	[58]	54	64	66	68	[53]
CH	—	[52]	58	68	69	71	51
CN		—	62	71	73	75	54
F			—	61	63	64	[57]
I				—	52	51	[66]
PD					—	[52]	[68]
SES						—	[70]

NOTES: Common language effect size indicators for mean *r* values. Bracketed Values favor vertical axis; unbracketed values favor horizontal axis. AGR = age, gender, race; CH = criminal history/history of antisocial behavior; CN = criminogenic need factors; F = family factors; I = intellectual functioning; PD = personal distress; SES = socioeconomic status or social class of origin; SA = social achievement.

scores favor the vertical axis predictor. For example, in comparing criminogenic needs (CN) with personal distress (PD), one can see that 73% of the time CN produced higher correlations with recidivism than did PD.

In the case of the static-dynamic comparison (Table 2), the CL score was 54% in favor of the dynamic predictor domain.

ACTUARIAL MEASURES

Table 4 summarizes the mean effect sizes of the composite risk and personality scales with recidivism. All of the instruments predicted recidivism significantly different from zero. Amongst the risk scales, the LSI-R produced the highest correlation with recidivism (r = .35), but it was not significantly greater than the SFS, Wisconsin, or Other risk scale domains [$F(3, 119) = 1.52$]. The Other domain consisted of SFS clones, that is, instruments containing about 5 to 10 items, almost all of which were static in nature.

The LSI-R produced CL scores of 76% and 67% with the Wisconsin and SFS, respectively, when mean r was the dependent variable.

A comparison of the mean r values associated with the antisocial personality measures revealed a significant difference between measures [$F(2, 59) = 4.01$]. The SNK multiple comparison test reported that the PCL was a significantly better predictor than either the MMPI-based measures or Other domain.

The CL analysis indicated that 83% of the time the PCL produced

larger Pearson *r* correlations with recidivism than did the MMPI.[2]

DISCUSSION AND RECOMMENDATIONS

Prior to discussing the results it must be noted that the generalization of the results of any meta-analysis is limited by the nature of the studies examined.

Some valuable studies (e.g., Gendreau et al., 1979a) could not be used because the researchers reported their results in formats (e.g., regression analyses) from which Pearson *r*s could not be calculated. In addition, little attempt was made to retrieve unpublished studies that were not immediately available. A common assumption is that one of the reasons some

---

2. As a result of collecting the literature and analyzing the data, some other comparisons came to light that merited closer examination.

Personal Distress: Within the personal distress domain, 24 of 66 effect sizes tapped the psychiatric symptomatology dimension through items such as schizophrenia, psychosis, and prior psychiatric history. The mean *r* (S.D.) for this subset with recidivism was .00 (.17).

Family Factors: Our analysis of family factors did not include studies from the gene-crime relationship because Walters (1992) has already conducted a thorough meta-analysis in this area. He reported small correlations between genetic background and criminal behavior. We determined whether, in fact, genetic background predicted criterion significantly greater than 0. Only those studies that were twin and adoption studies (the most stringent comparison of the gene-crime relationship) and used an official measure of outcome were assessed. Fifteen effect sizes from Tables 2 and 3 of Walter's (1992) study were generated and analyzed using the Hedges and Olkin (1985) formulas. The mean *r* with recidivism was .08. The z^+ was also .08, indicating that genetic background was a significant predictor of recidivism.

Measuring Change: Andrews and Bonta (1994) and Bonta (1996) have stressed the importance of measuring change with dynamic predictors. Six studies were located that assessed offenders at two points in time and derived a change score, which was then correlated with future recidivism. A meta-analysis of their results was not possible because five of the six studies did not report data in a suitable form or had very small cell frequencies. The following narrative will have to suffice.

Recidivism rates changed between 30% and 50% when an offender's status moved from high to low risk or vice versa (Motiuk et al., 1986; Motiuk, 1991). Change scores predicted recidivism as well as measures taken at either entry to prison or prior to release (Gendreau et al., 1979b). The effect size for change scores may be quite substantial. Data from Table 3 of Bonta (1996) were recalculated (for the low-high/high-low cells) yielding a X^2 (1, N = 808) = 116.41, which is equivalent to an *r* of .38.

Type of Outcome: While the issue is rarely, if ever, raised in the research literature, one is occasionally asked by practitioners which official measure of recidivism is the most sensitive. Four criteria—arrest, conviction, incarceration, and parole violation—were compared as to differences in mean effect size, where values ranged from .13 to .19. There was a significant difference among the mean values [F(3, 894) = 6.71]. The SNK multiple comparison test reported that the mean *r* values associated with incarceration were significantly greater than those of conviction or parole violation. The CL scores for the four outcome indices were calculated. In all comparisons, however, the CL scores were less than 60%.

Table 4. Mean Effect Sizes for Risk and Antisocial Personality Scales

Predictor (k)	N	$M\ r$	$M\ z^+$
Risk Scales[a]			
1. LSI-R (28)	4,579	.35(.08)	.33*
2. SFS (15)	9,850	.29(.10)	.26*
3. Wisconsin (14)	14,092	.27(.08)	.32*
4. Other (66)	29,290	.30(.17)	.30*
Antisocial Personality Scales[b]			
5. MMPI Based (16)	3,420	.16(.09)	.21*
6. PCL (9)	1,040	.28(.09)[c]	.29*
7. Other (37)	8,875	.16(.13)	.16*

NOTES: k = effect sizes per predictor domain; N = subjects per predictor domain; $M\ r$ = mean Pearson r (S.D.); $M\ z^+ = \Sigma\ [(z_r) \times (n - 3)] \div \Sigma\ [(n - 3)]$, where n = number of subjects per effect size.

[a] $M\ r$: $F\ (3, 119) = 1.52$, $p > .05$.

[b] $M\ r$: $F\ (2, 59) = 4.01$, $p < .05$.

[c] 6 vs. 5, 7; Student–Newman–Keuls post hoc comparison, $p < .05$.

* $p < .05$.

studies are not published is that they lack methodological rigor, which in turn, affects the magnitude of effect sizes (see Lipsey and Wilson, 1993). Lipsey and Wilson's (1993) analysis applied to treatment studies, but so far, prediction studies have not shown similar results (Goggin and Gendreau, 1995).

Another methodological point concerns one of the goals of meta-analysis. Hunter and Schmidt (1990) are interested in determining the maximum value that can be obtained in prediction if all variables are perfectly measured. Others insist that the goal of meta-analysis is to "teach us better what *is*, not what might some day be in the best of all possible worlds . . ." (italics added; Rosenthal, 1991:25). We are of the latter view and did not attempt to adjust statistically for methodological artifacts, which may or may not have had an impact on the magnitude of the effect sizes obtained.

The data base was, regrettably, virtually silent on the prediction of recidivism among female offenders, minority groups, white-collar offenders, and some important sample characteristics, such as risk level and the psychological make-up of the subjects studied. Much of the effect size data on dynamic predictor domains came from Canada, where there has been a strong emphasis on the assessment of individual differences (Andrews and Bonta, 1994).

One should not assume that many of the correlations found in this meta-analysis (e.g., .10 — .30) are inconsequential. In fact, mean *r* values in this range can be indicative of substantial practical import (Hunter and Schmidt, 1990; Rosnow and Rosenthal, 1993). Indeed, the percentage improvement in predicting recidivism can equal the value of *r*, assuming base rates and selection ratios that are not in the extreme (Rosenthal, 1991:134).

The fact that the data base consisted of just over 1,000 effect sizes involving almost 750,000 subjects suggests that reasonable confidence can be placed in the results. Additional research, in our view, is not likely to change the direction or ordering of the results of the predictor domains to any marked degree.

The remainder of this discussion addresses the questions raised in the introduction.

PREDICTOR DOMAINS

The meta-analysis provided further confirmation of the narrative reviews, which concluded that variables such as age, criminal history, companions, family factors, gender, social achievement, and substance abuse are significant and potent predictors of recidivism. On the other hand, it offered some important insights into several other predictor domains.

The time is long past when those offender risk factors that are dynamic in nature can be cavalierly ignored. Indeed, criminogenic needs produced higher correlations with recidivism (see Table 3) a much higher percentage of the time than did several other predictor domains. When considering all predictor domains, a statistically significant difference was found in favor of dynamic risk factors, but the CL effect size indicator was only 54%. Moreover, the two major static and dynamic categories, criminal history and criminogenic needs, were almost identical in predicting recidivism. While very few studies have assessed how well changes over time within dynamic factors predict recidivism, the data suggest that changes in criminogenic needs may produce strong correlations in that regard.

Early family factors and history of preadult antisocial behavior are rarely included in adult offender risk prediction instruments.[3] Fortuitously, a number of estimable studies (producing 103 effect sizes) were located that followed offenders from early years to adulthood. The combined family factors domain (Table 2) and preadult history of antisocial behavior (Table 1) produced correlations of .12 and .13 with recidivism,

3. Typically, risk prediction instruments for adults assess just one aspect of this predictor domain and employ one or two items in so doing. For example, the LSI-R has one item (no. 5) in this regard (i.e., "arrested under age 16").

respectively, demonstrating once again that antisocial risk factors in childhood can have far-reaching influence (e.g., Stattin and Magnusson, 1989).

Much controversy has focused upon how well personal distress, intelligence, and SES predict recidivism (Andrews and Bonta, 1994; Herrnstein and Murray, 1994; Tittle and Meier, 1990). From a treatment standpoint, the important result centered on the fact that personal distress turned out to be quite a weak predictor of recidivism. Moreover, one of the components of this domain, psychiatric symptomatology, which has characteristically been perceived as an important predictor of reoffending in the field of psychiatry (Phillips et al., 1988), did not correlate (r = .00) with recidivism. This finding was based on few effect sizes; more research is needed to confirm this tentative result. It would be reasonable, therefore, to assume that programs that insist on alleviating offenders' personal distress, as many do (Gendreau et al., 1994), will have little success in reducing offender recidivism. Meta-analyses of the offender treatment literature (e.g., Andrews et al., 1990b) are also supportive of this conclusion.

The studies in the meta-analysis that included measures of IQ were of the "traditional" sort, that is, standard paper and pencil tests that measured linguistic and mathematical abilities. Although these sorts of IQ measures can produce modest correlations with criminal behavior over long periods of time (Moffitt et al., 1994), it is generally agreed that this type of IQ assessment has reached its limits (Gardner, 1995). A much more productive strategy would be to focus on what is called practical or tacit intelligence, which is defined as the ability to learn and profit from experience, effectively monitor one's own and other's feelings and needs, and solve the problems of everyday life (Gardner, 1983; Sternberg et al., 1995).

This meta-analysis extended Tittle and Meier's (1990, 1991) pessimistic conclusions regarding the social class-crime link with delinquent samples to that of adult offenders. It is difficult to judge how social class theories will evolve in the future; for speculations on this matter see Andrews and Bonta (1994) and Tittle and Meier (1990). The most probable scenario is that social class theories will incorporate more psychological concepts (e.g., Agnew, 1992).

How well might the results from the meta-analysis generalize to specialized offender groups? Few violence prediction studies that predicted the occurrence of violence versus no criminal activity were retrieved. Our reading of the literature indicates that the strongest predictors identified in this meta-analysis also apply to violent offenders (Harris et al., 1993; Reiss and Roth, 1993). As well, composite measures of general recidivism (i.e., LSI-R) correlate highly (r = .78) with measures intended to predict violence (i.e., PCL-R) (Loza and Simourd, 1994). One area in which the predictors of violent offending may be quite different is that of impulsivity

combined with overly hostile attributions of other people's intent (Serin and Kuriychuk, 1994). Sex offenders present a somewhat different picture. At the risk of generalizing across such a complex group, there do appear to be a few predictors, centering on the offense itself, that are unique to this population (Hanson and Bussière, 1995).

In regard to theory development, the results from the meta-analysis are most supportive of recent advancements in differential association and social learning theories (see Andrews and Bonta, 1994:104–124). These authors assert that it is absolutely essential that criminogenic needs and antisocial associates are two of the strongest correlates of criminal conduct. Criminogenic needs establish the standards of conduct and generate the rationale for engaging in antisocial behavior. Antisocial associates provide the opportunity for antisocial modeling to occur, govern the rewards and costs of such behavior, and influence antisocial attitudes.

The less potent predictors in this meta-analysis (e.g., SES, personal distress, intellectual functioning) have traditionally been associated with the anomie/strain and subcultural theories and biologically oriented theories.

ACTUARIAL MEASURES FOR PREDICTING RECIDIVISM

Composite measures of risk, on average, produced substantially greater correlations with recidivism than antisocial personality scales. This is not surprising, because risk measures generally sample from a much wider variety of predictor domains than personality scales.

Among the former, the LSI-R produced higher correlations with recidivism than the SFS, the Wisconsin, or the Other category. While the mean differences among the four measures were not statistically significant,[4] the CL effect size indicator provided a result of practical importance. The LSI-R produced larger correlations with recidivism than did the three other risk measures between 62% and 76% of the time. The LSI-R, therefore, appears to be the current measure of choice. An impressive number of studies confirming its predictive validity with recidivism and prison adjustment have been generated for a variety of offender populations (i.e., adults, juveniles, natives, females) (Andrews and Bonta, 1995).

In the area of antisocial personality assessment, a noteworthy finding was that Hare's (1991) PCL-R produced significantly greater correlations with recidivism than the widely used MMPI-based systems. The PCL-R specializes in assessing the psychopathic dimension of antisocial personality. It is recommended by clinicians who are concerned with predicting violence (Harris et al., 1993).

4. See J. Cohen (1994) and Schmidt (1992) for a criticism of the use of standard significance testing, which they claim, often results in Type II errors and a failure to account for a realistic estimate of the magnitude of the effect sizes under study.

CONCLUSION AND RECOMMENDATIONS

In conclusion, the modest contribution from this meta-analysis has been to clarify which predictor domains and actuarial measures of risk will be most useful to practitioners and policymakers. In regard to the assessment of static predictors, protocols should contain any reliable information that accurately captures early family life and social adjustment risk factors. Dynamic risk factors, particularly those of criminogenic needs, must be included and reassessed over time. The choice of criterion (e.g., reconviction) should depend on the goals of the assessment. Of the available risk measures, the LSI-R is recommended. In the case of specialized offender populations, additional measures (e.g., PCL-R) might be used in conjunction with a general measure of risk.

REFERENCES

Agnew, Robert
1992 Foundation for a general strain theory of crime and delinquency. Criminology 30:47–86.

Andrews, D.A. and James Bonta
1994 The Psychology of Criminal Conduct. Cincinnati, Ohio: Anderson.
1995 LSI-R: The Level of Service Inventory - Revised. Toronto, Ont.: Multi-Health Systems, Inc.

Andrews, D.A. and J. Stephen Wormith
1989 Personality and crime: Knowledge destruction and construction in criminology. Justice Quarterly 6:289–309.

Andrews, D.A., James Bonta, and Robert D. Hoge
1990a Classification for effective rehabilitation: Rediscovering psychology. Criminal Justice and Behavior 17: 19–52.

Andrews, D.A. Ivan Zinger, Robert D. Hoge, James Bonta, Paul Gendreau, and Francis T. Cullen
1990b Does correctional treatment work? A psychologically informed meta-analysis. Criminology 28:369–404.

Baird, Christopher S.
1981 Probation and parole classification: The Wisconsin model. Corrections Today 43:36–41.

Bonta, James
1996 Risk-needs assessment and treatment. In Alan Harland (ed.), Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply. Thousand Oaks, Calif.: Sage.

Bonta, James and Paul Gendreau
1990 Reexamining the cruel and unusual punishment of prison life. Law and Human Behavior 14:347–372.

Clear, Todd R. and Kenneth W. Gallagher
1985 Probation and parole supervision: A review of current classification. Crime and Delinquency 31:423–443.

Cohen, Albert K.
1955 Delinquent Boys: The Culture of the Gang. Glencoe, Ill.: Free Press.

Cohen, Jacob
1994 The earth is round ($p < .05$). American Psychologist 49:997–1003.

Cooper, Harris and Larry V. Hedges
1994 Handbook of Research Synthesis. New York: Russell Sage Foundation.

Cronbach, Lee J.
1990 Essentials of Psychological Testing. New York: HarperCollins.

Feeley, Malcolm M. and Jonathon Simon
1992 The new penology: Notes on the emerging strategy of corrections and its implications. Criminology 30:449–474.

Gardner, Howard
1983 Frames of Mind: The Theory of Multiple Intelligence. New York: Basic Books.
1995 Cracking open the IQ box. The American Prospect 20:71–80.

Gendreau, Paul
1996 Offender rehabilitation: What we know and what needs to be done. Criminal Justice and Behavior 23:144–161.

Gendreau, Paul and D.A. Andrews
1990 Tertiary prevention: What the meta-analysis of the offender treatment literature tells us about "what works." Canadian Journal of Criminology 32:173–184.

Gendreau, Paul and Robert R. Ross
1987 Revivification of rehabilitation: Evidence from the 80's. Justice Quarterly 4:349–407.

Gendreau, Paul, Brian A. Grant, Mary Leipciger, and Steven Collins
1979a Norms and recidivism rates for the MMPI and selected experimental scales on a Canadian delinquent sample. Canadian Journal of Behavioural Science 11:21–31.

Gendreau, Paul, Brian A. Grant, and Mary Leipciger
1979b Self-esteem, incarceration and recidivism. Criminal Justice and Behavior 6:67–75.

Gendreau, Paul, Claire Goggin, and Helen Annis
1990 Survey of existing substance abuse programs. Forum on Corrections Research 2:6–8.

Gendreau, Paul, D.A. Andrews, Claire Goggin, and Françoise Chanteloupe
1992 The Development of Clinical and Policy Guidelines for the Prediction of Criminal Behaviour in Criminal Justice Settings. Ministry Secretariat, User Report. Ottawa, Ont.: Solicitor General of Canada.

Gendreau, Paul, Francis T. Cullen, and James Bonta
1994 Intensive rehabilitation supervision: The next generation in community corrections? Federal Probation 58:72–78.

Goddard, Henry Herbert
1920 Human Efficiency and Levels of Intelligence. Princeton, N.J.: Princeton University Press.

Goggin, Claire E. and Paul Gendreau
1995 Predicting psychiatric rehospitalization: A meta-analysis. Unpublished manuscript, University of New Brunswick, Saint John, N.B., Canada.

Gough, Harrison G.
1957 Manual for the California Psychological Inventory. Palo Alto, Calif.: Consulting Psychologists Press.

Gough, Harrison G., E.A. Wenk, and V.V. Rozynko
1965 Parole outcome as predicted from the CPI, the MMPI, and a Base Expectancy Table. Journal of Abnormal Psychology 70:432–441.

Hanson, R. Karl and Monique Bussière
1995 Predictors of sexual offender recidivism. Ministry Secretariat. Ottawa, Ont.: Solicitor General of Canada.

Hare, Robert D.
1991 The Revised Psychopathy Checklist. Toronto: Multi-Health Systems.

Harris, Grant T., Marnie E. Rice, and Vernon L. Quinsey
1993 Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. Criminal Justice and Behavior 20:315–335.

Hedges, Larry V. and Ingram Olkin
1985 Statistical Methods for Meta-Analysis. San Diego, Calif.: Academic Press.

Herrnstein, Richard J. and Charles A. Murray
1994 The Bell Curve: Intelligence and Class Structure in American Life. New York: Free Press.

Hirschi, Travis and Michael J. Hindelang
1977 Intelligence and delinquency: A revisionist review. American Sociological Review 42:571–587.

Hoffman, Peter B.
1983 Screening for risk: A revised salient factor score (SFS 81). Journal of Criminal Justice 11:539–547.

Hoge, Robert D., Alan W. Leschied, and D.A. Andrews
1993 An Investigation of Young Offender Services in the Province of Ontario: A Report of the Repeat Offender Project. Toronto, Ont.: Ontario Ministry of Community and Social Services.

Hunter, John E. and Frank L. Schmidt
1990 Methods of Meta-Analysis: Correcting Error and Bias in Research Findings. Newbury Park, Calif.: Sage.

Jones, Peter R.
1996 Risk prediction in criminal justice. In Alan T. Harland (ed.), Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply. Thousand Oaks, Calif.: Sage.

Kaminer, Wendy
1992 I'm Dysfunctional - You're Dysfunctional. New York: Addison–Wesley.

LeBlanc, Marc, Marcel Ouimet, and Richard E. Tremblay
1988 An integrative control theory of delinquent behaviour: A validation 1976–1985. Psychiatry 51:164–176.

Lipsey, Mark W.
1992 Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T.D. Cook, H. Cooper, D.S. Cordray, H. Hartmann, Larry V. Hedges, R.J. Light, T.A. Louis, and F. Mosteller (eds.), Meta-analysis for Explanation. New York: Russell Sage Foundation.

Lipsey, Mark W. and David B. Wilson
1993 The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. American Psychologist 48:1181–1209.

Little, Kenneth B. and Edwin S. Shneidman
1959 Congruencies among interpretations of psychological test and anamnestic data. Psychological Monographs: General and Applied 73:1–42.

Loza, Wagdy and David J. Simourd
1994 Psychometric evaluation of the Level of Supervision Inventory (LSI-R) among male Canadian federal offenders. Criminal Justice and Behavior 21:468–480.

Matza, David
1964 Delinquency and Drift. 2d ed. New York: John A. Wiley & Sons.

Mauer, Marc
1994 Americans behind bars: The international use of incarceration, 1992–1993. Washington, D.C.: Sentencing Project.

McGraw, Kenneth O. and S.P. Wong
1992 A common language effect size. Psychological Bulletin 111:361–365.

Meehl, Paul E.
1954 Clinical Versus Statistical Prediction. Minneapolis: University of Minnesota Press.

Merton, Robert K.
1957 Social Theory and Social Structure. New York: Free Press.

Moffitt, Terrie E., Donald R. Lynam, and Phil A. Silva
1994 Neuropsychological tests predicting persistent male delinquency. Criminology 32:277–300.

Motiuk, Laurence L.
1991 Antecedents and consequences of prison adjustment: A systematic assessment and reassessment approach. Unpublished doctoral dissertation, Carleton University, Ottawa, Ontario, Canada.
1993 Where are we in our ability to assess risk? Forum on Corrections Research 5:14–19.

Motiuk, Laurence L., James Bonta, and D.A. Andrews
1986 Dynamic predictive criterion validity in offender assessment. Paper presented at the June meeting of the Canadian Psychological Association, Ottawa, Ontario, Canada.

Phillips, Michael R., Aron S. Wolf, and David J. Coons
1988 Psychiatry and the criminal justice system: Testing the myths. American Journal of Psychiatry 145:605–610.

Reiss, Albert J., Jr. and Jeffrey A. Roth
1993 Understanding and Preventing Violence. Washington, D.C.: National Academy Press.

Rice, Marnie E. and Grant T. Harris
1995 Violent recidivism: Assessing predictive validity. Journal of Consulting and Clinical Psychology 63:737–748.

Rosenthal, Robert
1991 Meta-Analytic Procedures for Social Research. Newbury Park, Calif.: Sage.

Rosnow, Robert L. and Robert Rosenthal
1993 Beginning Behavioral Research: A Conceptual Primer. New York: Macmillan.

Rowe, David C. and D. Wayne Osgood
1984 Heredity and sociological theories of delinquency: A reconsideration. American Sociological Review 49:526–540.

Schmidt, Frank L.
1992 What do data really mean? American Psychologist 47:1173–1181.

Serin, Ralph C. and Morris Kuriychuk
1994 Social and cognitive processing deficits in violent offenders. International Journal of Law and Psychiatry 17:431–441.

Serin, Ralph C., Ray D. Peters, and Howard E. Barbaree
1990 Predictors of psychopathy and release outcome in a criminal population. Psychological Assessment: A Journal of Consulting and Clinical Psychology 2:419–422.

Simourd, David J., James Bonta, D.A. Andrews, and Robert D. Hoge
1991 Criterion validity of assessments of psychopathy: A meta-analysis. Unpublished manuscript, Carleton University, Ottawa, Ontario, Canada.

Stattin, Håkan and David Magnusson
1989 The role of early aggressive behaviour in the frequency, seriousness and types of later crime. Journal of Consulting and Clinical Psychology 57:710–718.

Sternberg, Robert J., Richard K. Wagner, Wendy M. Williams, and Joseph A. Horvath
1995 Testing common sense. American Psychologist 50:912–926.

Tittle, Charles R. and Robert F. Meier
1990 Specifying the SES/delinquency relationship. Criminology 28:271–299.
1991 Specifying the SES/delinquency relationship by social characteristics of contexts. Journal of Research in Crime and Delinquency 28:430–455.

Walters, Glenn D.
1992 A meta-analysis of the gene-crime relationship. Criminology 30:595–613.

Widom, Cathy S. and Hans Toch
1993 The contribution of psychology to criminal justice education. Journal of Criminal Justice Education 4:251–272.

Wilson, James Q. and Richard J. Herrnstein
1985 Crime and Human Nature. New York: Simon & Schuster.

Yeaton, William H. and Paul M. Wortman
1993 On the reliability of meta-analytic reviews: The role of intercoder agreement. Evaluation Review 17:292–309.

Paul Gendreau is a Professor in the Department of Psychology and Director, Centre for Criminal Justice Studies, at the University of New Brunswick in Saint John. He has been addressing the "nothing works" issue in offender treatment during the past two decades. His other research interests include the evaluation of the quality of offender treatment programs and assessing the effects of prison life. Correspondence concerning this article should be addressed to this author at the Department of Psychology, University of New Brunswick, P.O. Box 5050, Saint John, New Brunswick, Canada, E2L 4L5.

Tracy Little recently completed her M.A. thesis in psychology on issues pertaining to the prediction of offender recidivism.

Claire Goggin has completed her M.A. in psychology. She is currently consultant to New Brunswick's Mental Health Commission and Department of the Solicitor General. Her research interests concern the prediction of psychiatric rehospitalization and offender recidivism.

APPENDIX

CODING CRITERIA FOR PREDICTOR DOMAINS: CATEGORY I

STATIC PREDICTORS

1. Age: at time of data collection/assessment.
2. Criminal history: adult-prior arrest, probation, jail, conviction, incarceration, prison misconducts.
3. History of antisocial behavior: preadult—prior arrest, probation, jail, conviction, incarceration, alcohol/drug abuse, aggressive behavior, conduct disorder, behavior problems at home and school, delinquent friends.
4. Family criminality: parents and/or siblings in trouble with the law.
5. Family rearing practices: lack of supervision and affection, conflict, abuse.
6. Family structure: separation from parents, broken home, foster parents.
7. Gender.
8. Intellectual functioning: WAIS/WISC, Raven, Porteous Q score, learning disabilities, reading level.
9. Race: white vs. black/Hispanic/native.
10. Social class of origin: socioeconomic status (SES) of parents (parental occupation, education, or income).

DYNAMIC PREDICTORS

11. Antisocial personality/sociopathy/psychopathy scales: MMPI Pd, Megargee system, EPI-Psychoticism, CPI-Soc, PCL-R, DSM-III personality disorders, any indices of egocentric thinking.
12. Companions: identification/socialization with other offenders.
13. Criminogenic needs: antisocial attitudes supportive of an antisocial lifestyle and behavior regarding education, employment.
14. Interpersonal conflict: family discord, conflict with significant others.
15. Personal distress: anxiety, depression, neuroticism, low self-esteem, psychiatric symptomatology (i.e., psychotic episodes, schizophrenia, not guilty by reason of insanity, affective disorder), attempted suicide, personal inadequacy.
16. Social achievement: marital status, level of education, employment history, income, address changes.
17. Substance abuse: recent history of alcohol/drug abuse.

COMPOSITE MEASURES

18. LSI-R, SFS, Wisconsin, Other risk scales.

BIBLIOGRAPHY OF STUDIES IN THE META-ANALYSIS

Adams, Don and Joel Fischer
1976 The effects of prison residents' community contacts on recidivism rates. Corrective and Social Psychiatry and Journal of Behavioral Technology, Methods, and Therapy 22:21–27.

Andrews, D.A. and Walter Friesen
1987 Assessments of anticriminal plans and the prediction of criminal futures: A research note. Criminal Justice and Behavior 14:33–37.

Andrews, D.A., Jerry J. Kiessling, Susan Mikus, and David Robinson
1986 The construct validity of interview-based risk assessment in corrections. Canadian Journal of Behavioural Science 18:460–470.

Babst, Dean V., Mary Koval, and M.G. Neithercutt
1972 Relationship of time served to parole outcome for different classifications of burglars based on males paroled in fifty jurisdictions in 1968 and 1969. Journal of Research in Crime and Delinquency 9:99–116.

Babst, Dean V., William H. Moseley, James Schmeidler, M.G. Neithercutt, and Mary Koval
1976 Assessing the length of institutionalization in relation to parole outcome. Criminology 14:41–54.

Baker, Laura A., Wendy Mack, Terrie E. Moffitt, and Sarnoff Mednick
1989 Sex differences in property crime in a Danish adoption cohort. Behavior Genetics 19:355–370.

Barbaree, Howard E. and William L. Marshall
1988 Deviant sexual arousal, offense history, and demographic variables as predictors of reoffense among child molesters. Behavioral Sciences and the Law 6:267–280.

Barnett, Arnold, Alfred Blumstein, and David P. Farrington
1989 A prospective test of a criminal career model. Criminology 27:373–388.

Bayer, Betty M., James L. Bonta, and Laurence L. Motiuk
1985 The PD subscales: An empirical evaluation. Journal of Clinical Psychology 41:780–788.

Bennett, Lawrence A.
1974 Self-esteem and parole adjustment. Criminology 12:346–363.

Bieber, Stephen L., Richard A. Pasewark, Katherine Bosten, and Henry J. Steadman
1988 Predicting criminal recidivism of insanity acquittees. International Journal of Law and Psychiatry 11:105–112.

Black, Tony and Penny Spinks
1985 Predicting outcomes of mentally disordered and dangerous offenders. In David P. Farrington and Roger Tarling (eds.), Prediction in Criminology. Albany: State University of New York Press.

Bogenberger, Robert P., Richard A. Pasewark, Howard Gudeman, and Stephen L. Bieber
1987 Follow-up of insanity acquittees in Hawaii. International Journal of Law and Psychiatry 10:283–295.

Bonta, James
1989 Native inmates: Institutional response, risk, and needs. Canadian Journal of Criminology 31:49–61.

Bonta, James and Laurence L. Motiuk
1985 Utilization of an interview-based classification instrument: A study of correctional halfway houses. Criminal Justice and Behavior 12:333–352.
1987 The diversion of incarcerated offenders to correctional halfway houses. Journal of Research in Crime and Delinquency 24:302–323.
1990 Classification to halfway houses: A quasi-experimental evaluation. Criminology 28:497–506.

Bonta, James, Stan Lipinski, and Michael Martin
1992 The characteristics of aboriginal recidivists. Canadian Journal of Criminology 38:517–522.

Broadhurst, Roderick G. and Ross A. Maller
1990 The recidivism of prisoners released for the first time: Reconsidering the effectiveness question. Australian and New Zealand Journal of Criminology 23:88–104.
1991 Sex Offending and Recidivism. Report No. 3. Crime Research Center. Nedlands: The University of Western Australia.

Broadhurst, Roderick G., Ross A. Maller, Maxwell G. Maller, and Jennifer Duffecy
1988 Aboriginal and non-aboriginal recidivism in Western Australia: A failure rate analysis. Journal of Research in Crime and Delinquency 25:83–108.

Brown, Lawrence D.
1978 The development of a parolee classification system using discriminant analysis. Journal of Research in Crime and Delinquency 15:92–108.

Buikhuisen, W. and H.A. Hoekstra
1974 Factors related to recidivism. British Journal of Criminology 14:62–69.

Buikhuisen, W. and B.W.G.P. Meijs
1983 A psychosocial approach to recidivism. In Katherine Teilmann Van Dusen and Sarnoff A. Mednick (eds.), Prospective Studies of Crime and Delinquency. Boston: Kluwer–Nijhoff.

Carlson, Kenneth A.
1973 Some characteristics of recidivists in an Ontario institution for adult male first incarcerates. Canadian Journal of Criminology and Corrections 15:397–409.

Carroll, John S., Richard L. Wiener, Dan Coates, Jolene Galegher, and James J. Alibrio
1982 Evaluation, diagnosis, and predicting in parole decision making. Law & Society Review 17:199–228.

Cloninger, C. Robert and Samuel B. Guze
1973 Psychiatric disorders and criminal recidivism. Archives of General Psychiatry 29:266–269.

Donnelly, Henry C. and Gerald H. Bala
1991 1985 Releases: Five Year Post Release Follow-Up. Albany, N.Y.: Department of Correctional Services.

Eron, Leonard D. and L. Rowell Huesmann
1984 The relation of prosocial behaviour to the development of aggression and psychopathology. Aggressive Behavior 10:201–211.

Farrington, David P.
1983 Offending from 10-25 years of age. In Katherine Teilmann Van Dusen and Sarnoff Mednick (eds.), Prospective Studies of Crime and Delinquency. Boston: Kluwer–Nijhoff.

Farrington, David P., Louise Biron, and Marc LeBlanc
1982 Personality and delinquency in London and Montreal. In John Gunn and David P. Farrington (eds.), Abnormal Offenders, Delinquency, and the Criminal Justice System. New York: John Wiley & Sons.

Farrington, David P., Gwen Gundry, and David J. West
1975 The familial transmission of criminality. Medicine, Science, and the Law 15:177–186.

Feder, Lynette
1991 A comparison of the community adjustment of mentally ill offenders with those from the general prison population: An 18-month follow up. Law and Human Behavior 15:477–493.

Gabrielli, William F. and Sarnoff A. Mednick
1983 Genetic correlates of criminal behaviour. American Behavioral Scientist 27:59–74.

Gendreau, Paul, Brian A. Grant, and Mary Leipciger
1979 Self-esteem, incarceration and recidivism. Criminal Justice and Behavior 6:67–75.

Gendreau, Paul, Patrick Madden, and Mary Leipciger
1979 Norms and recidivism for first incarcerates: Implications for programming. Canadian Journal of Criminology 21:416–441.

Gibbens, T.C.N.
1984 Borstal boys after 25 years. British Journal of Criminology 24:49–62.

Gordon, Alistair M.
1983 Drugs and delinquency: A ten year follow-up of drug clinic patients. British Journal of Psychiatry 142:169–173.

Gottfredson, Michael R. and Don R. Gottfredson
1980 Decision Making in Criminal Justice: Toward the Rational Exercise of Discretion. Cambridge, Mass.: Ballinger.

Gottfredson, Don M., Leslie T. Wilkins, and Peter B. Hoffman
1978 Guidelines for Parole and Sentencing: A Policy Control Method. Lexington, Mass.: D.C. Heath.

Greenberg, William M., Pritesh J. Shah, and Marilyn Seide
1993 Recidivism on an acute psychiatric forensic service. Hospital and Community Psychiatry 44:583–585.

Grygier, Tadeusz, Frank Blum, and O.R. Porebski
1971 Decision and outcome: Studies in parole prediction. Canadian Journal of Criminology and Corrections 13:133–146.

Gunn, J., G. Robertson, S. Dell, and C. Way
1978 Psychiatric Aspects of Imprisonment. London: Academic Press.

Hamparian, Donna
1987 How well can we predict for juveniles? Juvenile delinquency and adult crime. In Fernand N. Dutile and Cleon H. Foust (eds.), The Prediction of Criminal Violence. Chicago: Charles C Thomas.

Hann, Robert G. and William G. Harman
1988 Release risk prediction: Testing the Nuffield scoring system for native and female inmates. Ottawa: Ministry of the Solicitor General of Canada.
1992 Predicting General Release Risk for Canadian Penitentiary Inmates. Ottawa: Corrections Branch of the Ministry of the Solicitor General of Canada.

Hare, Robert D., Leslie M. McPherson, and Adelle E. Forth
1988 Male psychopaths and their criminal careers. Journal of Consulting and Clinical Psychology 56:710–714.

Hart, Stephen D., Philip R. Kropp, and Robert D. Hare
1988 Performance of male psychopaths following conditional release from prison. Journal of Consulting and Clinical Psychology 56:227–232.

Hodgins, Sheilagh
1983 A follow-up study of persons found incompetent to stand trial and/or not guilty by reason of insanity in Quebec. International Journal of Law and Psychiatry 6:399–411.
1987 Men found unfit to stand trial and/or not guilty by reason of insanity: Recidivism. Canadian Journal of Criminology 29:51–70.

Hoffman, Peter B.
1983 Screening for risk: A revised Salient Factor Score (SF81). Journal of Criminal Justice 11:539–547.

Hoffman, Peter B. and James L. Beck
1974 Parole decision-making: A salient factor score. Journal of Criminal Justice 33:195–206.
1976 Salient Factor Score validation: A 1972 release cohort. Journal of Criminal Justice 4:69–76.
1980 Revalidating the Salient Factor Score: A research note. Journal of Criminal Justice 8:185–188.
1984 Burnout-age at release from prison and recidivism. Journal of Criminal Justice 12:617–623.
1985 Recidivism among released federal prisoners: Salient Factor Score and five-year follow-up. Criminal Justice and Behavior 12:501–507.

Hoffman, Peter B., Barbara Stone-Meierhoefer, and James L. Beck
1978 Salient Factor Score and release behaviour: Three validation samples. Law and Human Behavior 2:47–62.

Holcomb, William R. and Paul R. Ahr
1988 Arrest rates among young adult psychiatric patients treated in inpatient and outpatient settings. Hospital and Community Psychiatry 39:52–57.

Holland, Terrill R., Norman Holt, Mario Levi, and Gerald E. Beckett
1983 Comparison and combination of clinical and statistical predictions of recidivism among adult offenders. Journal of Applied Psychology 68:203–211.

Huesmann, L. Rowell, Leonard D. Eron, Monroe M. Lefkowitz, and Leopold O. Walder
1984 Stability of aggression over time and generations. Developmental Psychology 20:1120–1134.

Janes, Cynthia L., Victor M. Hesselbrock, Darcy Gilpin Myers, and Janet H. Peniman
1979 Problem boys in young adulthood: Teacher ratings and twelve-year follow-up. Journal of Youth and Adolescence 8:453–472.

Jenkins, W.O., E.K. deValera, and J.B. Muller
1977 The Behavioral Evaluation Treatment and Analysis (BETA) system in the prediction of criminal and delinquent behavior. Quarterly Journal of Corrections 1:44–50.

Kassebaum, Gene G., David A. Ward, and Daniel Wilner
1971 Prison Treatment and Parole Survival: An Empirical Assessment. Toronto: John Wiley & Sons.

Klein, Stephen P. and Michael N. Caggiano
1986 The Prevalence, Predictability, and Policy Implications of Recidivism. Santa Monica, Calif.: Rand Corporation.

Kolvin, Israel, F.J.W. Miller, M. Fleeting, and P.A. Kolvin
1988 Social and parenting factors affecting criminal-offense rates. British Journal of Psychiatry 152:80–90.

Lambert, Leah R. and Patrick G. Madden
1976 The adult female offender: The road from institution to community life. Canadian Journal of Criminology and Corrections 18:3–15.

Lane, David A.
1987 Personality and antisocial behaviour: A long-term study. Personality and Individual Differences 8:799–806.

LeBlanc, Marc and Marcel Fréchette
1989 Male Criminal Activity from Childhood Through Youth: Multilevel and Developmental Perspectives. New York: Springer–Verlag.

LeBlanc, Marc, Évelyne Vallières, and Pierre McDuff
1993 The prediction of males' adolescent and adult offending from school experience. Canadian Journal of Criminology 25:459–478.

Liberton, Michael, Mitchell Silverman, and William R. Blount
1992 Predicting probation success for the first-time offender. International Journal of Offender Therapy and Comparative Criminology 36:335–347.

Little, Gregory L. and Kenneth D. Robinson
1989 Relationship of DUI recidivism to moral reasoning, sensation seeking, and MacAndrew alcoholism scores. Psychological Reports 65:1171–1174.

Loney, Jan, Mary Anne Whaley–Klahn, Todd Kosier, and Jay Conboy
1983 Hyperactive boys and their brothers at 21: Predictors of aggressive and antisocial outcomes. In Katherine Teilmann Van Dusen and Sarnoff A. Mednick (eds.), Prospective Studies of Crime and Delinquency. Boston: Kluwer–Nijhoff.

Macnaughton-Smith, P.
1976 Permission to be Slightly Free. Ottawa, Ont., Canada: Law Reform Commission of Canada.

McCord, Joan
1982 A longitudinal view of the relationship between paternal absence and crimes. In John Gunn and David Farrington (eds.), Abnormal Offenders, Delinquency, and the Criminal Justice System. Chichester: John Wiley & Sons.
1983 A forty year perspective on effects of child abuse and neglect. Child Abuse and Neglect 7:265–270.

McGarvey, Bill, William F. Gabrielli, Peter M. Bentler, and Sarnoff A. Mednick
1981 Rearing, social class, education, and criminality: A multiple indicator model. Journal of Abnormal Psychology 90:354–364.

McGurk, B.J., N. Bolton, and M. Smith
1978 Some psychological, educational, and criminological variables related to recidivism in delinquent boys. British Journal of Social and Clinical Psychology 17:251–254.

McWilliams, William
1975 Sentencing and recidivism: An analysis by personality type. British Journal of Social Work 5:311–324.

Mednick, Birgitte R., Robert L. Baker, and Linn E. Carothers
1990 Patterns of family instability and crime: The association of timing of the family's disruption with subsequent adolescent and young adult criminality. Journal of Youth and Adolescence 19:201–219.

Mednick, Birgitte R., Charlotte Reznick, Dennis Hocevar, and Robert L. Baker
1987 Long term effects of parental divorce on young adult male crime. Journal of Youth and Adolescence 16:31–45.

Minor, Kevin I. and David J. Hartmann
1992 An evaluation of the Kalamazoo probation enhancement program. Federal Probation 56:30–35.

Mitchell, Sheila and Peter Rosa
1981 Boyhood behaviour problems as precursors of criminality: A fifteen year follow-up study. Journal of Child Psychology and Psychiatry 22:19–33.

Moffitt, Terrie E.
1987 Parental mental disorder and offspring criminal behaviour: An adoption study. Psychiatry 50:346–360.

Moffitt, Terrie E., William F. Gabrielli, Sarnoff A. Mednick, and Fini Schulsinger
1981 Socioeconomic status, IQ, and delinquency. Journal of Abnormal Psychology 90:152–156.

Motiuk, Laurence L.
1991 Antecedents and consequences of prison adjustment: A systematic assessment and reassessment approach. Unpublished doctoral dissertation, Carleton University, Ottawa, Ontario, Canada.

Motiuk, Laurence L. and James M. Bonta
1991 Prediction and matching in corrections: An examination of the risk principle in case classification. Paper presented at the Canadian Psychological Association Annual Convention (June), Calgary, Alberta, Canada.

Motiuk, Laurence L. and Shelley L. Brown
1993 Survival Time until Suspension for Sex Offenders on Conditional Release. Report No. R-31. Research and Statistics Branch. Ottawa: Correctional Service Canada.
1993 The Validity of Offender Needs Identification and Analysis in Community Correction. Report No. R-34. Research and Statistics Branch. Ottawa: Correctional Service Canada.

Motiuk, Laurence L. and Frank J. Porporino
1988 Offender Risk/Needs Assessment: A Study of Conditional Releases. Report No. R-01. Ottawa: Correctional Service of Canada.
1989 Field Test of the Community Risk/Needs Management Scale: A Study of Offenders on Case Load. Report No. R-06. Research Branch, Communications and Corporate Development. Ottawa: Solicitor General Canada/ Correctional Service Canada.

Motiuk, Laurence L., James Bonta, and D.A. Andrews
1986 Classification in correctional halfway houses: The relative and incremental predictive criterion validities of the Megargee-MMPI and LSI-R systems. Criminal Justice and Behavior 13:33–46.

O'Donnell, Clifford R. and Kathleen G. Stanley
1974 An adult furlough center: Correlates of parole success. Journal of Community Psychology 2:83–85.

Pallone, Nathaniel J. and James J. Hennessey
1977 Empirical derivation of a scale for recidivism proneness among parolees: A multivariate model. Offender Rehabilitation 2:95–110.
1977 Some correlates of recidivism among misdemeanants and minor felons: A 12-month follow-up. Journal of Social Psychology 101:321–322.

Payne, Clive, Sarah McCabe, and Nigel Walker
1974 Predicting offender-patients' reconvictions. British Journal of Psychiatry 125:60–64.

Petersén, K. Ingemar, M. Matousek, Sarnoff A. Mednick, Jan Volavka, and V. Pollack
1982 EEG antecedents of thievery. Acta Psychiatrica Scandinavica 65:331–338.

Petersilia, Joan and Susan Turner
1987 Guideline-based justice: Prediction and racial minorities. In Don M. Gottfredson and Michael Tonry (eds.), Prediction and Classification: Criminal Justice Decision Making. Chicago: University of Chicago Press.

Platt, Jerome J. and Christina Labate
1976 Recidivism in youth heroin offenders and characteristics of parole behaviour and environment. International Journal of Addictions 11:651–657.

Porporino, Frank J., Edward Zamble, and Susan Higginbottom
in press Assessing models of predicting risk for criminal recidivism. Journal of Criminal Justice.

Putniņš, Aldis L.
1982 The Esyenck Personality questionnaires and delinquency prediction. Personality and Individual Differences 3:339–340.

Quinsey, Vernon L., Manfred Pruesse, and Robert Fernley
1975 Oak Ridge patients: Prerelease characteristics and postrelease adjustment. Journal of Psychiatry and Law 3:63–77.

Rice, Marnie E. and Grant T. Harris
1991 A comparison of criminal recidivism among schizophrenic and nonschizophrenic offenders. International Journal of Law and Psychiatry 15:397–408.
1993 Psychopathy, Schizophrenia, Alcohol Abuse and Violent Recidivism. Research Report Vol. X, No. 3. Penetanguishene, Ontario, Canada: Mental Health Centre.

Rice, Marnie E., Vernon L. Quinsey, and Grant T. Harris
1991 Sexual recidivism among child molesters released from a maximum security psychiatric institution. Journal of Consulting and Clinical Psychology 59:381–386.

Rice, Marnie E., Grant T. Harris, Carol Lang, and Valerie Bell
1990 Criminological and Psychiatric Predictors of Recidivism among Male Insanity Acquittees. Research Report Vol. VII, No. 4. Penetanguishene, Ontario, Canada: Mental Health Centre.

Robins, Lee N. and Kathryn S. Ratcliffe
1980 Childhood conduct disorders and later arrest. In Lee N. Robins, Paula J. Clayton, and John Kenneth Wing (eds.), The Social Consequences of Psychiatric Illness. New York: Brunner/Mazel.

Robinson, David and Frank J. Porporino
1989 Validation of an Adult Offender Classification System for Newfoundland and Labrador. Report No. R.04. Research Branch. Ottawa: Solicitor General of Canada.

Rogers, Sally
1981 Factors Related to Recidivism among Adult Probationers in Ontario. Toronto: Ontario Ministry of Correctional Services.

Roundtree, George A., Dan W. Edwards, and Jack B. Parker
1984 A study of the personal characteristics of probationers as related to recidivism. Journal of Offender Counselling, Services, and Rehabilitation 8:53–61.

Sampson, Allan
1974 Post-prison success prediction. Criminology 12:155–173.

Serin, Ralph C.
1990 Recidivism Prediction: A Comparison of the Psychopathy Checklist and Actuarial Risk Scales. Ottawa: Correctional Service Canada.

Serin, Ralph C. and Howard E. Barbaree
1993 Decision issues in risk assessment. Federal Probation 5:22–25.

Serin, Ralph C., Ray DeV. Peters, and Howard E. Barbaree
1990 Predictors of psychopathy and release outcome in a criminal population. Psychological Assessment: A Journal of Consulting and Clinical Psychology 2:419–422.

Shannon, Lyle W.
1985 Risk assessment vs. real prediction: The prediction problem and public trust. Journal of Quantitative Criminology 1:159–185.

Stattin, Håkan and David Magnusson
1989 The role of early aggressive behaviour in the frequency, seriousness and types of later crime. Journal of Consulting and Clinical Psychology 57:710–718.

Steadman, Henry J.
1973 Follow-up on Baxstrom patients returned to hospitals for the criminally insane. American Journal of Psychiatry 130:317–319.

Steadman, Henry J. and Joseph Cocozza
1974 Careers of the Criminally Insane: Excessive Social Control of Deviance. Lexington, Mass.: D.C. Heath.

Steadman, Henry J., Joseph J. Cocozza, and Mary Evans Melick
1978 Explaining the increased arrest rate among mental patients: The changing clientele of state hospitals. American Journal of Psychiatry 135:816–820.

Tarling, Roger and John A. Perry
1985 Statistical methods in criminological prediction. In David P. Farrington and Roger Tarling (eds.), Prediction in Criminology. Albany: State University of New York Press.

Tennent, Gavin and Cynthia Way
1984 The English special hospital - A 12-17 year follow up study: A comparison of violent and nonviolent re-offenders and non-offenders. Medicine, Science, and the Law 24:81–91.

Thornberry, Terence P. and Joseph E. Jacoby
1979 The Criminally Insane: A Community Follow-up of Mentally Ill Offenders. Chicago: University of Chicago Press.

Wallander, Jan L.
1988 The relationship between attention problems in childhood and antisocial behaviour eight years later. Journal of Child Psychology and Psychiatry 29:53–61.

Werner, Eric and Ted Palmer
1976 Psychological characteristics of successful and unsuccessful parolees: Implications of heteroscedastic and non-linear relationships. Journal of Research in Crime and Delinquency 13:165–178.

West, Donald J. and David P. Farrington
1977 The Delinquent Way of Life. London: Heinemann.

Wheeler, Gerald R. and Rodney V. Hissong
1990 Transferability of probation risk-assessment instruments: A case for caution. Evaluation and Program Planning 13:399–406.

Widom, Cathy S.
1989 The cycle of violence. Science 244:160–166.

Wilbanks, William L.
1985 Predicting failure on parole. In David P. Farrington and Roger Tarling (eds.), Prediction in Criminology. Albany: State University of New York Press.

Wormith, J. Stephen and Colin S. Goldstone
1984 The clinical and statistical prediction of recidivism. Criminal Justice and Behavior 11:3–34.

Wormith, J. Stephen and Monica Ruhl
1986 Preventive detention in Canada. Journal of Interpersonal Violence 1:399–430.

Wright, Kevin N., Todd R. Clear, and Paul Dickson
1984 Universal applicability of probation risk-assessment instruments. Criminology 22:113–134.

Zamble, Edward and Frank Porporino
1990 Coping, imprisonment, and rehabilitation: Some data and their implications. Criminal Justice and Behavior 17:53–70.

Correlates of delinquency: A look at gender differences

by Linda Simourd[1] and D.A. Andrews[1]
Department of Psychology, Carleton University

Contemporary research on delinquency is challenging fundamental assumptions regarding female delinquency made by early theorists and researchers. As a result, correctional theory and research appears to be at a crossroad in terms of gender issues.

One important issue is the identification of risk factors for female youth. The apparent social bias of early female-delinquency theories suggests that exclusive reliance on personal and familial problems in the assessment of female risk may be inadequate. Yet, the relevance of male-based risk factors has not been fully evaluated.

Our study attempted to fill this gap through a systematic review of research that has examined the same risk factors for male and female youths separately. It should be noted here that our research and its findings focused on youth criminality (delinquency) rather than on adult criminality.

Female delinquency has historically been perceived as relatively rare and less serious than male delinquency.[2] Early research using police and court records showed that for each female delinquent, there were three to seven male delinquents.[3] In terms of offence type, female youths were perceived as committing relatively minor offences, such as running away, truancy and sexual acting out, while male youths were perceived as committing a much wider and more serious range of offences.[4]

Early female-delinquency theories were also dramatically different from those developed for male youths. Personal maladjustment was viewed as a fundamental cause of problem behaviour in female youths. Psychological problems, inadequate performance of the proper sex role, and a problematic home life were some of the proposed explanations for female deviance.[5] In contrast, early male-delinquency theories focused on more external risk factors, such as peer group, lower social class, and lack of educational or occupational resources.[6]

Within the past 20 years, research has challenged some of these traditional views. Contemporary self-report studies, for example, have suggested a more accurate gender ratio of no more than three male delinquents to each female delinquent.[7] Recent studies have also found that female youth are involved in a broad range of criminal behaviours, not just minor offences. One exception to the similarity of offences is the use of physical aggression; here, female involvement remains lower than that of male youths.[8]

To explain and predict delinquency, contemporary research has expanded its scope to consider personal factors (such as behaviour, personality, and cognition), interpersonal factors (such as

[1] Linda Simourd and D.A. Andrews, Department of Psychology, Carleton University, Loeb Building B-542, Ottawa, Ontario K1S 5B6.

[2] R. Canter, "Family Correlates of Male and Female Delinquency," *Criminology: An Interdisciplinary Journal*, 20 (1982): 149–167. See also P. Richards, "Quantitative and Qualitative Sex Differences in Middle-class Delinquency," *Criminology: An Interdisciplinary Journal*, 18 (1981): 453–470.

[3] R. Canter, "Sex Differences in Self-report Delinquency," *Criminology: An Interdisciplinary Journal*, 20 (1982): 373–393.

[4] Canter, Sex Differences in Self-report Delinquency.

[5] P. Giordano, "Girls, Guys and Gangs: The Changing Social Context of Female Delinquency," *The Journal of Criminal Law and Criminology*, 69 (1978): 126–132.

[6] Giordano, Girls, Guys and Gangs: The Changing Social Context of Female Delinquency.

[7] The precise ratios are 1.2:1 to 2.5:1 as reported in Canter, Sex Differences in Self-report Delinquency.

[8] I. Gomme, "Predictors of Status and Criminal Offences among Male and Female Adolescents in an Ontario Community," *Canadian Journal of Criminology*, 27 (1985): 147–159.

family and peers), and structural factors (such as school and church). Although the tendency to omit female youths continues, there has been an emerging interest in female involvement in delinquency. Increasingly, studies on risk are including male and female youths in their research samples. Until now, these studies were never examined as a group.

Our study

The primary goal of our research was to review this portion of the delinquency literature. Specifically, we looked at published and unpublished studies conducted over the past 30 years. This review gave us an opportunity to comprehensively examine female youths on a wide range of risk factors. It also allowed us to examine risk factors not typically investigated among male samples. We sought answers to two basic questions:

- What are the important risk factors for **each** gender? For example, does associating with criminal peers place male youths at risk for delinquency, and does this factor also place female youths at risk?
- Are specific risk factors **more important** for a particular gender? For example, are family problems more strongly related to female delinquency than to male delinquency, and are school difficulties more strongly associated with male delinquency than with female delinquency?

We used a technique called "meta-analysis" to conduct this review.[9] One of the main advantages of meta-analysis over the traditional (narrative) form of literature review is that meta-analysis summarizes large bodies of literature and reaches more definitive conclusions. This quantitative approach combines numerical information (actual data) from selected studies and computes an average result for each risk factor for males and females separately. This information then indicates the degree of association between delinquency and a particular risk factor.[10]

Table 1

Risk Factors

- Lower social class
- Family structure or parental problems (broken home, marital problems)
- Personal distress (anxiety, low self-esteem, apathy)
- Minor personality variables (empathy, moral reasoning)
- Poor parent–child relations (attachment, supervision)
- Educational difficulties (poor grades, dropout)
- Temperament or misconduct problems (psychopathy, impulsivity, substance use)
- Antisocial peers or attitudes

Each study included in this review met three criteria:

- Male and female youths were sampled.
- Each gender was examined on the same risk factor.
- The data for each gender were reported separately.

These criteria ensured that male–female comparisons were based on the same delinquency and risk measures, thus eliminating biases that could result from using different measures for each gender.

Sixty studies met these requirements and provided 464 correlations between delinquency and risk factors. We then grouped the risk factors examined within these correlations into eight general risk categories based on previous reviews and common themes in the delinquency literature.

Table 1 lists the eight risk factors and provides a brief indication of their content. Some factors will be familiar to FORUM readers, as a previous issue on delinquency highlighted some of these risk areas.[11]

What risk factors are most important for delinquency in each gender?

[9] For a complete discussion of meta-analysis see R. Rosenthal, "Meta-analysis: A Review," *Psychosomatic Medicine*, 53 (1991): 247–271.

[10] Measures of variability are available from Linda Simourd; however, they are not discussed in this paper.

[11] See T. Nouwens and F. Porporino, eds., *Forum on Corrections Research*, 3, 3 (1991).

Table 2

Average Correlations for Eight Risk Factors

Risk factor	Female youths *r (n)*	Male youths *r (n)*
Lower social class	0.07 (19)	0.06 (19)
Family structure or parental problems	0.07 (17)	0.09 (17)
Personal distress	0.10 (14)	0.09 (14)
Minor personality variables	0.18 (9)	0.22 (9)
Poor parent–child relations	0.20 (41)	0.22 (41)
Educational difficulties	0.24 (34)	0.23 (34)
Temperament or misconduct problems	0.35 (45)	0.36 (45)
Antisocial peers or attitudes	0.39 (53)	0.40 (53)

r = average correlation; n = number of correlations that contributed to that average

Table 2 provides a clear indication of the combined results of these 60 studies. Correlations of 0.15 or larger would be of substantive interest.

For female youths, the most important risk factors in descending order were antisocial peers or attitudes, temperament or misconduct problems, educational difficulties, poor parent–child relations, and minor personality variables. Personal distress, family structure or parental problems, and lower social class did not appear to be strongly related to delinquency.

A similar pattern emerged for the male youths: that is, the first three risk factors in the table were not strongly related to male delinquency, and the last five factors were important.[12]

Are specific risk factors more important for a particular gender?

A comparison of the two columns of correlations in Table 2 indicates there were no differences in the risk factors across gender. Statistical tests supported this seemingly obvious finding. In other words, the general risk factors that were important for male delinquency were also important for female delinquency.

Practical importance

In practical terms, these data suggest that knowing a youth's socioeconomic status or family structure would provide little information about his or her risk of delinquency.

However, information indicating difficulties in the area of family relations, conduct or peers would provide valuable information about that individual's risk of delinquency. These findings are consistent with social psychological models of criminal conduct that suggest that a variety of factors are associated with delinquency.[13]

Control variables

The next step in our research process was to assess whether particular aspects of these 60 studies contained systematic bias. For example, would the numbers in Table 2 change dramatically if we took into account the source of the information — that is, whether delinquency was measured by self-reports from the delinquents or by officials from the justice system? Would it matter if the sample consisted of high-school students or a group of offenders?

We considered 15 aspects related to the studies and their samples.[14] While certain aspects did influence the size of the correlations in Table 2, the overall ranking of factors did not change. That is, despite taking into account various study characteristics, the data on social class, family structure or parent problems and personal distress still provided little information

[12] The findings for minor personality variables should be viewed with caution, given the small number of correlations representing this factor and fluctuations noted within this category.

[13] D.A. Andrews and J. Bonta, *The Psychology of Criminal Conduct* (Cincinnati: Anderson, in press). See also M. LeBlanc, M. Ouimet and R.E. Tremblay, "An Integrative Control Theory of Delinquent Behaviour: A Validation 1976–1985," *Psychiatry*, 51 (1988): 164–176.

[14] These were 1) sample, 2) design, 3) sample size, 4) period of data collection, 5) gender focus of study, 6) predominant ethnicity, 7) source of delinquency information, 8) source of risk information, 9) retrieval source, 10) gender of author, 11) psychometrics of risk measure, 13) social class, 14) type of delinquency, and 15) nature of delinquency.

about an individual's risk for delinquency. Parent–child difficulties and school problems remained important. Finally, temperament or misconduct problems and antisocial peers or attitudes remained the factors most significantly related to delinquency, regardless of study characteristics.

To summarize, this meta-analytical review yielded three conclusions with respect to eight general risk factors:

- The general risk factors that were important for female youths were also important for male youths. Further, no risk factor was more important for a particular gender.
- The most important risk factors for both genders, in descending order, were antisocial attitudes and peers, temperament or misconduct problems, educational difficulties, poor parent–child relations, and minor personality variables. In contrast, lower social class, family structure or parental problems, and personal distress were not strongly related to delinquency.
- When various sample and study characteristics were taken into account, the general pattern of results remained the same.

The general risk factors that were important for female youths were also important for male youths. Further, no risk factor was more important for a particular gender.

Other risk factors

Although this review suggests that the same risk factors are important for men and women, some might argue that gender differences may exist for other factors not captured by the eight areas we examined.

To explore this possibility, we grouped 96 correlations not captured by the eight general risk factors into 10 additional risk areas and assessed them for possible gender differences. Table 3 lists these factors and the number of times each factor was found in the studies examined. Undoubtedly more research has been conducted within each area; however, recall that we selected the studies based on three criteria listed earlier.

Since only a total of 96 correlations pertained to these 10 categories, we treated this set of findings as exploratory. A few highlights will be mentioned.

Two very promising categories of risk were lack of attachment to convention and sexual behaviour. Lack of attachment to convention pertained to an individual's lack of affiliation with prosocial people (such as parents and teachers) or institutions (such as family, school and church). Because these measures assessed attachment to multiple individuals or institutions, they could not be incorporated in any of the earlier categories, such as education or parent–child relations.

The data indicated that lack of attachment to convention was associated with delinquency for each gender (male average correlation = 0.23; female average correlation = 0.22). This was to be expected given that our earlier analysis found that lack of attachment to specific institutions (such as educational difficulties) and people (poor parent–child relations) were important risk factors.

Although the data for sexual behaviour were obtained from only three studies, there was

Table 3

Other Risk Factors

- Victimization (1)
- Illegitimate opportunity (2)
- Lack of legitimate opportunity (3)
- Sexual behaviour (3)
- Lack of hobbies or involvement (4)
- Accommodation problems (4)
- Self-concept issues (6)
- Race (7)
- Sex-role orientation (8)
- Lack of attachment to convention (10)

consensus about the importance of this factor. The lowest correlations for male and female youths were 0.35 and 0.26, respectively. What remains unclear is whether gender differences exist — one study found large gender differences, another found minor differences, and the third found none.

The significance of sexual behaviour as a correlate of delinquency may be somewhat surprising, given current rates of sexual activity among teenagers. However, these three studies were conducted between 1966 and 1971. Sexual behaviour in those days may have served as an indicator of an individual's tolerance of rules or norms. An interesting question is whether the same degree of association would be noted for high-school students in the 1990s.

Two less promising categories were race and sex-role orientation. The issue of race refers solely to Black versus Caucasian samples, as no other racial comparisons were reported. The information from seven studies suggests no link between race and delinquency (male average correlation = - 0.02; female average correlation = 0.03).

Sex-role orientation (masculinity or femininity) was also found to be unrelated to delinquency (male average correlation = 0.05; female average correlation = 0.07). There may be a minor association of delinquency with sex-role orientation when lack of feminine traits (communal traits, such as sharing and caring) are examined. A narrative literature review of research in this area found similar results and concluded this line of research should be abandoned.[15]

The jury is still out for the six remaining risk categories. For these factors (victimization, illegitimate opportunity, lack of legitimate opportunity, lack of hobbies or involvement, accomodation problems, and self-concept issues), either too few studies were obtained or various study characteristics made interpretation of the findings difficult. Victimization, for example, was only measured by one study, which asked students between the ages of 12 and 16 whether they personally were a victim of five types of crimes ranging from theft of unguarded possessions to attack and assault.[16] The correlations, based on overall victimization, not on victimization specific to violence, yielded no gender difference (male correlation = 0.14; female correlation = 0.16). Given the growing clinical interest in sexual abuse and victimization, particularly for females,[17] future gender research should examine this issue empirically.

Another category, accommodation problems (e.g., crowding, high-crime neighbourhood), was assessed by four studies, whose results varied across sample and design, thus making interpretation difficult. The two cross-sectional studies indicated no association between accomodation problems and delinquency in male youths, but a minor association between them for female youths. The longitudinal study reported a significant association for men but not for women. Finally, the offender versus non-offender sample found accommodation problems to be a significant risk factor for both genders, although more important for women.

More questions are raised than answered when variation occurs to this extent in the data. Clearly, for these six categories, more research must be gathered before any conclusions can be drawn.

In summary, our conclusions about the 10 categories described in this section are limited by the small number of correlations and the sometimes varying results. Future research could examine some of these factors for their use as risk factors and for possible gender differences.

[15] N. Naffin, "The Masculinity–Femininity Hypothesis: A Consideration of Gender-based Personality Theories of Female Crime," *British Journal of Criminology*, 25 (1985): 365–381.

[16] R. Mawby, "Sex and Crime: The Results of a Self-report Study," *British Journal of Sociology*, 31 (1980): 525–543.

[17] J. Darke and H. McLean, "Invisible Women: The Treatment and Research Needs of Female Offenders," *The Treatment of Female Offenders*, ed. A. Loucks, Symposium conducted at the annual meeting of the Canadian Psychological Association (Ottawa: May 1990). See also A. Einbender and W. Friedrich, "Psychological Functioning and Behaviour of Sexually Abused Girls," *Journal of Consulting and Clinical Psychology*, 57 (1989): 155–157.

Conclusions

The results of this literature review are clear. The risk factors that are important for male delinquency are also important for female delinquency. Of the risk factors examined, the most important are antisocial peers or attitudes, temperament or misconduct problems, educational difficulties, poor parent–child relations, and minor personality variables. In contrast, lower social class, family structure or parental problems, and personal distress are not strongly related to delinquency for either gender.

These results support recent social psychological models of criminal conduct that suggest a variety of personal, interpersonal and structural factors are related to delinquent behaviour in males and females.

However, our results seriously challenge the value of early delinquency theories. Most importantly, notions of female delinquency as exclusively symptomatic of personal distress or familial difficulties have been shown to be inadequate. Early male theories, which focused on lower social class as a major route to criminal behaviour, can also be questioned. Others previously challenged this social-location perspective, and a contemporary understanding of male delinquency has moved beyond this limited view.[15]

Several challenges remain for researchers and scientist-practitioners. First, these results do not eliminate the possibility that other factors are associated with delinquency in one or both genders. Future research could investigate, for example, the role of some of the less frequently measured factors explored in this study. Weaker personality variables, sexual behaviour, and sexual abuse and victimization are a few areas worthy of further consideration.

Second, these findings should lead to reformulated ideas and directions about theory and research on female delinquency. Those interested in female delinquency should learn from male-delinquency literature that has progressed beyond the early reliance on social class as the major explanation of male deviance. It is time to set aside antiquated ideas and to consider a larger group of factors as causes of delinquency. Our review examined gender differences and similarities in the correlates of delinquency, so the next step is to examine causal factors for individual female and male youths in the context of their experiences. That research would inform us of the need for gender-specific or gender-neutral theories of delinquency.

In conclusion, for some, the results of this review may simply state the obvious. For years, a small portion of the literature on delinquency has examined the same risk factors for male and female youths and independently, yet repeatedly, found the same results across gender. This review of the literature has pulled those results together in a quantitative fashion, and the similarity across gender can no longer be ignored. The factors examined to date suggest a unique set of correlates may not be required for female delinquency.

Future research will inform us about the role gender plays with respect to predictors and theories of criminal conduct. Consistent with this approach is the need to incorporate gender into the research design and to present the male and female data separately. Only then can a clearer understanding of the similarities and differences across gender be obtained. ■

Coming up in *Forum on Corrections Research . . .*

The theme of the next issue of FORUM is "Special Needs Offenders"

[15] This conclusion is consistent with that of a study conducted by C. Tittle and R. Meir, "Specifying the SES/Delinquency Relationship," *Criminology: An Interdisciplinary Journal*, 28 (1990): 271–296.

Best Practices of Classification and Assessment

by Edward J. Latessa*

The Evolution of Classification

For the purposes of this article, "risk" refers to the probability that an offender will reoffend. Thus, high-risk offenders have a greater probability of reoffending than do low-risk offenders. How offender risk is determined is, thus, very important, because it can affect public protection and the way and manner in which offenders are supervised in—or whether they are even released into—the community.

"Gut Feelings." The prehistory of risk assessment in criminal justice refers to the use of "gut feelings" to make decisions about the risk an offender presents. With this process, information is collected about the offender, usually through an interview or file review. The information is then reviewed and a general assessment or global prediction is made: "In my professional opinion. . . ." The problems with this method are considerable and have been delineated by Wong (1997) and Kennedy (1998), who find that:

- Predictions are subject to personal bias;
- Predictions are subjective and often unsubstantiated;
- Decision rules are not observed;
- The process can lead to bias decisions;
- It is difficult to distinguish levels of risk; and
- Information is overlooked or overemphasized.

The First Generation of Classification: The "Burgess Scale." The first generation of formal classification instruments was pioneered by Bruce et al. in 1928. The development of this standardized and objective instrument was brought about by the request of the Illinois Parole Board, which wanted to make more informed decisions about whom to release on parole. Bruce and his colleagues reviewed the records of nearly 6,000 inmates. The so-called "Burgess scale" included items such as criminal type (first timer, occasional, habitual, professional), social type (farm boy, gangster, hobo, ne'er-do well, drunkard), age when paroled, and other static factors. Although many of these categories seem out-of-date today, the Burgess scale was one of the first attempts to develop an actuarial instrument to predict offender risk. There are several advantages and disadvantages to this approach (Kennedy, 1998; Wong, 1997). The advantages are that the categories:

- Are objective and accountable;
- Cover important historical risk factors;
- Are easy to use and are reliable; and
- Distinguish levels of risk of reoffending.

The disadvantages are that they:

- Consist primarily of static predictors;
- Do not identify target behaviors; and
- Are not capable of measuring change in the offender.

The Second Generation of Classification: The CMC. The second generation of risk prediction recognized that risk is more than simply static predictors. The best example can be seen in the Wisconsin Client Management Classification System. First developed and used in Wisconsin in 1975, the Client Management Classification System (CMC) is designed to help identify the level of surveillance needed for each case, as well as to determine the needs of the offender and the resources necessary to meet them. With adequate classification, limited resources can be concentrated on the most critical cases—those of high risk (Wright et al., 1984). Following Wisconsin's development of the CMC, the National Institute of Corrections (1983) adopted it as a model system and began advocating and supporting its use throughout the country. It has been proven satisfactory in many jurisdictions, including Austin, Texas (Harris, 1994).

The foundation of the system is a risk/needs assessment instrument that is completed for each probationer at regular intervals. Cases are classified into high, medium, or low risk/needs. These ratings are, in turn, used to determine the level of supervision required for each case.

Once an offender is classified into a risk/needs level, a profiling interview makes a more detailed assessment that helps to determine what the relationship should be between the officer and the offender. This element of the system is called the Client Management Classification System, and it consists of four unique treatment modalities:

- **Selective Intervention.** This group is designed for offenders who enjoy relatively stable and prosocial lifestyles (e.g., employed, established in community, and minimal criminal records). Such offenders have typically experienced an isolated and stressful event or neurotic problem. With effective intervention, there is a higher chance of avoiding future difficulty. The goals of treatment for these individuals include the development of appropriate responses to temporary crises and problems and the reestablishment of pro-life patterns.
- **Environmental Structure.** Offenders in this group are predominantly characterized by deficiencies in social, vocational, and intellectual skills. Most of their problems stem from their inability to succeed in their employment or to be comfortable in most social settings and from their overall lack of social skills and intellectual cultivation/ability. The goals for these persons include: (a) developing

See *PRACTICES*, next page

The CMC component is time consuming to administer and the scoring is somewhat involved. In practice, many probation departments that use this instrument rely more heavily on the risk component, which consists of mainly static predictors.

Edward J. Latessa is professor and head of the Division of Criminal Justice at the University of Cincinnati, Cincinnati, OH. He can be reached at Division of Criminal Justice, 600 Dyer Hall, P.O. Box 210389, Cincinnati, Ohio 45221-0389; phone: (513) 556-5827; email: Edward.Latessa@uc.edu.

PRACTICES, from page 4

basic employment and social skills; (b) selecting alternatives to association with criminally oriented peers; and (c) improving social skills and impulse control.

- **Casework Control.** Offenders in this group manifest instabilities in their lives as evidenced by failures in employment and domestic problems. A lack of goal-directedness is present, typically associated with alcohol and drug problems. Offense patterns include numerous arrests, although marketable job skills are present. Unstable childhoods, family pressure, and financial difficulties are typically present. The goals appropriate for offenders in this group include promoting stability in their professional and domestic endeavors and achieving an improved utilization of the individual's potential, along with an elimination of self-defeating behavior and emotional/psychological problems.
- **Limit Setting.** Offenders in this group are commonly considered to be successful career criminals because of their long-term involvement in criminal activities. They generally enjoy "beating the system"; they frequently act for material gain; and they show little remorse or guilt. Because of their value system, they easily adapt to prison environments and return to crime upon release. Goals for this group are problematic, but they include changing the offender's basic attitudes and closely supervising his or her behavior within the community.

The information for the CMC is based on a structured interview with the offender. After a case has been classified, an individual treatment plan is developed. Results from the CMC have found that approximately 40% of probation caseloads are assigned to Selective Intervention, 15% to Environmental Structure, 30% to Casework Control, and 15% to Limit Setting.

Despite the advantages of the CMC, there are several shortcomings. One is that risk and needs are separately assessed and not fully integrated. Another is that the CMC component is time consuming to administer and the scoring is somewhat involved. In practice, many probation departments that use this instrument rely more heavily on the risk component, which consists of mainly static predictors.

Recent Classification Instruments: The LSI-R. The latest generation of classification instruments has successfully combined risk and needs and is relatively easy to use. One example is the Level of Service Inventory-Revised (LSI-R) designed by Andrews and Bonta (1995). The LSI-R is based on social learning theory and has been extensively tested and validated across North America. The LSI-R consists of 54 items in 10 areas. These areas are:

- Criminal history;
- Education and employment;
- Financial;
- Family and marital;
- Accommodation;
- Leisure/Recreation;
- Companions;
- Alcohol/Drug Problem;
- Emotional/Personal; and
- Attitudes/Orientation.

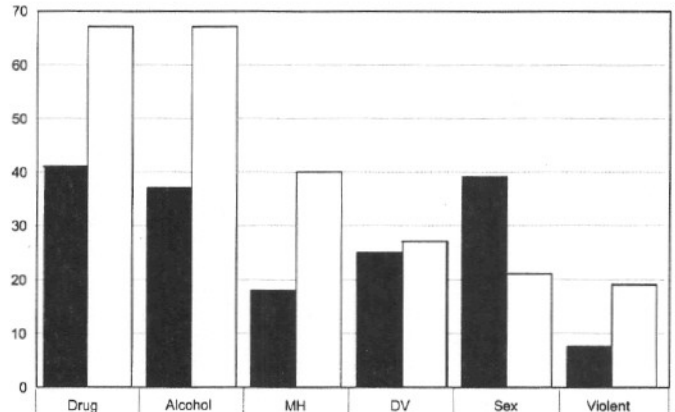
A leading advantage of actuarial risk and need assessment tools is that they are standardized and objective and help distinguish levels of risk or need. Because they are based on statistical studies, they also reduce bias and false positive and false negative rates.

Information is collected primarily through a structured interview process. The LSI-R has been found to be one of the most valid instruments for predicting recidivism. For example, a recent study compared the LSI-R to the Psychopathy Checklist-Revised (PCL-R), an actuarial instrument developed by Hare (1996) and widely used in the United States and Canada to classify and assess psychopaths. Gendreau and his colleagues (2001) found that the LSI-R surpassed the PLC-R in its ability to predict both general recidivism ($r = 0.38$ vs. 0.23) and violent recidivism ($r = 0.26$ vs. 0.22). There is also a juvenile version of the LSI-R called the Youthful Level of Service/Case Management Inventory (Y-LSI; Hoge & Andrews, 1996).

Specialized Classification Systems: The OPI. There are also classification systems designed for certain types of offenders or need areas, such as the mentally disordered, sex offenders, or substance abusers. Some of these systems help to classify cases and recommend levels of intervention. One example of the latter is the Offender Profile Index (OPI) developed by the National Association of State Alcohol and Drug Abuse Directors (Inciardi et al., 1993). The OPI is a broad classification instrument useful for determining which type of drug abuse treatment intervention is most appropriate.

See *PRACTICES*, next page

Figure 1: Use of Specialized Assessment Tools by Probation, Parole, and Community Corrections Service Providers



	Drug	Alcohol	MH	DV	Sex	Violent
Probation/Parole	41	37	18	25	39	7.5
CC Providers	67	67	40	27	21	19

Source: D. Hubbard, L.F. Travis & E.J. Latessa, 2001. Case Classification in Community Corrections: A National Survey of the State of the Art.

PRACTICES, from page 5

Five profiles are produced by the OPI:

- Long-term residential treatment;
- Short-term residential treatment;
- Intensive outpatient treatment;
- Regular outpatient treatment; and
- Urine monitoring only.

The OPI is not designed to replace more comprehensive and clinically oriented treatment plans and case studies. However, for probation and parole agencies with large numbers of cases, the OPI can be a valuable tool in case management and appropriate treatment referral. Figure 1 shows the use of specialized assessment tools across the United States.

Advantages, Criticisms, and Use of Assessment Tools

Advantages of Assessment Tools. One of the leading advantages of actuarial risk and need assessment tools is that they are standardized and objective and help distinguish levels of risk or need (e.g. high, medium, low). Because they are based on statistical studies, they also reduce bias and false positive and false negative rates (Holsinger et al., 2001). In a recent study of sex offender assessment, several actuarial instruments were compared to clinical assessment (Hanson & Bussiere, 1998). Several different outcome measures were examined, and as is apparent in Figure 2, when pitted against the best actuarial tool, clinical assessment fell short in terms of predictive ability. As can be seen, the differences are not even close.

In a recent national survey of probation and parole agencies concerning the use and practices surrounding class classification, Hubbard et al. (2001) found that the vast majority of agencies reported using some actuarial instrument to assess and classify offenders. A summary of their findings is presented below:

- Almost 75% of the probation and parole agencies and about 56% of the community corrections service providers reported that they classify using standardized and objective instruments;
- Large agencies were more likely to classify than small agencies;
- More than 83% of the respondents reported that it was "absolutely" or "very necessary" to classify on risk, and 66% reported it was necessary to classify on needs;
- The most widely used instrument was the Wisconsin Risk and Need instrument, followed by the LSI-R.

A recent national survey of probation and parole agencies found that the vast majority of agencies reported using some actuarial instrument to assess and classify offenders.

- Nearly all respondents agreed that case classification makes their job easier, benefits the offender, creates a more professional environment, helps staff make better decisions, increases effectiveness of service delivery, and enhances fairness in decision-making;
- The most common use of these tools was to address officer workloads (75%), staff deployment (54%), development of specialized caseloads (47%), and sentencing decisions (20%);
- Nearly 80% of the agencies reported using the various instruments to reassess offenders.

Criticisms of Assessment Tools. Offender classification is not without its critics. Some argue that the instruments are nothing more than "educated guesses" (Smykla, 1986); others are more concerned about their proper use and accuracy (Greenwood & Zimring, 1985; Silver & Miller, 2001; Wilbanks, 1985). Another leading concern centers on the use of a risk instrument in one jurisdiction that has been developed and validated in another. Just because a risk instrument is accurate in one jurisdiction does not necessarily mean it will be effective in predicting outcome in another (Collins, 1990; Kratcoski, 1985; Sigler & Williams, 1994; Wright et al., 1984). As Travis (1989) has stated: "Ideally, a risk classification device should be constructed based on the population on which it is to be used."

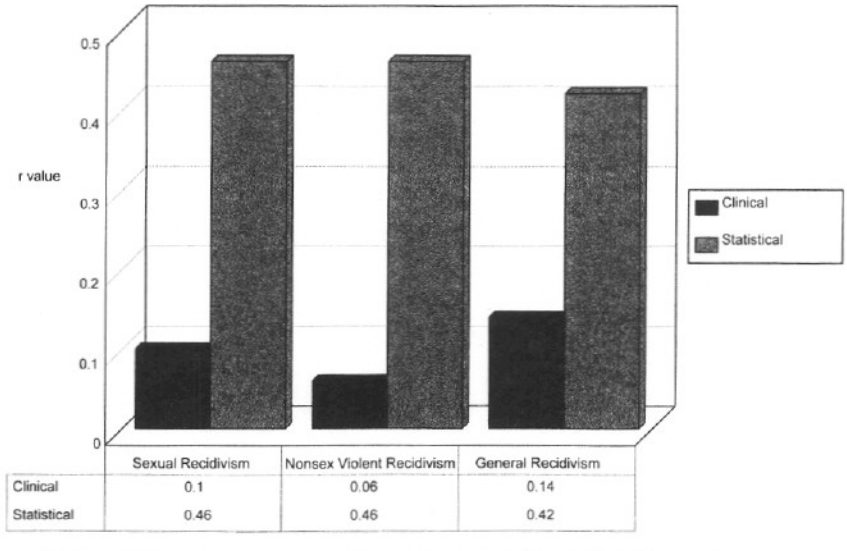
The Use of Assessment Tools. Clear (1988) maintains that the implementation of these prediction instruments has two main advantages: First, they improve the reliability of decisions made about offenders; in a sense, they make correctional officials more predictable. Second, they provide a basis on which corrections personnel can publicly justify both individual decisions and decision-making policies. In both cases, the advantage is grounded in the powerful appearance of "scientific" decision-making.

There are a number of reasons that the classification and assessment of offenders are important. Among these are that they help to:

- Guide and structure decision-making;
- Reduce bias;
- Improve the placement of offenders for treatment and public safety;

See *PRACTICES*, page 27

Figure 2: Comparison of Clinical versus Statistical Prediction of Sex Offenders



	Sexual Recidivism	Nonsex Violent Recidivism	General Recidivism
Clinical	0.1	0.06	0.14
Statistical	0.46	0.46	0.42

Source: R.K. Hanson & M.T. Bussiere, 1998. Predicting Relapse: A Meta Analysis of Sexual Offender Recidivism Studies. *Journal of Consulting and Clinical Psychology, 66*, 348-362.

PRACTICES, from page 6

- Manage offenders in a more effective manner;
- Mount legal challenges; and
- Utilize resources more effectively.

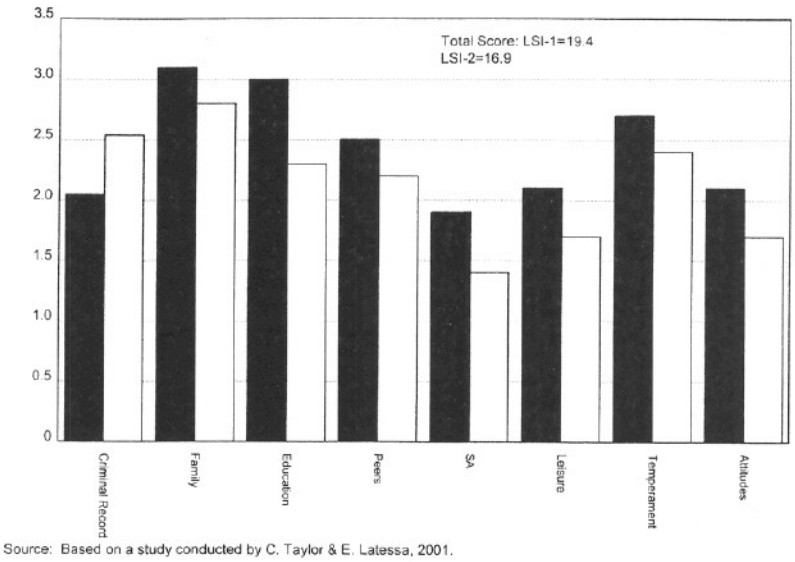
In addition to the above, another advantage of using assessment tools based on dynamic factors is the ability to reassess the offender and to determine whether or not there has been a reduction in risk score. This allows an agency to move beyond risk management to risk reduction—the ultimate goal of community corrections. Figure 3 illustrates the initial assessment and reassessment scores from a sample of youth supervised on probation. As can be seen, these data can help a probation department better focus its resources and strategies.

Another advantage of using assessment tools based on dynamic factors is the ability to reassess the offender and to determine whether or not there has been a reduction in risk score.

Another example is demonstrated in Figure 4, which shows the results from the reassessment of offenders sentenced to an Ohio community-based correctional facility. The purpose of such a facility is to provide up to six months of secure, structured treatment to felony offenders who would otherwise be incarcerated in a prison.

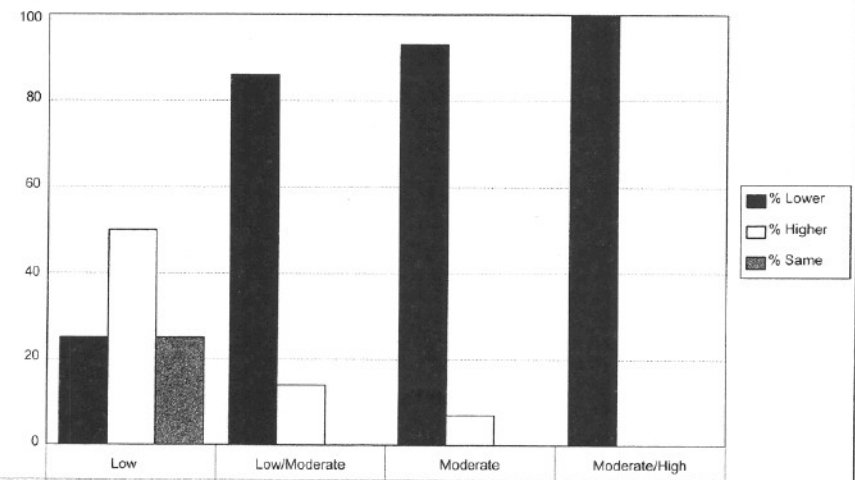
The results from the Ohio study show that the greatest reduction in risk scores was for the highest-risk offenders, whereas low-risk offenders actually saw their risk scores increase. These data demonstrate the risk principle, which states that intensive treatment services should be reserved for high-risk offenders. When low-risk offenders are placed in an intensive intervention program, the outcome is often detrimental to them. This occurs for two reasons. The first may be that the high-risk offenders have a negative influence on the low-risk, more prosocial individuals. The second probably results from the disruption of prosocial networks and other social support mechanisms that low-risk offenders usually possess (or they would not be low risk). For example, placement in a program such as the one described above usually results in loss of employment and disruption to the family.

Figure 3: Assessment and Reassessment Risk/Need Scores From 84 Youth Based on the Youthful Level of Service Inventory



Source: Based on a study conducted by C. Taylor & E. Latessa, 2001.

Figure 4: LSI Risk Reduction by Group for Ohio Community Correctional Facilities



	Low	Low/Moderate	Moderate	Moderate/High
% Lower	25	86	93	100
% Higher	50	14	7	0
% Same	25	0	0	0

Principles of Offender Classification

Andrews et al. (1990) have identified four principles of effective classification:

- **Risk:** Predicting future criminal behavior and matching levels of treatment/services to the risk level of the offender;
- **Need:** Matching offenders to programs that address their criminogenic needs;
- **Responsivity:** Delivering intervention in a style and mode that is consistent with the ability and learning style of the offender and recognizing that individuals may be more responsive to certain staff; and
- **Professional discretion:** Having considered risk, need, and responsivity, making decisions as are appropriate under existing conditions.

Through the work of a number of researchers, our understanding of classifica-

See *PRACTICES*, next page

PRACTICES, from page 27

tion and assessment, and of the important role it plays in community corrections, is becoming more apparent (Andrews, 1983, 1989; Bonta & Montiuk, 1985; Gendreau et al., 1996; Jones, 1996; Kennedy & Serin, 1997).

The latest generation of classification instruments provides the probation or parole department with an effective and fairly simple means of classifying and managing offenders. It is important to remember, however, that although instruments such as the CMC or LSI-R can be important and useful tools in assisting the community correctional agency and the supervising officer in case management, they will not solve all of the problems faced by probation and parole agencies, and they will not fully replace the sound judgment and experience of well-trained probation and parole officers (Klein, 1989; Schumacher, 1985).

Standards of Classification

Travis and Latessa (1996) have identified 10 elements of effective classification and assessment. They are:

- **Purpose.** Generally, the purpose of classification and assessment is to insure that offenders are treated differentially within a system so as to insure safety, adequate treatment, and understanding.
- **Organizational fit.** Organizations and agencies have different characteristics, capabilities, and needs.
- **Accuracy.** How well does the instrument correctly assess outcome? Is the offender correctly placed within the system? Reliability and validity are the key elements to accuracy. Glick et al. (1998, p. 73) explain reliability and validity "as hitting the same spot on a bull's eye all the time. If your system is reliable but not valid, you may be hitting the target consistently, but not the right spot."
- **Parsimony.** Parsimony refers to the ease of use, the economy of composition, and the achievement of accuracy with the least number of factors. In other words, short and simple.
- **Distribution.** How well does the system disperse cases across classification groups? If all offenders fall into the same group, there is little distribution.
- **Dynamism.** Dynamism is the instrument measuring dynamic risk factors that are amenable to change. Dynamic factors also allow for the measurement of progress and change in the offender and aid in reclassification.
- **Utility.** To be effective, classification systems must be useful. This means that the staff achieve the purposes of classification and the goals of the agency.
- **Practicality.** Closely related to utility is the practical aspect of classification. The system must be practical and possible to implement. A process that is 100% accurate but impossible to apply in an agency does not help that agency. Similarly, a system that is easy to use but does not lead to better decisions is of no value.
- **Justice.** An effective classification and assessment process should produce just outcomes. Offender placement and service provision should be based upon offender differences that are real and measurable and should yield consistent outcomes, regardless of subjective impressions.
- **Sensitivity.** Sensitivity is really a goal of the classification process. If all elements are met, the most effective classification and assessment process is sensitive to the differences of offenders. At the highest level, this would mean individualizing case planning.

Although the debate will likely continue, it appears that instruments such as the LSI-R can indeed be used to assess and classify both male and female offenders.

Classification and Female Offenders

Several scholars have questioned the notion that the risk factors used to predict antisocial behavior for male offenders are similar to those needed for female offenders (Chesney-Lind, 1989, 1997; Fund, 1999; Mazerolle, 1998). The neglect of female offenders has consistently been criticized in areas of criminological and criminal justice research ranging from theory development to the development of correctional interventions (Belknap & Holsinger, 1998; Chesney-Lind & Shelden, 1992; Funk, 1999). Furthermore, the lack of instruments that discriminate between males and females has been a common criticism of current risk/needs assessment efforts (Funk, 1999). The basis for this criticism is twofold:

- The factors involved in risk assessment for females may differ from those for males; and
- The risk factors may be similar for females and males, but exposure to these factors may present different challenges for female and male offenders (Chesney-Lind, 1989; Funk, 1999; Gilligan & Wiggins, 1988).[

There is no question that there has been considerably less research conducted on female offenders than on male offenders. However, several studies that have examined risk factors and gender have found that instruments such as the LSI-R can be useful in assessing and classifying female offenders (Andrews, 1982; Bonta & Motiuk, 1985; Coulson et al., 1996; Hoge & Andrews, 1996; Motiuk, 1993; Shields & Simourd, 1991). In a recent study examining risk prediction for male and female offenders, Lowenkamp et al. (2001) added to this research by looking at 317 males and 125 females. They found that the LSI-R was a valid predictive instrument for female offenders. They also found that a history of prior abuse (sexual or physical), although more prevalent in female offenders, was not correlated with outcome. Although the debate will likely continue, it appears that the evidence is mounting that instruments such as the LSI-R can indeed be used to assess and classify both male and female offenders.

Conclusions

In conclusion, several points can be made with regard to offender assessment:

- **First,** there is not a "one size fits all" strategy for offender assessment. Once a general risk/needs assessment has been completed, it is often necessary to conduct secondary assessments on specific target areas (i.e., substance abuse or mental health).
- **Second,** assessment is not a "one-time" event. Offender risk and need factors change, and it is thus important to consider assessment as a process rather than an event. Reassessing offenders helps promote public safety.
- **Third,** offender assessments based on standardized and objective factors are more reliable, easier to use, less time consuming, and less expensive than clinical approaches.

See *PRACTICES*, next page

PRACTICES, from page 28

- **Fourth,** staff training is vital if assessment is going to achieve its full potential. Along with training is the importance of quality assurance mechanisms to monitor the use and application of assessment tools and processes.

Finally, it is important to remember that assessment involves making decisions. Although instruments give guidance and information, it is people who decide what to do.

References

Andrews, D.A. (1982). *The Level of Services Inventory (LSI): The First Follow-Up*. Toronto: Ontario Ministry of Correctional Services.

Andrews, D.A. (1983). The assessment of outcome in correctional samples. In M.K. Lambert, E.R. Christensen, & S.S. DeJulio (Eds.). *The Measurement of Psychotherapy Outcome in Research and Evaluation*. New York: Wiley.

Andrews, D.A. (1989). Recidivism is predictable and can be influenced: Using risk assessments to reduce recidivism. *Forum on Correctional Research, 1(2)*, 11–17.

Andrews, D.A., & Bonta, J. (1995). *LSI-R: The Level of Service Inventory—Revised*. Toronto: Multi-Health Systems, Inc.

Andrews, D.A., Bonta, J., & Hoge, R.D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior, 17*, 19–52.

Belknap, J., & Holsinger, K. (1998). An overview of delinquent girls: How theory and practice have failed and the need for innovative changes. In R.T. Zaplin (Ed.). *Female Crime and Delinquency: Critical Perspectives and Effective Interventions*. Gaithersburg, MD: Aspen, pp. 31–64.

Bonta, J., & Motiuk, L.L. (1985). Utilization of an interview-based classification instrument: A study of correctional halfway houses. *Criminal Justice and Behavior, 12*, 333–352.

Bruce, A.A., Harno, A.J., Burgess, E.W., & Landesco, J. (1928). *The Workings of the Intermediate-Sentence Law and the Parole System in Illinois*. Springfield, IL: State of Illinois, Board of Parole.

Chesney-Lind, M. (1989). Girls' crime and women's place: Toward a feminist model of female delinquency. *Crime and Delinquency, 35*, 5–29.

Chesney-Lind, M. (1997). *The Female Offender*. Thousand Oaks, CA: Sage.

Chesney-Lind, M., & Sheldon, R. (1998). *Girls, Delinquency, and Juvenile Justice* (2nd ed.). Belmont, CA: Wadsworth.

Clear, T.R. (1988). Statistical prediction in corrections. *Research in Corrections, 1*, 1–39.

Collins, P.M. (1990). Risk classification and assessment in probation: A study of misdemeanants. Master's thesis, University of Cincinnati, Cincinnati, OH.

Funk, S.J. (1999). Risk assessment for juveniles on probation. *Criminal Justice and Behavior, 26*, 44–68.

Gendreau, P., Goggin, C., & Little, T. (1996). *Predicting Adult Offender Recidivism: What Works?* Ottawa: Solicitor General Canada.

Gendreau, P., Goggin, C., & Smith, P. (2001). PCL-R: The "unparalleled" measure of offender risk? *Criminal Justice and Behavior* (forthcoming).

Gilligan C., & Wiggins, G. (1988). The origins of morality in early childhood relationships. In C. Gilligan, J. Ward & J. Taylor (Eds.). *Mapping the Moral Domain: A Contribution of Women's Thinking to Psychological Theory and Education*. Cambridge, MA: Harvard University Press, pp. 111–138.

Glick, B., Sturgeon, W., & Venator-Santiago, C.R. (1998). *No Time to Play: Youthful Offenders in Adult Correctional System*. Lantham, MD: American Correctional Association.

Greenwood, P.W., & Zimring, F.E. (1985). *One More Chance: The Pursuit of Promising Intervention Strategies for Chronic Juvenile Offenders*. Santa Monica, CA: Rand Corporation.

Hanson, R.K., & Bussiere, M.T. (1998). Predicting relapse: A meta analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*, 348–336.

Hare, R.D. (1996). Psychopathy: A clinical construct whose time has come. *Criminal Justice and Behavior, 23*, 25–54.

Harris, P. (1994). Client management classification and prediction of probation outcome. *Crime & Delinquency, 40*, 154–174.

Hoge, R.D., & Andrews, D.A. (1996). *The Youthful Level of Service/Case Management Inventory: Description and Evaluation*. Paper presented at the Annual Conference of the American Psychological Association in Toronto, August 1996.

Holsinger, A.M., Lurigio, A.J., & Latessa, E.J. (2001). Practitioner's guide to understanding the basis of assessing offender risk. *Federal Probation (June)*, 46–50.

Hubbard, J.D., Travis, L.F., & Latessa, E.J. (2001). *Case Classification in Community Corrections: A National Survey of the State of the Art*. Washington, DC: National Institute of Justice, U.S. Department of Justice.

Inciardi, J.A., McBride, D.C., & Weinman, B.A. (1993). *The Offender Profile Index: A User's Guide*. Washington, DC: National Association of State Alcohol and Drug Abuse Directors.

Jones, P.R. (1996). Risk prediction in criminal justice. In A.T. Harlan (Ed.). *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply*. Thousand Oaks, CA: Sage, pp. 33–68.

Kennedy, S. (1998). *Effective Interventions with Higher Risk Offenders*. Longmont, CO: National Institute of Corrections.

Kennedy, S., & Serin, R. (1997). Treatment responsivity: Contributing to effective correctional programming. *ICCA Journal on Community Corrections, 7(4)*, 46–52.

Klein, A.R. (1989). The curse of caseload management. *Perspectives, 13 (Winter)*, 27–28.

Kratcoski, P.C. (1985). The functions of classification models in probation and parole: Control or treatment-rehabilitation? *Federal Probation, 49(4)*, 49–56.

Lowenkamp, C.T., Holsinger, A.M., & Latessa, E.J. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior, 28*, 543–563.

Mazerolle, P. (1998). Gender, general strain, and delinquency: An empirical examination. *Justice Quarterly, 15*, 65–91.

Motiuk, L.L. (1993). Where are we in our ability to assess risk? *Forum on Correctional Research, 5(1)*, 14–18.

National Institute of Corrections. (1983) *Model Classification System*, Washington, DC: Author.

Schumacher, M.A. (1985). Implementation of a client classification and case management system: A practitioner's view. *Crime & Delinquency, 31*, 445–455.

Shields, I.W., & Simourd, D.J. (1991). Predicting predatory behavior in a population of incarcerated young offenders. *Criminal Justice and Behavior, 18*, 180–194.

Sigler, R., & Williams, J. (1994). A study of the outcomes of probation officers and risk-screening instrument classifications. *Journal of Criminal Justice, 22*, 495–502.

Silver, E., & Miller, L.L. (2001). A cautionary note on the use of actuarial risk assessment tools for social control. *Crime and Delinquency, 48*, 138–161.

Smykla, J.O. (1986). Critique concerning prediction in probation and parole: Some alternative suggestions. *International Journal of Offender Therapy and Comparative Criminology, 30-31*, 125–139.

Travis, L.F. (1989). *Risk Classification in Probation and Parole*. Risk Classification Project, University of Cincinnati, Cincinnati, OH.

Travis, L.F., & Latessa, E.J. (1996). Classification and needs assessment module. In *Managing Violent Youthful Offenders in Adult Institutions Curriculum*. Longmont, CO: National Institute of Corrections.

Wilbanks, W.L. (1985). Predicting failure on parole. In D. Farrington & R. Tarling (Eds.). *Prediction in Criminology*. Albany, NY: State University of New York Press, pp. 78–94.

Wong, S. (1997). Risk assessing the risk of violent recidivism. Presentation at the American Probation and Parole Association meeting in Boston, MA.

Wright, K.N., Clear, T.R., & Dickson, P. (1984). Universal applicability of probation risk assessment instruments. *Criminology, 22*, 113–134. ■

Finally, it is important to remember that assessment involves making decisions. Although instruments give guidance and information, it is people who decide what to do.

SUBSCRIPTION INFORMATION

Journal of Community Corrections is published quarterly. A basic one-year subscription to **Journal of Community Corrections** is $159 plus postage and handling. Non-exempt New Jersey and New York residents please add appropriate sales tax.

TO ORDER
Complete the information below and mail to: Civic Research Institute 4478 U.S. Route 27 • P.O. Box 585 Kingston, NJ 08528. Or e-mail your order to: order@civicresearchinstitute.com

- ❑ Enter my one-year subscription to **Journal of Community Corrections** at $159 plus $10.95 postage and handling.
- ❑ Enter my one-year subscription to **Correctional Law Reporter** at $159 plus $10.95 postage and handling.
- ❑ Enter my order for **Sex Offender Registration and Community Notification: A "Megan's Law" Sourcebook** by Karen J.Terry, Ph.D., and Jack S. Furlong, J.D. at $287 plus $12.50 shipping and handling.
- ❑ Enter my order for **Forensic Mental Health** by Gerald Landsberg and Amy Smiley at $125 plus $8.95 shipping and handling.

Name

Title

Agency

Address

City

State Zip Code

Phone Number

E-Mail Address

Purchase Order #

Missing or damaged issues?
Call Customer Service at 609-683-4450.

RESPONSIVITY, from page 26

Prochaska, J.O., & DiClemente, C.C. (1992). Stages of change in the modification of problem behaviors. *Progress in Behavior Modification, 28*, 183–218.

Prochaska, J.O., DiClemente, C.C., & Norcross, J.C. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist, 47(9)*, 1102–1114.

Schofield, W. (1964). *Psychotherapy, The Purchase of Friendship*. Englewood Cliffs, NJ: Prentice-Hall.

Serin, R., & Kennedy, S. (1997). Treatment readiness and responsivity: Contributing to effective correctional programming. Correctional Services of Canada Research Report, R–54. Ottawa: Correctional Service of Canada Research Branch.

Serin, R., & Kennedy, S. (1998). *Assessment Protocol for Treatment Readiness, Responsivity, and Gain: Guidelines for Clinician Ratings*. Ottawa: Correctional Service of Canada Research Branch.

Shearer, R., Myers, R., & Ogan, G. (2001). Treatment resistance and ethnicity among female offenders in substance abuse treatment programs. *The Prison Journal, 81(1)*, 55–72.

Sykes, G.M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review, 22*, 664–670.

Van Voorhis, P. (1997). Correctional classification and the "responsivity principle." *Forum on Corrections Research, 9(1)*, 46–50.

Warren, M. (1983). Application of interpersonal maturity theory to offender populations. In W. Laufer & J. Day (Eds.). *Personality Theory, Moral Development, and Criminal Behavior*. Lexington, MA: Lexington Books, pp. 23–49.

Whitehead, J., & Lab, S. (1989) A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime and Delinquency, 26*, 276–295.

Wlodkowski, R., & Ginsberg, M. (1995). A framework for culturally responsive teaching. *Educational Leadership, 53*, 17–21.

Wormith, J.S. (2001) Assessing offenders assessments: Contributing to effects of correctional treatment. *ICCA Journal on Community Corrections, July 2001*, 12–23. ■

EFFECTIVE, from page 21

Rice, M.E., Harris, G.T., Varney, G.W., & Quinsey, V.L. (1989). *Violence in Institutions: Understanding, Prevention, and Control*. Toronto: Hans Huber.

Robinson, D., Porporino, F.J., & Simourd, L. (1992). *Staff Commitment in the Correctional Service of Canada*. Research Report R-21. Ottawa: Correctional Service of Canada.

Robinson, D., Porporino, F.J., & Simourd, L. (1996). Do different occupational groups vary on attitudes and work adjustment in corrections? *Federal Probation, 60(3)*, 45–53.

Serin, R.C., & Preston, D.L. (2000). Managing and treating violent offenders. In J.B. Ashford, B.D. Sales & W. Reid (Eds.). *Treating Adult and Juvenile Offenders with Special Needs*. Washington, DC: American Psychological Association.

Serin, R.C., & Preston, D.L. (2001). Designing, implementing and managing treatment programs for violent offenders. In G.A. Bernfeld, D.P. Farrington & A.W. Leischied (Eds.). *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs*. New York: Wiley, pp. 205–221.

Shapiro, D.L. (1991). *Forensic Psychological Assessment: An Integrative Approach*. Boston: Allyn and Bacon.

Silverman, M. (1993). Ethical issues in the field of probation. *International Journal of Offender Therapy and Comparative Criminology, 37(1)*, 85–94.

Thurber, A. (2000). *Accreditation of Correctional Programs: Recognizing and Maintaining Effective Program Interventions for Offenders*. ICCA Monograph. Lacrosse, WI: International Community Corrections Association.

Wittenburg, P.M. (1998). Successful mentoring in a correctional environment. *Federal Probation, 62(2)*, 75–80. ■

Understanding the Risk Principle: How and Why Correctional Interventions Can Harm Low-Risk Offenders

Christopher T. Lowenkamp,
Assistant Director,
The Corrections Institute,
University of Cincinnati
and
Edward J. Latessa,
Professor and Head,
Division of Criminal Justice,
University of Cincinnati

Over the last several years, the importance of the risk principle has been well established in many correctional settings. Simply stated, the risk principle indicates that offenders should be provided with supervision and treatment levels that are commensurate with their risk levels. However, there continues to be some confusion regarding the implications of the risk principle and why the trends predicted by the risk principal are observed. The purpose of this article is to discuss what the risk principle is, what it means for corrections, and why we see intensive treatments and supervision leading to no effect or increased recidivism for low-risk offenders.

Perhaps it is important that we begin by defining the concept of "risk" as it pertains to offender recidivism. For some, "risk" is a concept associated with the seriousness of the crime—for example, in the sense that a felon poses a higher risk than a misdemeanant. In actuality, however, though a felon has been convicted of a more serious offense than a misdemeanant, his or her relative risk of reoffending may have nothing to do with the seriousness of the crime.

For our purposes, "risk" refers to the probability of reoffending. A low-risk offender is one with a relatively low probability of reoffending (few risk factors), while a high-risk offender has a high probability (many risk factors). The application of the concept in corrections is similar to that in most actuarial sciences. For example, life insurance is cheaper for a nonsmoker in his 40s than for a smoker of the same age. The reason insurance costs more for the smoker is that smokers have a risk factor that is significantly correlated with health problems. Similarly, an offender who uses drugs has a higher chance of reoffending than someone who does not use drugs.

In 1990, Andrews, Bonta, and Hoge discussed the importance of the risk principle as it relates to the assessment of offenders. Their article makes clear that the risk principle calls for the administration and delivery of more intense services and supervision to higher-risk offenders. In contrast, lower-risk offenders should receive lower levels of supervision and treatment. Since 1990, considerable research has investigated how adhering to the risk principle can impact a correctional program's effectiveness.

Meta-Analyses Involving the Risk Principle

Meta-analysis after meta-analysis has revealed a similar trend when the risk principle is empirically investigated. Table 1, page 4, shows the results of seven meta-

analyses conducted on juvenile and adult offenders in correctional programs or school-aged youth in school-based intervention programs.

The first row of the table lists the results from a study conducted by Andrews, Zinger, Hoge, et al. (1990). This study investigated the effects of correctional interventions from 85 studies. Overall, they found that the correctional programs were much more effective when the correctional program took in mostly higher-risk offenders. Reductions in recidivism of 11% were noted in programs that had mostly higher-risk offenders versus 2% reductions for programs that took in both low- and high-risk offenders (re-analysis by Andrews and Bonta, 1998).

The second, third, and fourth rows summarize the findings of studies conducted by Dowden and Andrews. These three meta-analyses all indicate that programs serving a greater percentage of higher-risk offenders were more effective than those that did not. This finding was observed when looking at juvenile offenders, female offenders, and violence as an outcome measure.

The fifth row reports on the results of a meta-analysis that reviewed the effectiveness of drug courts. Again, drug courts where over half the offenders served had a prior record were twice as effective (10% versus 5% reduction) as drug courts where more than half the offenders served were first-time offenders. Finally, two meta-analyses report on the effectiveness of school-based interventions in reducing delinquent and analogous behaviors (Wilson, Gottfredson, and Najaka, 2002) and aggressive behavior (Wilson, Lipsey, and Derzon, 2003). Both studies indicate better effects when targeting youths who are at risk for the particular behaviors that are to be prevented.

Table 1. Summary of Meta-Analyses Investigating the Risk Principle

Study	No. of Studies Reviewed	Type of Studies Reviewed	Findings
Andrews et al. (1990)	85	Juvenile, mixed	Effect size 5 times as great when focusing on high-risk
Dowden and Andrews (1999a)	26	Juvenile and adult female, or mainly female	Effect size 6 times as great when following risk principle
Dowden and Andrews (1999b)	229	Young offenders	Effect size 4 times as great when when following risk principle
Dowden and Andrews (2000)	35	Juvenile and adult violent outcomes only	Effect size 2 times as great when when following risk principle
Lowenkamp et al. (2002)	33	Juvenile and adult drug courts	Effect size 2 times as great when when following risk principle
Wilson et al. (2002)	165	School-based interventions	Effect size 3 times as great when when targeting high-risk youth
Wilson et al. (2003)	221	School-based interventions targeting aggression	Effect size 4 times as great when when targeting high-risk youth

Differing Treatment Effects for High- and Low-Risk Offenders

While Table 1 provides plenty of support for the risk principle, a recent study that Lowenkamp and Latessa (2002) conducted in Ohio offers even more evidence. This study is the largest ever conducted of community-based correctional treatment facilities. The authors tracked a total of 13,221 offenders who were placed in one of 38 halfway houses and 15 community-based correctional facilities throughout the state. A 2-year follow-up was conducted on all offenders, and recidivism measures included new arrests and incarceration in state penal institutions. Treatments effects were calculated, which represent the difference in recidivism rates for the treatment group (those offenders with a residential placement) and the comparison group (those offenders that received just supervision with no residential placement).

Figure 1 shows the effect for low-risk offenders, using incarceration as the outcome measure. The negative numbers show the programs that were associated with increases in recidivism rates for low-risk offenders. The positive numbers show the few programs that were actually associated with reductions in recidivism for low-risk offenders. As you can see from this figure, the majority of programs in this study were associated with increases in the failure rates for low-risk offenders. Only a handful of programs reduced recidivism for this group, and the largest reduction was 9%.

Fig. 1 Changes in the Probability of Recidivism by Program for Low-Risk Offenders

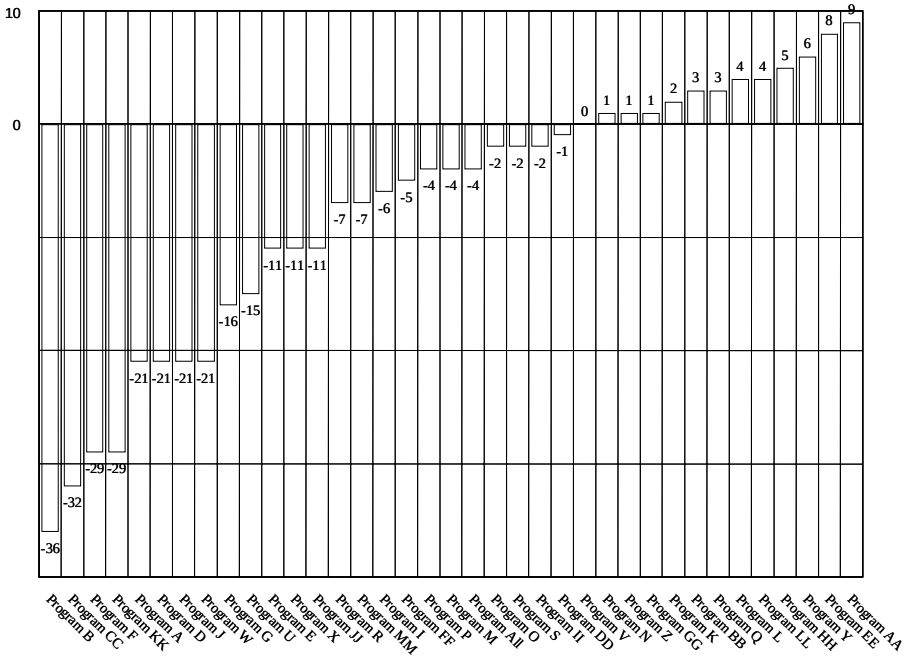
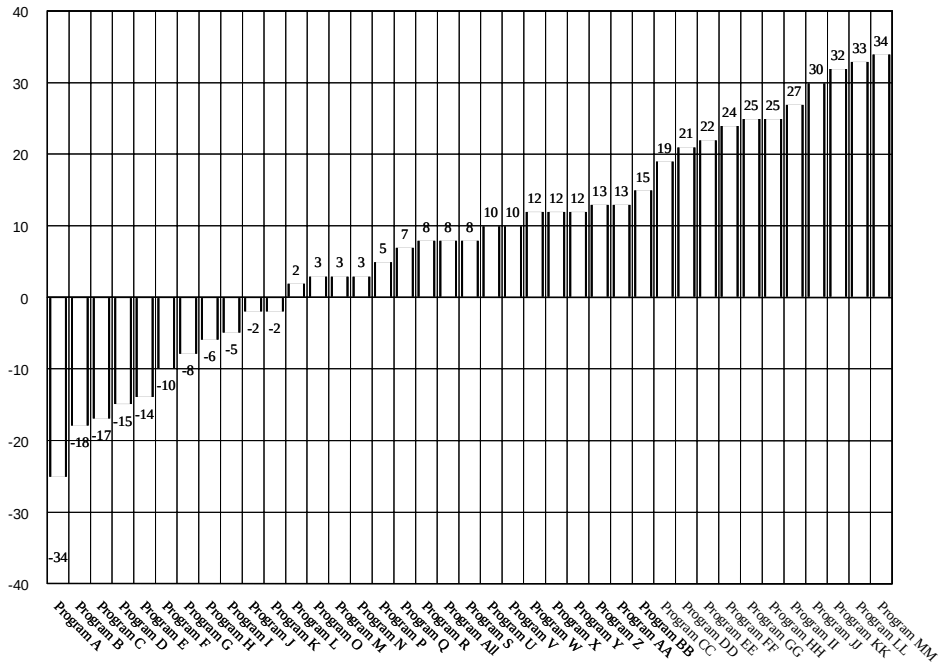


Figure 2 shows the results for high-risk offenders. Not only were most programs associated with reductions in recidivism for this group, but there were also eight programs that reduced recidivism over 20% and three programs that reduced recidivism over 30%. (Note that there were some programs in Ohio that did not reduce recidivism at any level of risk. This is likely related to program integrity. See Lowenkamp and Latessa, 2004.)

Fig. 2. Change in the Probability of Recidivism by Program for High-Risk Offenders



The best illustration of the risk principle can be seen by looking at the programs that had the greatest effect on high-risk offenders. Programs KK and MM each reduced recidivism for high-risk offenders by over 30%, yet looking at their effect for low-risk offenders, we see that Program MM increased recidivism for this group by 7% and Program KK by 29%. Thus, the same programs that reduced recidivism for higher-risk offenders actually increased it for low-risk offenders. The risk principle held across geographic location (rural, metro, urban) and with sex offenders (Lowenkamp and Latessa, 2002).

When taken together, these meta-analyses and individual studies provide strong evidence that more intense correctional interventions are more effective when delivered to higher-risk offenders, and that they can increase the failure rates of low-risk offenders. Recall the meta-analyses and the Ohio study, as well as Hanley (2003) and Bonta, Wallace-Capretta, and Rooney (2000), which both found that intensive supervision reduces recidivism for higher-risk offenders but increases the recidivism rates of lower-risk offenders.

Why Interventions Are More Successful with High-Risk Offenders

A question that continues to arise is why an intervention can have the intended consequences for a high-risk offender but have undesired and unintended consequences for a low-risk offender. To answer this question, one only need look at the risk factors for offending behavior. A review of the meta-analyses on the risk predictors consistently reveals antisocial attitudes, associates, personality, and a history of antisocial behavior as the strongest predictors (Andrews and Bonta, 1998). Other risk factors include substance abuse and alcohol problems, family characteristics, education, and employment (Gendreau, Little, and Goggin, 1996).

Given these risk factors, consider what a high-risk and a low-risk offender would look like. High-risk offenders would have antisocial attitudes, associates, and personalities, or a long criminal history, or substance abuse problems, or poor family relations, and would likely be unemployed. Low-risk offenders, on the other hand, would be fairly prosocial and have good jobs with some, if not many, prosocial contacts. That is, low-risk offenders likely have good jobs, good relationships with their families, good relationships with prosocial acquaintances, fairly prosocial attitudes, a limited criminal history, and few if any substance abuse problems. What happens to that low-risk offender when he/she is placed in a residential facility with high-risk offenders? You have likely come to an explanation for why we see low-risk offenders being harmed by intense correctional interventions.

The increased failure rates of low-risk offenders can largely be understood when considering the following three explanations:

- When we place low-risk offenders in the more intense correctional interventions, we are probably exposing them to higher-risk offenders, and we know that who your associates are is an important risk factor. Practically speaking, placing high- and low-risk offenders together is never a good idea. If you had a son or daughter who got into some trouble, would you want him or her placed in a group with high-risk kids?

- When we take lower-risk offenders, who by definition are fairly prosocial (if they weren't, they wouldn't be low-risk), and place them in a highly structured, restrictive program, we actually disrupt the factors that make them low-risk. For example, if I were to be placed in a correctional treatment program for 6 months, I would lose my job, I would experience family disruption, and my prosocial attitudes and prosocial contacts would be cut off and replaced with antisocial thoughts and antisocial peers. I don't think my neighbors would have a "welcome home from the correctional program" party for me when I was released. In other words, my risk would be increased, not reduced.

- Other factors such as IQ, intellectual functioning, and maturity might be at work. We rarely find programs that assess these important responsivity factors when they place offenders into groups. It could be the case that there

are some low-functioning, low-risk offenders who are manipulated by more sophisticated, higher-risk, predatory offenders.

What all this means for corrections is that low-risk offenders should be identified and excluded, as a general rule, from higher-end correctional interventions. We are pragmatists and therefore say "general rule," as we realize that programs are often at the mercy of the court or parole board in terms of who is referred to the program. Even so, programs that end up receiving low-risk offenders should make sure that those offenders are returned back to the environments that made them "low-risk." This can be achieved by developing programming (both treatment and supervision) that is based on the risk level of the offender.

In addition, the research reviewed here and the risk principle also dictate that we should direct the majority of services and supervision to higher-risk offenders because it is with this group of offenders that such interventions are most effective. The first step in meeting the risk principle is identifying the appropriate targets (higher-risk offenders). To achieve this, agencies must assess offenders with standardized and objective risk assessment instruments. Risk assessment is now considered the cornerstone of effective correctional intervention. ■

For more information:

Christopher T. Lowenkamp and Edward J. Latessa
Division of Criminal Justice
University of Cincinnati
P.O. Box 210389
Cincinnati, Ohio 45221-0389
(513) 556-0855
christopher.lowenkamp@uc.edu
edward.latessa@uc.edu

References

Andrews, D. A. & Bonta, J. (1998). *The Psychology of Criminal Conduct.* Cincinnati, OH: Anderson Publishing Co.

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

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology* 8, 369-404.

Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior* 27, 312-329.

Dowden, C. & Andrews, D. A. (1999a). What works for female offenders: A meta-analytic review. *Crime and Delinquency* 45, 438-452.

Dowden, C. & Andrews, D. A. (1999b). What works in young offender treatment: A meta-analysis. *Forum on Corrections Research* 11, 21-24.

Dowden, C. & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, 449-467.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology* 34, 575-607.

Lowenkamp, C. T. and Latessa, E. J. (2002). Evaluation of Ohio's halfway house and community based correctional facilities. Cincinnati, Ohio: University of Cincinnati.

Lowenkamp, C. T. , Holsinger, A. M., and Latessa, E. J. (2004). Are drug courts effective? A meta-analytic review. Unpublished manuscript, Cincinnati, Ohio: University of Cincinnati.

Lowenkamp, C. T. , Holsinger, A. M., and Latessa, E. J. (2004). Investigating the relationship between program integrity and correctional program effectiveness. *Ohio Department of Rehabilitation and Correction Research Compendium* 2, 208-213.

Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2002). School-based prevention of problem behaviors: A meta-analysis. Journal of Quantitative Criminology, 17, 247-272.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive and disruptive behavior: A meta-analysis. Journal of Consulting & Clinical Psychology, 2003, 71, 136-149.

DOES CORRECTIONAL PROGRAM QUALITY REALLY MATTER? THE IMPACT OF ADHERING TO THE PRINCIPLES OF EFFECTIVE INTERVENTION*

CHRISTOPHER T. LOWENKAMP
EDWARD J. LATESSA
PAULA SMITH
University of Cincinnati

Research Summary:
This study analyzed data on 3,237 offenders placed in 1 of 38 community-based residential programs as part of their parole or other post-release control. Offenders terminated from these programs were matched to, and compared with, a group of offenders (N = 3,237) under parole or other post-release control who were not placed in residential programming. Data on program characteristics and treatment integrity were obtained through staff surveys and interviews with program directors. This information on program characteristics was then related to the treatment effects associated with each program.

Policy Implications:
Significant and substantial relationships between program characteristics and program effectiveness were noted. This research provides information that is relevant to the development of correctional programs, and it can be used by funding agencies when awarding contracts for services.

KEYWORDS: Correctional Program Integrity, Correctional Program Assessment Inventory (CPAI), Treatment Effects, Halfway Houses

A report released recently by the U.S. Department of Justice indicated that the correctional population reached a new high in this country with almost 6.9 million offenders under correctional control at the end of 2003

* This research was supported by a contract from the State of Ohio Department of Rehabilitation and Correction. The views and opinions expressed are those of the authors and do not necessarily reflect the views and opinions of the State of Ohio Department of Rehabilitation and Correction. The authors wish to thank Reginald Wilkinson, Evalyn Parks, Linda Janes, and the rest of the staff at the State of Ohio Department of Rehabilitation and Correction who assisted with the Halfway House/ Community-Based Correctional Facility Study. The authors also wish to thank the anonymous reviewers and Paul Gendreau for their comments and advice on earlier versions of this manuscript.

(Glaze, 2004). Although this figure only represented a 2% increase from the previous year, it capped a 50% increase in the correctional population since 1990. The prison population experienced the greatest increase during this time period (76%), but probation and parole populations grew as well (44% and 37%, respectively) (U.S. Department of Justice, 2001).

This growth in offender populations under community supervision, coupled with reductions in resources available for community-based correctional agencies, causes concern. Research on recidivism in probation samples indicates recidivism rates as high as 65% (Petersilia, 1985). Although other studies indicate much lower recidivism rates (e.g., McGaha et al., 1987; Vito, 1986), a national sample from 1986 indicates that 62% of the sample had a disciplinary hearing for a violation of probation or were rearrested for another felony during a three-year follow-up. Additionally, 46% had been sentenced to prison or jail or had absconded from supervision within that time period (United States Department of Justice, 1992).

Other statistics pertaining to probation and parole populations indicate that a substantial number of offenders continue to exhibit criminal behaviors while under community supervision. To illustrate, the number of offenders entering state prison systems on violations during community supervision rose from 18,000 to 142,000 between 1975 and 1991. By 1991, probation and parole violators constituted 45% of the prison population. Most community supervision violators (77%) in prison were sentenced to prison for a new felony conviction (Cohen, 1995).

In 1991, probation and parole violators in prison equated to roughly 318,000 prisoners. These 318,000 offenders committed 13,200 new murders, 12,900 new rapes, 19,200 new assaults, and 39,500 new robberies while under community supervision. Furthermore, although felony probationers and parolees are not permitted to carry firearms, 21% of the community supervision violators in prison for committing a new offense reported possessing a firearm while under post-release control, and most were armed when committing the new offense (Cohen, 1995). Additional data indicate many other transgressions committed by offenders under community supervision, such as substance abuse, absconsion, failure to pay fines, and failure to meet other conditions of supervision (Bonzcar, 1997; Cohen, 1995; U.S. Department of Justice, 1992). These data certainly question the ability of community supervision to effect meaningful behavioral change in a direction favorable to public safety.

What can be done to enhance the effectiveness of community-based services in order to rehabilitate offenders? Research on intensive supervision programs and other supervision enhancements based on custody, control, and/or deterrence has failed to show promise in reducing the recidivism of offenders supervised in the community (Cullen et al., 1996; Petersilia and

Turner, 1993). One promising method of reducing the recidivism of offenders under community supervision is to provide "clinically relevant" correctional interventions (Andrews et al., 1990b). A popular method for delivering these services is through the use of community residential facilities (e.g., halfway houses). There is, however, considerable variation in the structure and effectiveness of these facilities (see Latessa and Travis, 1992).

Previous research helps to explain this variation in correctional program effectiveness. Cullen (2002) describes "the Canadians' theory of rehabilitation," including predictions about which correctional programs should be effective in reducing recidivism and which programs are likely to fail. A group of Canadian researchers have developed the basis for this theory, which in turn provides specifications for a successful correctional program [see also the important contributions of Quay (1979) and Palmer (1991, 1994, 1995)]. More specifically, correctional programs and interventions should focus on higher risk offenders; deliver cognitive-behavioral or behavioral interventions that focus on relevant criminogenic needs; attend to the qualifications, skills, and values of staff; and evaluate what they do (Andrews et al., 1990a; Andrews et al., 1990b; Cullen, 2002; Gendreau, 1996). Adherence to the parameters of the theory can be measured with the Correctional Program Assessment Inventory (CPAI)[1] (see Gendreau and Andrews, 1996). However, there is little empirical research on the relationship between the scores yielded by the CPAI and program effectiveness [and the research that has been conducted to date includes Gray (1997), Holsinger (1999), Lowenkamp (2004), and Nesovic (2003)]. The current research investigates the relationship between the CPAI and program outcome using program integrity and offender data from 38 halfway house (HWH) programs in Ohio.

The findings of this research are important for several reasons. First, the offender samples used in this research comprise parolees. Given the current focus on the importance of offender re-entry, this research is timely and relevant to current issues in corrections. Second, this research focuses on identifying the characteristics of effective correctional programs. Once these characteristics have been identified, a blueprint of effective correctional programs can be developed and refined, providing invaluable information to existing correctional programs or agencies that are in the process of developing and implementing interventions. Third, this research provides information that is important to policy makers and funding sources when making funding decisions regarding correctional programs.

1. The most recent version of the CPAI is the CPAI-2000 (Gendreau and Andrews, 2001)

Fourth, it allows for a partial test of "the Canadians' theory of rehabilitation" mentioned above and discussed in greater detail below.

METHODOLOGY[2]

The main research question is: "Does a relationship exist between program integrity and program effectiveness?" Previous research indicates that certain types of interventions are more effective than others in reducing recidivism (Gendreau, 1996). However, fewer studies have provided strong evidence of the empirical link between program integrity and program effectiveness. This study will provide empirical analyses of the relationship between dimensions of program integrity as measured by the CPAI and program effectiveness.

To answer this research question, data on program integrity and program effectiveness were collected for each of the 38 programs included in the analyses. To develop data on program integrity, individuals trained in the application of the CPAI conducted site visits to each program. Many of the program characteristics included on the CPAI[3] were scored for each program based on interviews with the program director and a review of relevant program materials. Additional data on program integrity was gathered from surveys completed by all staff at each program. It is important to emphasize that the CPAI was not scored using the standard assessment protocol, nor were data collected to score the entire instrument for all programs. Specifically, some items on the CPAI were not scored as site visits were relatively brief and did not involve an examination of multiple sources of data (e.g., structured interviews with staff and offenders, observation of group treatment sessions, and so on). In addition, the scoring of the CPAI was based on information collected in 2002. A concern arises in that the programs may have changed from fiscal year 1999 when the treatment group received the services. To address this issue, interviews with the program directors were based on their recollections of program operations in 1999. Furthermore, program materials relevant to the CPAI assessment were also limited to that time period. This research should not be considered a validation of the CPAI; rather the instrument was used to structure data collection and analysis of variables related to program integrity.

A comparison of the recidivism rates of the treatment and comparison groups for each program served as the measure of program effectiveness. The recidivism data were based on incarceration rates collected during a

2. For more detailed information on the methodology, see Lowenkamp and Latessa (2002) and Lowenkamp (2004).

3. Due to the limited nature of the interviews with staff and program directors, the total number of items to be scored on the CPAI was reduced from 77 to 62.

two-year follow-up period and were collected from the database maintained by the Ohio Department of Rehabilitation and Correction.

SAMPLE

The sample of programs included all Community Corrections Act funded programs in the State of Ohio that provided residential services to state parolees through a halfway house program (HWH). To be included in the sample, the program had to be in operation during fiscal year 1999 (July 1, 1998 to June 30, 1999). These parameters yielded a total of 34 service providers operating 38 distinct offender programs.

To calculate program effectiveness, data were collected on individual offenders who participated in these programs and were related to a comparison group of similar offenders. Specifically, the treatment group consisted of all offenders who were served by the aforementioned programs with a release date during fiscal year 1999, and who were referred to and entered a HWH operated by the State Parole Board. The offenders sampled (N = 3,237) were spread across the 38 distinct programs with an average of 139 offenders per program. The range was 12 to 329, with over 75% of the programs having served 50 or more offenders in fiscal year 1999. Approximately 65% of the sample was classified as successful terminations, with the number of successful terminations from each program ranging from 6 to 272.

The parolees/post-release control releases in the treatment group were matched to parolees/post-release control releases who were under parole supervision, but who were not placed in a HWH. The comparison cases (N = 3,237) were drawn from a larger sampling frame (N = 6,781) and were matched with the treatment cases on county of conviction, sex, and risk level as determined by a modified version of the Salient Factor Score (Hoffman, 1994).[4] When multiple cases were available, the first eligible

4. To match cases from the treatment and comparison group and determine whether programs provided greater services to high-risk or low-risk offenders, a measure of risk needed to be developed. To create a risk scale, a modified version of the Salient Factor Score (SFS) was developed. This scale is based on the original items and weights used in the SFS, with some slight modifications. Modifications included replacing the item measuring commitment-free periods with an item that taps employment status, replacing the item measuring current CJ status with an item that measures community supervision violations, and modifying the substance abuse item (which previously focused only on opiates) to include any drug addiction. The only other change was the use of arrests rather than convictions. Although slight adjustments were made, overall the two scales are fairly similar and still range in value from 0 to 10. Only the measure of employment changed substantively. The use of employment status at arrest is supported by reviewing the risk factors measured on other known risk assessment instruments (e.g., The Wisconsin Risk Scale, The Statistical Index of Recidivism, and the Level of Service Inventory). Previous research on risk factors also provides support

case was selected for inclusion in the comparison group.

TABLE 1. DESCRIPTIVE STATISTICS FOR ALL HALFWAY HOUSES BY GROUP MEMBERSHIP

Variable	Treatment Group	Comparison Group
Age*	Mean (*N*) 34 years (3,237)	Mean (*N*) 37 years (3,237)
Race Black White	% (*N*) 62 (2,017) 38 (1,220)	% (*N*) 61 (1,959) 39 (1,278)
Sex Male Female	% (*N*) 91 (2,959) 9 (278)	% (*N*) 91 (2,959) 9 (278)
Prior Arrest (Yes)*	% (*N*) 93 (3,022)	% (*N*) 87 (2,822)
Prior Incarcerations (Yes)*	% (*N*) 50 (1,618)	% (*N*) 40 (1,299)

*Difference is significant at $p < 0.05$.

The descriptive statistics on demographic characteristics for the treatment and comparison groups are reported in Table 1. Data reported in this table indicate that the comparison group was significantly older than the treatment group (37 versus 34 years old). The two groups were equivalent in terms of racial composition. Although the two groups differed in terms of the specific criminal history measures, most participants had been in contact with the criminal justice system before their inclusion in this study.

RECIDIVISM

The measures of recidivism used in this research included a return to an Ohio Correctional Facility for any reason (technical violation or new arrest). In addition, the data were analyzed for incarceration as a result of a technical violation versus a new arrest separately. The decision to use these measures over others was based on the reliability of incarceration data. Although incarceration is a conservative measure of future criminal behavior, the data on prison intakes are complete and easy to query.

The actual measure of effectiveness is the logged odds ratio, which is a good measure to use when dealing with two groups of offenders (treatment versus comparison) and two distinct outcomes (success versus

for the use of a measure of employment when assessing risk and predicting recidivism (Gendreau et al., 1996).

failure).[5]

PROGRAM INTEGRITY DATA

Program integrity was measured using an abbreviated version of the CPAI, measures developed from responses on the staff surveys, and data available from the CCIS database maintained by the State of Ohio Department of Rehabilitation and Correction. The areas covered with these measures, and those specifically found on the CPAI, include program implementation, client pre-service assessment, program characteristics, staff characteristics, evaluation, and an "other" category containing miscellaneous items.

Program implementation measures the qualifications of the program director, his/her involvement in the program, community support for the program, planning and research, as well as funding. Client pre-service assessments are concerned with the appropriateness of the clients received by the program and assessment practices related to risk, need, and responsivity. The program characteristics section measures the type of treatment, treatment targets, duration and dosage of treatment, matching of offenders and staff to programming, use of rewards and sanctions, the presence of aftercare, and whether the program varies the intensity and duration of services by risk. The staff section of the CPAI measures the education and experience of staff, evaluation and supervision of staff, staff attitudes toward treatment, staff training, and the ability of staff to have input into the program. The evaluation section measures how well a program evaluates itself through the use of quality assurance mechanisms and outcome evaluations. The final section of the CPAI includes miscellaneous items pertaining to the program such as disruptive changes in the program, funding, or community support, ethical guidelines, and the comprehensiveness of client files.

ANALYSIS

To determine whether a relationship exists between program integrity and program effectiveness, several separate analyses were conducted. First, all items on the CPAI were tallied to generate a total score for each program. To calculate an overall score, the total number of points received across the six sections were totaled and then divided by the total number of points possible across the entire instrument. Second, each section on the CPAI was tallied to provide a score for each section. To calculate scores for each section on the CPAI, the total points received in each section were divided by the total number of points possible in each section. The

5. For more information on the logged odds ratio, see Lipsey and Wilson (2001).

CPAI scores were then correlated with treatment effects to determine the relationship between program integrity and treatment effectiveness. The use of 95% confidence intervals (*CI*s), as recommended in the 2001 APA Publication Manual (American Psychological Association, 2001), was emphasized in the interpretation of results. For a discussion of the different ways the *CI* can be interpreted, see Cumming and Finch (2005). It should be noted that the *CI* can be interpreted as a significance test; that is, a *CI* that contains 0 is not statistically significant. It should also be emphasized, however, that because a *CI* includes 0 does not mean that there is no effect; it is extremely rare for 0 to be the point estimate (Hunter and Schmidt, 2004). The primary utility of the *CI* rests on the interpretation of its width. As the width of the *CI* increases, the precision of the estimate of μ decreases (i.e., the estimate is associated with more uncertainty) (Hunter and Schmidt, 2004).

RESULTS

This section reviews the results of this research. The findings are presented in three parts: The first contains information on the treatment effects of the 38 programs in this study, the second provides the information relevant to the relationship between the CPAI and program effectiveness, and the third investigates how these programs can be grouped by their CPAI scores and whether these groupings have practical value. The implications of these findings are then summarized and discussed.

TREATMENT EFFECT SIZES

The first set of analyses focused on calculating the treatment effect sizes and weights for each program. Recall that the measure of effectiveness is calculated using the formula for the logged odds ratio, whereas the weight was the inverse of the estimate variance. A logged odds ratio equal to zero indicates no difference in recidivism rates between the treatment and the comparison groups. A negative logged odds ratio indicates that the comparison group had a lower recidivism rate relative to the treatment group. Finally, a positive logged odds ratio indicates a program where the treatment group outperformed the comparison group (i.e., recidivated at a lower rate than the comparison group). A logged odds ratio was calculated using all program participants and their respective comparison cases and then calculated using only those offenders who were successfully terminated and their respective comparison cases (i.e., cases were yoked).

When reviewing the effect sizes calculated with all program participants (both successful and unsuccessful terminations included in the analysis), the smallest effect is –1.54, and the largest effect size is 2.15. A total of 28 of the 38 programs are associated with effect sizes equal to or less than 0,

which means that for 73% of the programs, the comparison group recidivated at lower rates (or equal rates in the case of 0) than the treatment group. The mean effect size is –0.43 with a 95% confidence interval that ranges from –0.53 to –0.32 (see Table 2). A logged odds ratio of –0.43 would roughly equate to a recidivism rate for the treatment group that is 10 percentage points higher than the comparison group.

Effect sizes calculated with successful terminations only provided a range of –1.17 to 4.48 with an average of 0.15 and a 95% confidence interval from 0.01 to 0.29 (see Table 2). With these data, only 39% (15 out of 38) of the programs were associated with treatment effects that favored the comparison group. A mean logged odds ratio of 0.15 approximates a recidivism rate for the treatment group that is four percentage points lower than that of the comparison group.

TABLE 2. DESCRIPTIVE STATISTICS FOR EFFECT SIZES AND WEIGHTS

Item	*N*	Effect Size	95% *CI*	Weight	Weight Standard Deviation
All Program Participants	38	–0.43	–0.53 to –0.32	9.18	8.40
Successful Program Participants	38	0.15	0.01 to 0.29	5.42	5.88

Comparing these two measures, it is obvious that, when considering all program participants, most programs are not associated with very positive results. This conclusion is evident from the mean treatment effect calculated (–0.43) and the percentage of programs that had no effect or that were associated with an increase in recidivism rates in relation to the comparison group. When the treatment effect size is calculated using only successful terminations, the data yield a somewhat different trend. Using this measure, on average, the programs are associated with a reduction in recidivism; in other words, most program participants performed better than their matched counterparts.

The finding that treatment effects are linked to the termination status of offenders in the sample could be attributable to the relationship between unsuccessful discharge from the program and reincarceration. That is, those offenders who were unsuccessfully discharged from the program likely had technical violations filed against them. This filing, in turn, led to revocation of community supervision and incarceration in a state facility after termination. This explanation helps to explain the higher rates of incarceration among unsuccessful terminations and the greater treatment effects noted when only successful terminations are used to calculate the effects of the programs on recidivism. Alternatively, this finding could be

attributable to the increased dosage of treatment received by the successful terminations in comparison with the unsuccessful terminations (see Bourgon and Armstrong, 2005).

The goal of this research was to determine whether there was any relationship between measures of program integrity and treatment effectiveness. Of particular interest is whether the measures of program integrity can be used to identify those programs that are effective in reducing recidivism from those that are not. The practical implications associated with this question are obvious. The next section discusses the observed relationships between program integrity and treatment effects using both successful and unsuccessful terminations combined, as well as the successful terminations alone.

PROGRAM INTEGRITY AND OUTCOME

The next set of analyses involved tabulating and calculating the scores on the CPAI and calculating the correlations between the CPAI and treatment effectiveness. The CPAI is divided into six sections. The data for all sections, including the total score and the correlations with treatment effectiveness, are reported in Tables 3 and 4. Table 3 reports the results for on all participants (successful and unsuccessful terminations combined), whereas Table 4 reports the results for the successful terminations only.

TABLE 3. CORRELATIONS AMONG CPAI COMPONENT SCORES, TOTAL SCORES, AND TREATMENT EFFECT FOR ALL PROGRAM PARTICIPANTS (SUCCESSFUL AND UNSUCCESSFUL TERMINATIONS COMBINED)

		New Offense		Technical Violation		Return to Prison[a]	
	Mean	*r*	*CI*	*r*	*CI*	*r*	*CI*
Program Implementation	53.22	0.33*	0.01 to 0.58	0.58*	0.32 to 0.76	0.55*	0.28 to 0.74
Pre-Service Client Assessment	39.71	0.36*	0.05 to 0.61	0.38*	0.07 to 0.62	0.39*	0.08 to 0.63
Program Characteristics	43.11	0.20	–0.13 to 0.48	0.20	–0.13 to 0.48	0.22	–0.11 to 0.50
Staff Characteristics	43.29	–0.16	–0.45 to 0.17	0.01	–0.31 to 0.33	–0.12	–0.42 to 0.21
Evaluation	47.80	0.45*	0.15 to 0.67	0.24	–0.09 to 0.52	0.32	0.00 to 0.58
Other Characteristics	71.05	–0.06	–0.37 to 0.27	–0.05	–0.36 to 0.27	–0.05	–0.36 to 0.27
Total Score	45.51	0.35*	0.03 to 0.60	0.44*	0.14 to 0.67	0.42*	0.12 to 0.66

*Significant at $p < 0.05$.
[a]Return to prison combines incarceration for both new offenses and technical violations.

TABLE 4. CORRELATIONS AMONG CPAI COMPONENT SCORES, TOTAL SCORES, AND TREATMENT EFFECT FOR SUCCESSFUL TERMINATIONS ONLY

		New Offense		Technical Violation		Return to Prison[a]	
	Mean	*r*	*CI*	*r*	*CI*	*r*	*CI*
Program Implementation	53.22	0.31	–0.01 to 0.57	0.34*	0.02 to 0.60	0.45*	0.15 to 0.68
Pre-Service Client Assessment	39.71	0.31	–0.01 to 0.57	0.32	0.00 to 0.58	0.30	–0.02 to 0.57
Program Characteristics	43.11	0.03	–0.29 to 0.34	0.07	–0.26 to 0.38	0.10	–0.23 to 0.41
Staff Characteristics	43.29	0.01	–0.31 to 0.33	–0.02	–0.34 to 0.30	–0.05	–0.36 to 0.27
Evaluation	47.80	0.24	–0.09 to 0.52	0.21	–0.12 to 0.49	0.25	–0.08 to 0.53
Other Characteristics	71.05	0.15	–0.18 to 0.45	0.07	–0.26 to 0.38	0.13	–0.20 to 0.43
Total Score	45.51	0.26	–0.06 to 0.54	0.27	–0.06 to 0.54	0.33*	0.01 to 0.59

*Significant at $p < 0.05$.
[a]Return to prison combines incarceration for both new offenses and technical violations.

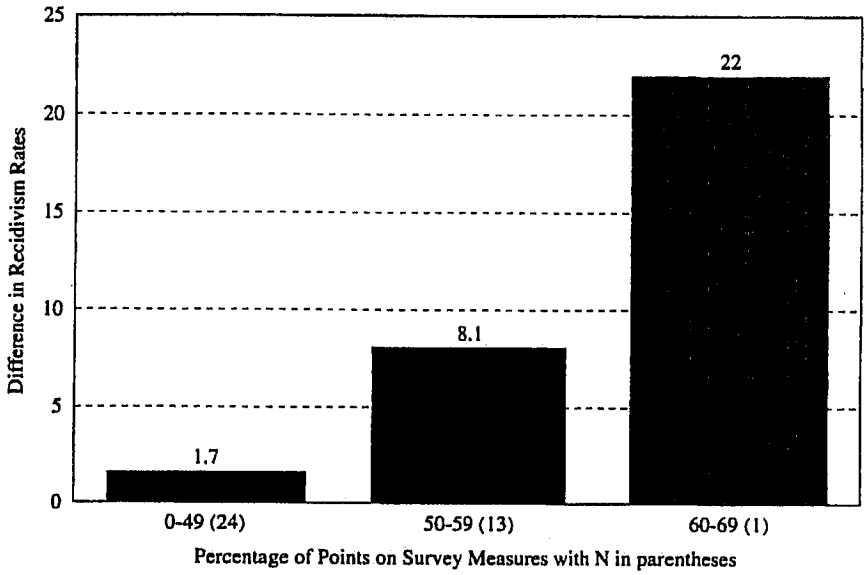
First, three subcomponents of the CPAI are correlated with at least one outcome measure. The scores in the program implementation area and pre-service client assessment are correlated with all three outcome measures, whereas evaluation is only correlated with incarceration for a new offense. Despite the fact that only three sections are significantly correlated with the outcome measures, the overall score (i.e., total score) is significantly correlated with all three outcome measures. The associated *CI*s did not include 0, but they were relatively wide. It should also be noted that, for later discussion, the average total score for the programs included in this study was 45.51%.

In summary, the total CPAI score significantly correlates with all three outcome measures (*r* values of 0.35, 0.44, and 0.42). Examination of the subscales indicates that several of the components are not significantly correlated to the outcome measures. In addition, several items are inversely correlated with the outcome measures.[6]

The pattern of results is somewhat different when the relationships are examined for successful terminations only. The total CPAI score does not

6. Reducing the CPAI score to include only those items that are positively and significantly correlated with the outcome measures increases the strength of the relationship as would be expected (i.e., *r* values of 0.60 and 0.47). For a more detailed description, see Lowenkamp (2003).

FIGURE 1. DIFFERENCE IN RECIDIVISM RATES BETWEEN TREATMENT AND COMPARISON GROUPS BASED ON THE CPAI MEASURE TOTAL SCORE



significantly correlate with new offenses or technical violations but is associated with return to prison. Only one subcomponent, program implementation, is correlated with at least one outcome measure. It should be noted, however, that the confidence intervals are wide and all overlap with one another.

CATEGORY OF PROGRAMS BASED ON TOTAL PROGRAM QUALITY SCORES

The final analyses in this research focus on determining whether the cutoff scores developed for the CPAI provide meaningful groupings of programs based on their associated reductions in recidivism. To this end, categories based on the recommended cutoff scores were used in this research. As per the assessment protocol for the CPAI, programs that scored 0% to 49% of the points are associated with the "unsatisfactory" category, a score of 50% to 59% falls into the "satisfactory but needs improvement" category, a score of 60% to 69% is categorized as "satisfactory", whereas a score of 70% or higher leads to placement in the "very satisfactory" group. The next step was to calculate a weighted average treatment effect for each group (based on the treatment effects of successful terminations

only). Once the average effect for each group was calculated, the logged odds ratio was converted back into a percentage point change in recidivism rates. The results of this analysis are displayed in Figure 1.

Figure 1 indicates that the majority (24 out of 38, or 68%) of the programs fell into the "unsatisfactory" category. Although this finding represents a large proportion of the programs, it is consistent with other research on the CPAI, which indicates that most programs fail to attain scores above 50%. Recall from earlier that the average score on the assessment for the programs included in this research is approximately 45.51%. Approximately 35% of the programs fell into the "satisfactory but needs improvement" category, and only one program was rated "satisfactory." What is interesting to note is that the percentage point reductions in returns to prison increase from one category to the next. Although the "unsatisfactory" group of programs averaged a 1.7% reduction in returns to prison, the "satisfactory but needs improvement" group averaged an 8.1% reduction in returns to prison. Finally, the one program that scored over 60% demonstrated a 22% reduction in returns to prison. Assuming a base recidivism rate of 50% for the comparison group, the relative reductions in returns to prison are 4%, 16%, and 44%, respectively.

SUMMARY AND DISCUSSION

Over the past 25 years, there has been a focus on identifying which correctional interventions effectively reduce recidivism. This research has converged to identify cognitive-behavioral and behavioral programs as the most effective interventions with most offenders. This research, however, often indicated considerable heterogeneity in the effectiveness of correctional programs that were based on cognitive-behavioral or behavioral therapies (for instance, see Lipsey et al., 2001; Pearson et al., 2002; Wilson et al., 2000). That is, when pooling all the studies on correctional interventions where a cognitive-behavioral or behavioral program was studied, some programs were very effective, some were moderately effective, and some had no effect or iatrogenic effects. It has been argued that some of this heterogeneity can be explained by program integrity (Cullen, 2002; Gendreau, 1996; Palmer, 1995; Quay, 1977). Given the problems associated with criminal behavior and the rehabilitation of offenders in this country, it is imperative to understand the link among program implementation, program integrity, and program effectiveness. As such, this research set out to answer the question: "Is there a relationship between program integrity and program effectiveness?"

The analyses of the treatment effects calculated indicated that overall not much effect was demonstrated by the 38 programs included in the analyses (mean logged odds ratio of 0.15, indicating a 4% reduction in

returns to prison). This effect was observed when calculating treatment effects using data on successful program terminations only. When using data from all program participants, a 10% increase in returns to prison is noted for the programs collectively.

With both outcome measures, considerable differences are noted in the ability of the HWH programs to reduce recidivism. The purpose of this research was to determine whether the HWH programs could be categorized based on measures of program integrity. In other words, the question of interest is as follows: "Are the measures of program integrity able to explain some of the differences observed in the effectiveness of the programs?"

The CPAI-based measures of program integrity significantly correlated with the outcome measure using all program participants and using successful terminations only in several instances. Consistently, significant correlations were observed for program implementation, pre-service client assessment, evaluation, and total scores. The correlations for the total scores, depending on the outcome measure, ranged from 0.26 to 0.42. These correlations equate to *r*-squared values of 0.07 and 0.18, respectively, meaning the total score for program integrity explains between 7% and 18% of the variation in treatment effects depending on how the items in the CPAI are configured and the outcome measures used. The information provided by the analyses leads to the conclusion that program integrity is in fact related to program effectiveness in some fashion.

Overall this research indicated that there is a fairly strong correlation between program integrity (as measured by the CPAI) and reductions in recidivism. More specifically, the analyses conducted here indicate that program implementation, offender assessment, and evaluation are all important in determining the effectiveness of a correctional program. Although several items are not related to program effectiveness or are related but in an unexpected direction, there is some consistency in the positive relationships that do exist. There also seems to be some consistency in the identification of matching as an important factor to consider in programming given the significant correlations between pre-service client assessment and outcome. Monitoring offenders' whereabouts and peer interactions have also been identified as important, as is varying the services by risk level, providing aftercare, and the provision of criminogenic services. These factors are all implicated by the theory of rehabilitation discussed earlier. Specifically, the core principles of risk, need, and responsivity would predict that the identified factors are important in developing correctional programming that is effective in reducing recidivism rates.

LIMITATIONS OF CURRENT RESEARCH

This research has provided a considerable amount of information on the relationship between program integrity and program effectiveness. Although this research contributes to the literature on this topic, the current study has several limitations. First, the sample size is small and limited to only 38 programs in total. The sample is further limited to adult offenders and is dominated by programs that serve populations of either exclusively male offenders or a majority of male offenders. In addition, all programs in the sample were located in the State of Ohio and only provided services to offenders released from a state prison facility. Given these considerations, the finding of the research may not be applicable to other offender populations.

The second limitation involves the timing of the data collection on program integrity. As mentioned, the offenders were released from the programs during fiscal year 1999, but the program integrity data were collected in 2002. Although attempts were made to gather data on the program as it operated in 1999, this was not always possible due to staff turnover. As a check on the reliability of this process, several programs included in this research project were assessed during or around fiscal year 1999 with the CPAI by independent assessors. The results of the assessments conducted during or around 1999 and during 2002 (but based on recollections of program operations in 1999) were strongly correlated (r = 0.74) (see Lowenkamp, 2004).

Third, the CPAI was not scored in the standard format, nor were data collected to score the entire instrument. Due to the brevity of the site visits, some items on the CPAI were not scored. In addition, the scoring of the CPAI was based on information reported by the program director based on his or her recollection of program operations in 1999. The CPAI is typically scored based on interviews with the program director, program staff, and offenders in the program, and it incorporates information from supporting documentation and observation of treatment groups. As such, this research should not be considered a validation of the CPAI. The CPAI was simply used to structure data collection and the scoring of program characteristics.

Finally, the follow-up period was two years. Perhaps using a longer follow-up period would yield different results.

IMPLICATIONS FOR CORRECTIONAL POLICY

Notwithstanding the limitations, this research has important implications for correctional policy. Residential correctional facilities have become increasingly popular as a way to reintegrate offenders returning home from prison, to deliver correctional interventions to offenders under

supervision in the community, to punish offenders, and to reduce prison and/or jail populations. Previous research on the theories underlying many of these correctional interventions indicates that treatments are often not based on sound theories and research, and that the interventions typically fail to embody the principles of effective correctional interventions (Gendreau, 1996). When this occurs, even empirically based programs can have null or iatrogenic effects (Barnosky, 2004).

Although there is research on individual principles of effective correctional interventions, these studies are often limited by the data and information presented by the original researcher. The current research is unique in that the data collected specifically relates to program integrity, and it measures many principles of effective intervention. As a result, it was possible to contribute to the literature by testing these principles more completely. The results of such tests have several implications for both residential correctional interventions and correctional policy.

First, program integrity matters. Although several principles related to effective interventions have been verified through meta-analyses and traditional literature reviews, there is limited research that measures program integrity and its relationship to program effectiveness. This research has demonstrated, albeit with a limited sample, that program integrity is related to program effectiveness. Anecdotally, correctional practitioners question the utility of program assessment. Prior literature has also identified dysfunctional attitudes about evaluation research (Van Voorhis et al., 1995). This research indicates that program integrity can be measured and then used to predict the effectiveness of a correctional program.

Second, the factors measured to assess program integrity are malleable or dynamic. Although it would likely take considerable time and effort for a correctional program to enhance the quality of the program, it is certainly possible. Furthermore, assessing program integrity can facilitate change in correctional programs. After having program integrity assessed, a correctional program can focus on problem areas identified in the assessment and increase its effectiveness. This research also supports the use of the CPAI and the related body of literature on the development of correctional programs.

Finally, this research can help funding agencies determine which programs are likely to have substantial impacts on recidivism, those programs that are likely to have negligible effects, and those that are likely to have iatrogenic effects. Budgetary constraints have historically and continually plagued corrections; this research suggests that policy makers and funding agencies should make decisions regarding the financial support of programs based, at least in part, on program integrity. Similar research has indicated that even sound and empirically supported programs, when poorly implemented or delivered, can cost taxpayers additional money by

leading to increases in the likelihood of recidivism for offenders served by the program.

Research should continue to investigate the link between program integrity and program effectiveness. These efforts will help programs and funding agencies at all levels develop sound correctional options in both community-based and institutional settings. This information should be shared with correctional practitioners, and those practitioners should have support when implementing what is revealed by future research. Such efforts should lead to a pool of quality correctional interventions that can provide long-term and cost-effective public safety by helping offenders make behavioral changes.

REFERENCES

American Psychological Association
2001 Publication Manual of the American Psychological Association. 5th ed. Washington, D.C.: Author.

Andrews, Don A., James Bonta, and Robert D. Hoge
1990 Classification for effective rehabilitation: Rediscovering psychology. Criminal Justice and Behavior 17:19–52.

Andrews, Don A., Ivan Zinger, Robert D. Hoge, James Bonta, Paul Gendreau, and Francis T. Cullen
1990 Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. Criminology 8:369–404.

Barnosky, Robert
2004 Outcome Evaluation of Washington State's Research-Based Programs for Juvenile Offenders. Olympia, Wash.: Washington State Institute for Public Policy.

Bonzcar, Thomas P.
1997 Characteristics of Adults on Probation, 1995. Washington, D.C.: Bureau of Justice Statistics, U.S. Department of Justice.

Bourgon, Guy and Barbara Armstrong
2005 Transferring the principles of effective treatment into a "real world" prison setting. Criminal Justice and Behavior 32:3–25.

Cohen, Robyn L.
1995 Probation and Parole Violators in State Prison, 1991. Washington, D.C.: Bureau of Justice Statistics, U.S. Department of Justice.

Cullen, Francis T.
2002 Rehabilitation and treatment programs. In James Q. Wilson and Joan Petersilia (eds.), Crime and Public Policy. 2d ed. San Francisco, Calif.: ICS Press.

Cullen, Francis T., John P. Wright, and Brandon K. Applegate
1996 Control in the community: The limits of reform? In Alan T. Harland (ed.), Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply. Thousand Oaks, Calif.: Sage.

Cumming and Finch
2005 Inference by eye: Confidence intervals, and how to read pictures of data. American Psychologist 60:170-180.

Gendreau, Paul
1996 The principles of effective interventions with offenders. In Alan T. Harland (ed.), Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply. Thousand Oaks, Calif.: Sage.

Gendreau, Paul and Don A. Andrews
1996 Correctional Program Assessment Inventory (CPAI). 6th Edition. Saint John, New Brunswick: University of New Brunswick.
2001 The Correctional Program Assessment Inventory—2000.

Gendreau, Paul, Tracy Little, and Claire Goggin
1996 A meta-analysis of the predictors of adult offender recidivism: What works! Criminology 34:575–607.

Glaze, Lauren E.
2004 Probation and Parole in the United States, 2003. Washington, D.C.: U.S. Department of Justice.

Gray, Glenn Allan
1997 Does coercion play a significant role in community treatment programs that reduce offender recidivism? The University of New Brunswick: Master's Thesis.

Hoffman, Peter B.
1994 Twenty years of operation use of a risk prediction instrument: The United States parole commission's salient factor score. Journal of Criminal Justice 22:477–494.

Holsinger, Alexander M.
1999 Opening the 'black box': Assessing the relationship between program integrity and recidivism. The University of Cincinnati: Doctoral Dissertation.

Hunter and Schmidt
2004 Methods of Meta-analysis: Correcting Error and Bias in Research Findings. Thousand Oaks, Calif.: Sage Publications.

Latessa, Edward J. and Lawrence T. Travis
1992 Residential Community Correctional Programs. In James Byrne, Arthur Lurigio, and Joan Petersilia (eds.), Smart Sentencing: The Emergence of Intermediate Sanctions. Newbury Park, Calif.: Sage.

Lipsey, Mark W. and David B. Wilson
2001 Practical Meta-analysis. Thousand Oaks, Calif.: Sage.

Lipsey, Mark W., Gabrielle L. Chapman, and Nana A. Landenberger
2001 Cognitive-behavioral programs for offenders. The Annals [of the American Academy of Political and Social Science] 578:144–157.

Lowenkamp, Christopher T.
2004 Correctional program integrity and treatment effectiveness: A multi-site, program-level analysis. The University of Cincinnati: Doctoral Dissertation.

Lowenkamp, Christopher T. and Edward J. Latessa
2002 Evaluation of Ohio's community based correctional facilities and halfway house programs. Technical Report. University of Cincinnati.

McGaha, John, Michale Fitcher, and Peter Hirschburg
1987 Felony probation: A re-examination of public risk. American Journal of Criminal Justice 12:1–9.

Nesovic, Aleksandra
2003 Psychometric Evaluation of the Correctional Program Assessment Inventory (CPAI). Carleton University: Doctoral Dissertation.

Palmer, Ted
1991 The effectiveness of intervention: Recent trends and current issues. Crime & Delinquency 37:330–346.
1994 A profile of correctional effectiveness and new directions for research. Albany, NY: SUNY Press.
1995 Programmatic and nonprogrammatic aspects of successful intervention: New directions for research. Crime & Delinquency 41:100–131.

Pearson, Frank S., Douglas S. Lipton, Charles M. Cleland, and Dorline S. Yee
2002 The effects of behavioral/cognitive behavioral programs on recidivism. Crime & Delinquency 48:476–496.

Petersilia, Joan
1985 Probation and felony offenders. Federal Probation 49:4–9.

Petersilia, Joan and Susan Turner
1993 Intensive probation and parole. In Michael Tonry (ed.), Crime and Justice: A Review of Research, Vol. 17. Chicago, Ill.: University of Chicago Press.

Quay, Herbert C.
1977 The three faces of evaluation. Criminal Justice and Behavior 4:341–354.

U.S. Department of Justice
1992 Recidivism of Felons on Probation, 1986-89. Washington, D.C.: U.S. Department of Justice.
2001 National correctional population reaches new high. Press Release. Washington, D.C.: U.S. Department of Justice.

Van Voorhis, Patricia, Francis T. Cullen, and Brandon K. Applegate
1995 Evaluating interventions with violent offenders: A guide for practitioners and policy makers. Federal Probation 59:17–28.

Vito, Gennaro
1986 Felony probation and recidivism: Replication and response. Federal Probation 50:17–25.

Wilson, David B., Leana C. Allen, and Doris Layton MacKenzie
2000 A Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders. George Mason University: Unpublished Manuscript.

Christopher T. Lowenkamp is Director of the Center for Criminal Justice Research and Assistant Research Professor in the Division of Criminal Justice at the University of Cincinnati. His research interests include the evaluation of correctional programming and macro-level criminological theory. Recent publications have appeared in the *Journal of Criminal Justice*, *Criminal Justice and Behavior*, *Federal Probation*, and *Journal of Research in Crime and Delinquency*.

Edward J. Latessa is a Professor and Head of the Division of Criminal Justice at the University of Cincinnati. He has published over 75 works in the area of criminal justice, corrections, and juvenile justice. He is co-author of seven books, including the tenth edition of *Corrections in America* and the third edition of *Corrections in the Community*. Dr. Latessa served as President of the Academy of Criminal Justice Sciences (1989-1990). He has also received several awards, including the Simon Dinitz Criminal Justice Research Award from the Ohio Department of Rehabilitation and Correction (2002), the Margaret Mead Award for dedicated service to the causes of social justice and humanitarian advancement by the International Community Corrections Association (2001), and the Peter P. Lejins Award for Research from the American Correctional Association (1999).

Paula Smith is Assistant Director of the Corrections Institute and Assistant Professor in the Division of Criminal Justice at the University of Cincinnati. Her research interests include offender classification and assessment, correctional rehabilitation, the psychological effects of incarceration, program implementation and evaluation, the transfer of knowledge to practitioners and policy makers, and meta-analysis. She has authored several articles, book chapters, and conference presentations on the above topics. Recent publications have appeared in *Criminal Justice and Behavior*, *Corrections Management Quarterly*, and *International Journal of Offender Therapy and Comparative Criminology*.

Crime & Delinquency
Volume 51 Number 1
January 2006 1-17

10.1177/0011128705281747
http://cad.sagepub.com
hosted at
http://online.sagepub.com

The Risk Principle in Action: What Have We Learned From 13,676 Offenders and 97 Correctional Programs?

Christopher T. Lowenkamp
Edward J. Latessa
University of Cincinnati, Cincinnati, OH

Alexander M. Holsinger
University of Missouri, Kansas City

Over the recent past there have been several meta-analyses and primary studies that support the importance of the risk principle. Oftentimes these studies, particularly the meta-analyses, are limited in their ability to assess how the actual implementation of the risk principle by correctional agencies affects effectiveness in reducing recidivism. Furthermore, primary studies are typically limited to the assessment of one or two programs, which again limits the types of analyses conducted. This study, using data from two independent studies of 97 correctional programs, investigates how adherence to the risk principle by targeting offenders who are higher risk and varying length of stay and services by level of risk affects program effectiveness in reducing recidivism. Overall, this research indicates that for residential and nonresidential programs adhering to the risk principle has a strong relationship with a program's ability to reduce recidivism.

Keywords: *risk principle; community corrections; program effectiveness*

The risk principle, which simply states that the level of supervision and treatment should be commensurate with the offender's level of risk, has been confirmed by research in corrections for more than a decade. The first mention of the risk principle by Andrews, Bonta, and Hoge (1990) was followed by a number of meta-analyses that confirmed and supported the importance of focusing on offenders who are higher risk (Andrews et al., 1990; Andrews & Dowden, 1999; Dowden & Andrews 1999a, 1999b, 2000;

Lipsey & Wilson, 1998; D. B. Wilson, Gottfredson, & Najaka, 2001; S. J. Wilson, Lipsey, & Derzon, 2003). Even though this research is fairly unequivocal, it is limited. Its limitations stem from the fact that meta-analyses are typically constrained in the measurement of offender risk to the use of an aggregate sample-level measure of risk; that is, most meta-analysts are forced to measure risk by using the percentage of the sample that has a criminal history or a history of a particular type of behavior. Furthermore, the meta-analyst often is not able to code and investigate the impact of adhering to the risk principle above and beyond measuring the percentage of the sample in a particular study that is higher risk.

The current research improves on earlier attempts to assess the importance of the risk principle by analyzing data from two separate studies. Collectively these studies provide data from 97 programs and a total of 13,676 individual offenders. The current investigation sought to answer the following questions: Are programs that adhere to the risk principle by providing more services and/or referrals for treatment to offenders who are higher risk more effective? Are programs that provide more services and/or supervision to offenders who are higher risk for a longer period of time more effective?

There is considerable empirical evidence that programs that target offenders who are higher risk are more effective in reducing recidivism than those that do not (Andrews et al., 1990; Andrews & Dowden, 1999; Dowden & Andrews, 1999a, 1999b, 2000; Lowenkamp & Latessa, 2005b); however, the questions still remain: Are there aspects of the risk principle that require specific actions by a correctional agency and are those actions meaningful when the appropriate targets for intervention have been selected?

Method

Because the data for the current investigation came from two distinct studies we review the participants from each data set separately. The program-level measures for each study are identical and are, therefore, discussed only once. Similarly, for analyses purposes, the data from Studies 1 and 2 were combined.

Study 1

The first set of data in the current analyses was developed from research conducted by Lowenkamp and Latessa (2002) and included offenders served by halfway houses (HWH) and community-based correctional facilities (CBCF) in Ohio. The HWH facilities receive offenders paroled from state

institutions or those who are placed under postrelease control (PRC), parole and/or PRC violators placed in a HWH as a sanction, and/or offenders released from a state institution under transitional control. All of the offenders in HWHs were reentering the community following a length of incarceration in a state institution. In contrast, the CBCF programs receive offenders placed under probation supervision. CBCF programs were initially designed to receive offenders who ordinarily would have been sent to prison but were given the opportunity to participate in rehabilitation services offered by the program. The sources of referral for the CBCF programs are the Courts of Common Pleas (i.e., offenders are sentenced directly to CBCFs) whereas the releasing authority (the Department of Corrections) or a parole officer makes HWH placements. The average length of stay (LOS) in the HWHs was 135 days, while the average LOS in the CBCFs was 137 days. The offenders were placed in the HWH programs as part of their PRC following a period of incarceration in a state institution. The offenders referred to a CBCF were placed on community control and sent to a CBCF from the court in lieu of a prison sentence.

Participants

The offenders placed in a HWH or CBCF program (3,782) were compared to parolees and other PRC offenders that were not placed in one of these residential programs. Each offender in the treatment group was matched to an offender in the comparison group based on the county of supervision, sex, and risk level using a modified version of the Salient Factor Score (SFS; Hoffman, 1983, 1994; Hoffman & Beck, 1974, 1985). Table 1 displays the original Salient Factor Score items and their respective weightings, and the slightly modified version used in the current research. The primary difference between the two scales was the use of employment at arrest, as opposed to the recent commitment free period.

Table 2 displays the number of CBCF and HWH programs included in the current study and the number of offenders served by each type of program (to calculate the total number of offenders for either group in any particular program, simply multiply the number of offenders listed in Table 2 by 2). As is indicated in Table 2, 15 of the programs included in the current study were CBCF facilities and 38 were HWHs. This represents a total of 53 programs that accounts for 55% of the programs included in the current analyses. These 53 programs also account for approximately 55% of the offenders in the treatment groups used in the current analyses.

Table 1
Risk Assessment Factors and Weights

Factors in Salient Factor Score (Hoffman, 1994)	Weight	Factors in Modified Salient Factor	Weight
Prior Convictions		Prior Arrest	
None	3	None	3
One	2	One	2
Two or three	1	Two or three	1
Four or more	0	Four or more	0
Prior Commitments > 30 days		Prior State or Federal Commitments	
None	2	None	2
One or two	1	One or two	1
Three or more	0	Three or more	0
Age at current offense		Age at current offense	
26 years or older	2	26 years or older	2
20 to 25 years	1	20 to 25 years	1
19 years or younger	0	19 years or younger	0
Recent commitment free period		Employed at arrest	
3 years since last offense	1	Employed	1
Otherwise	0	Unemployed	0
Probation/Parole/Escape state at offense		History of community control violations	
No criminal justice status at offense	1	None	1
Otherwise	0	One or more	0
Heroin and/or opiate dependence		History of drug use	
No history	1	History indicated	1
Otherwise	0	No history	0

Table 2
Distribution of Programs and Offenders in Study 1 and 2

Program Type	Program *n*	Offender *n*
Community-based correctional facilities[a]	15	1,791
Halfway house[a]	38	1,991
Day reporting[b]	7	412
Intensive supervision probation[b]	30	2,240
Work release[b]	3	206
Other[b]	4	198
Total	97	6,838

Note: a = Study 1.
b = Study 2.

Measures

When comparison groups were developed for each of the 53 programs in the current study, an *r* value was calculated between group membership (experimental vs. comparison) and recidivism (measured with this data as any incarceration in a state prison within 2 years of termination date from the program). Independent measures included a series of dummy variables that measured (a) whether two thirds of the offenders in a program's sample were higher risk (moderate or high risk on the modified SFS), (b) the difference in the average LOS between offenders who were lower risk and higher risk, (c) the difference in the average number of services and/or groups provided between offenders who were lower and higher risk, and (d) whether the program was rated as being cognitive behavioral.

The risk measure developed for the prior research involving this sample had four categories: low, low moderate, moderate, and high. For the purposes of the current research, we combined the moderate-risk and high-risk groups into a higher-risk group while the low-risk and low-moderate risk categories were combined into a lower risk group (offenders who were lower risk had a recidivism rate of 17% while offenders who were higher risk had a recidivism rate of 40%). If two thirds (66%) of the offenders in a program's sample were categorized as higher risk using this classification method, the program was assigned a code of 1. The cutoff proportion of two thirds was chosen based on prior research conducted by Lowenkamp and Latessa (2005a) and several meta-analyses that examined the effect of the presence of offenders who were high risk in the correctional treatment environment (Andrews & Dowden, 1999; Dowden & Andrews 1999a, 1999b; Lipsey & Wilson, 1998). This prior research indicated that an ample portion of the offenders engaged in treatment should be high risk. A percentage of offenders who are high risk at 50% was examined previously, yet the measures used to determine risk were somewhat limited (Lipsey & Wilson, 1998). Because of the more comprehensive measure of overall risk utilized in the current research, it was determined that two thirds would be a better benchmark for determining an offender population that was high risk. This benchmark may be more realistic as well, in light of common sentencing and placement practices (i.e., the items measured by the risk scale that was constructed are representative of offenders that typically would be sentenced to a secure environment).

To measure the extent to which a program adhered to the risk principle in terms of duration of treatment, we calculated the difference in the average LOS between the offenders who were lower risk and higher risk. Because the distribution for this measure was highly irregular, with extreme outliers at both ends, we dummy coded this measure. The measure was coded as a 1 if,

on average, the offenders who were higher risk stayed in the program longer than or equal to the offenders who were lower risk. We recognize that this measure is somewhat deficient given the differences in the overall program average LOS. Using these values, however, allowed us to make a determination that a program was minimally cognizant of the risk principle and at the very least did not violate the risk principle by keeping offenders who were lower risk in programming longer than offenders who were higher risk. While the measure may be crude, in the format used it serves as evidence as to whether the program clearly violated the risk principle in terms of program duration.

The next measure captured the difference in the average number of services and/or groups provided between the two categories of risk. If the program, on average, provided at least .5 more services or groups for offenders who were higher risk, the program was coded as a 1, otherwise, this variable was coded as a 0. For example, if Program X provided 1.5 groups on average for the offenders who were higher risk and 1.0 groups on average for offenders who were lower risk, Program X would be given a rating of 1. Services and groups refer to programming for such needs as substance abuse and education although not all services had to target criminogenic needs. Type of treatment was without question an important issue; however, the current research focuses primarily on whether the number of services, and the duration of services, vary by risk level of the offender. The decision to use .5 as the cutoff was somewhat arbitrary, however, we believed that anything less than one half of a service "unit" per offender seemed meaningless. As with program duration, ultimately this variable served as an indicator as to whether the program clearly provided at least a (potentially) meaningful higher amount of services to the offenders who were higher risk.

Each program was coded based on the reported treatment model. To gather this information, program staff were surveyed as to which type of treatment model guided programming. Those programs where at least two thirds of the staff reported that a cognitive behavioral or behavioral model guided programming were coded as a 1. All other programs were given a value of 0 for the treatment model variable. The purpose of this variable was to determine whether there was a meaningful likelihood that cognitive behavioral or behavioral models were the driving force behind the overall treatment modality. Cognitive behavioral therapies (CBT) are designed to specifically address the cognitions, thought patterns, and attitudes that underlie antisocial behavior. These therapies also utilize behavioral reinforcement techniques whereby rewards and consequences are used to solidify behavioral change. Treatment services such as CBT typically differ from "standard' correctional treatment in that most common treatment interven-

tions fail to address cognitions specifically, and fail to incorporate behavioral techniques. An example of standard correctional treatment would be unstructured group "talk therapy" discussion designed to share common negative experiences associated with drug use and addiction (among other types of treatment and/or counseling services).

Finally, a variable was coded to identify the treatment setting. Residential programs were given a value of 0 on this measure whereas nonresidential programs were given a value of 1.

Study 2

The second set of data come from another study conducted by Lowenkamp and Latessa (2005a) that investigated the effects of several nonresidential programs in the State of Ohio. A total of 44 programs were included in that research. The programs served prison diversion (offenders convicted of a felony where a state prison sentence is a possible penalty) and jail diversion (offenders convicted of a misdemeanor where incarceration in a jail is a possible penalty) offenders (for more information about the different groups of offenders see Lowenkamp & Latessa, 2005a). The typical program in the current study was an intensive supervision probation program, and while most of the programs were nonresidential (39), some were residential (5).

Participants

These 44 programs provided services to a total of 3,056 offenders. Offenders were matched to regular supervision probationers from each county or municipality running the program, or, for three programs, the offenders served were matched to jail inmates released during the same time period as the treatment-group offenders. Offenders were matched on jurisdiction, sex, and risk level using a risk measure developed from collected data.[1] Again, Table 2 shows the distribution of offenders across the different types of programs.

Measures

For 33 of the programs an *r* value was calculated between group membership (experimental vs. comparison) and recidivism (measured with this data as any incarceration in a state prison within 2 years of termination date from the program). For the remaining 11 programs, an *r* value was calculated between group membership and recidivism measured as any new arrest after

termination and/or release date. The jail diversion program participants were typically offenders who were lower risk who committed lower level offenses. Subsequent incarceration rates were too low to calculate reliable estimates of program effectiveness. All offenders were followed for 2 years posttermination or from their release date.

The other program-level measures (percentage offenders who were higher risk, measures indicating whether a program met the risk principle, and treatment model) included in the current research for this second set of data are identical to those discussed earlier. The residential programs from the current study were given a code of 0 on the setting measure. The nonresidential programs from this study were given a code of 1. There were a total of 57 residential programs and 40 nonresidential programs included in this sample.

Analyses

The measure of program effectiveness in the current investigation is the *r* value between group membership and recidivism. The *r* values reported in this research are correlation coefficients calculated for each program and represent the correlation between group membership and the outcome measure (any arrest for some programs and any incarceration for others although any incarceration was used for 86 of the 97 programs included in this study). The *r* values were transformed to Fisher's Z_r for all calculations (descriptive statistics and weighted least squares [WLS]) and then transformed back to standard form. Weights were used to take into consideration the differing numbers of offenders served by each program.

Several formulae were used for these transformations and the calculation of standard errors and weights.[2] For a more complete discussion, see Rosenthal (1991) and Lipsey and Wilson (2001).

Given the consistencies in the data and measures from Study 1 and 2, we combined the data for the two studies for analyses. To analyze the data, we calculated a series of WLS regression models.[3] Ultimately the model we ended with included the four independent measures reported earlier. These measures were used to predict the *r* values calculated for each program.

Results

The first set of analyses involved calculating descriptive statistics on the independent measures. The number of programs that met each of the principles associated with program effectiveness was low. Only 34 of the 97 programs were coded as using a CBT or other behavioral model. Only 26 of the programs provided more services and/or referrals for offenders who were

higher risk, and while almost one half the programs kept offenders who were higher risk in programming longer, more than one half kept offenders who were lower risk in programming longer than offenders who were higher risk. Barely one fifth of the programs had more than 65% offenders who were higher risk in their programs. The last measure simply identifies nonresidential programs. A total of 39 programs, or 40%, were nonresidential.

Focusing on the factors that are related to the content and operations of the program, it should come as no surprise that the majority of programs are failing to meet these criteria. Research involving the Correctional Program Assessment Inventory (CPAI; Gendreau & Andrews, 1994), which measures, among other things, the factors noted above, indicates that correctional programs fail miserably, as a group, when measured against the principles of effective interventions (Gendreau & Goggin, 2000; Hoge, Leschied, & Andrews, 1993; Latessa & Holsinger, 1999; Matthews, Jones Hubbard, & Latessa, 2001). So, our findings are consistent with similar research that investigated how closely a program adheres to the principles of effective intervention.

Overall, the 97 programs were associated with a slight increase in recidivism rates relative to the comparison groups ($r = -.03$). Although this increase is small, it is significant at the $p < .05$ level. Turning to the average r values by treatment setting, it was quickly observed that the residential programs were far more effective in reducing recidivism than the nonresidential programs. The residential programs were associated with an average reduction in recidivism of .03 while the nonresidential programs were associated with a substantial increase in recidivism ($r = -.12$). It is apparent from these data that the residential treatment programs were more effective than the nonresidential programs. Regardless, the impact of the risk principle on treatment effectiveness remains our primary concern and interest.

Our next analyses involved calculating a WLS regression model predicting the r values using the program characteristics and setting. The results of these analyses are contained in Table 3. First, note that the overall model is significant, $F(5, 91) = 8.106$, $p < .10$, with an adjusted R^2_{adj} of .27.

Starting with the first independent measure listed, Table 3 reveals that nonresidential programs were apparently much less effective than residential programs. This is not surprising given the differences in the type of programs. In general, the nonresidential programs would be electronic monitoring, day reporting, or intensive supervision. Programs of these types have, in the past, been shown to be associated with null or iatrogenic effects (Gendreau & Goggin, 1996). In contrast, the residential programs were developed to provide services to reduce offender risk (Lowenkamp & Latessa, 2002). The effects were not negligible—nonresidential programs

Table 3
Weighted Least Squares Results Predicting *r* Values

Variable	*b*	*SE*	*p*	β
Constant	−.05	.02	.01	.00
Nonresidential setting	−.09	.02	.00	−.27
Cognitive behavioral model	.04	.02	.05	.10
Higher risk offenders	.09	.03	.00	.21
Risk principle Tx	.06	.02	.00	.17
Risk principle supervision	.03	.02	.08	.12

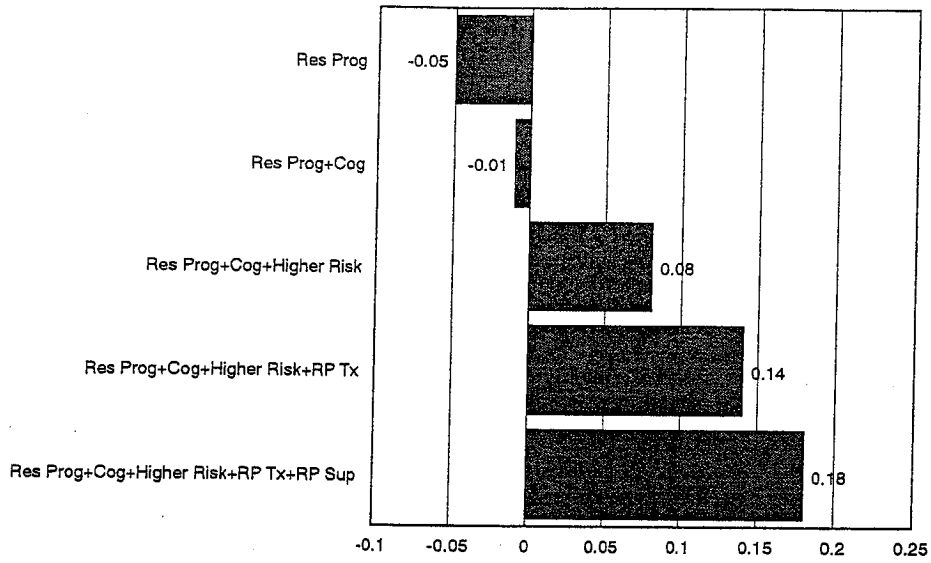
$F(5, 91) = 8.106, p < .10, R^2_{adj} = .27.$

were associated with an average reduction in recidivism that is 9 points smaller than for residential programs. Given that our dependent measure is an *r* value between group membership and a dichotomous outcome measure, *r* can be interpreted as the percentage point difference between the two groups in terms of the outcome measure (for greater detail, see the discussion on the binomial effect size display [BESD] in Rosenthal, 1991, and Lipsey & Wilson, 2001). Of greater importance and interest, however, is the fact that each of the program content factors is significant at $p < .10$.

The second measure in Table 3 controlled for the type of treatment the program reported to be the guiding philosophy or core of programming and/or services provided. The relationship between treatment type and program effectiveness was significant and in the direction indicating that those programs reported to be cognitive behavioral or behavioral were more effective than those reporting some other treatment modality.

The third measure in Table 3 captures whether 66% or more of the program's participants were higher risk. The relationship between this measure and outcome is the strongest for the substantive predictors. This finding is consistent with previous research, especially those meta-analyses that assessed the impact of the sample's risk level on outcome (Dowden & Andrews, 1999a, 1999b; Lipsey & Wilson, 1998). Also of importance is the fact that the two other measures implicated by the risk principle were significantly related to program effectiveness. Programs that provided at least .5 more units of service or referrals to offenders who were higher risk compared to offenders who were lower risk were more effective as were those that kept offenders who were higher -risk in the program as long as or longer than offenders who were lower risk. These effects are net the effects associated with treatment type and the risk composition of the offenders served by the program.

Figure 1
Predicted *r* Value Based on Adherence to Risk Principle and Treatment Type for Residential Programs

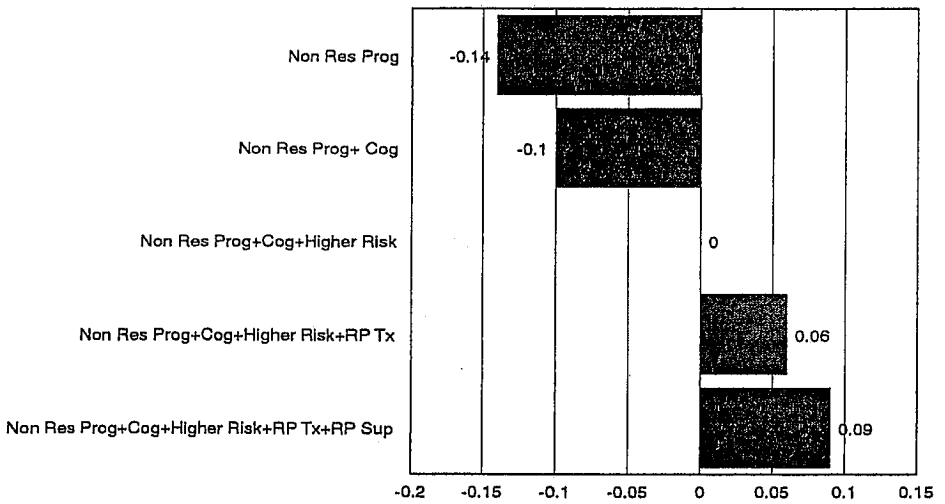


Note: Res Prog = Residential Program, Cog = Cognitive Behavioral or Behavioral Program, Higher risk = 66% or more of sample higher risk, RP Tx = more services for offenders who were higher risk, RP Sup = Longer or equal length of stay (LOS) for offenders who were higher risk.

Figure 1 illustrates the cumulative effects of meeting the criteria for the measures used in the WLS model. This first figure pertains to residential programs only and indicates that with the addition of each of the criteria a program's effectiveness continues to increase. Figure 1 indicates that a program's effectiveness climbs from an *r* value of –.05 (indicating an increase in recidivism rates) to an *r* value of .18 (indicating an 18-percentage-point reduction in recidivism rates relative to the comparison group) as a program continues to meet each of the criteria specified. Targeting offenders who are higher risk continues to be an important factor as indicated in Figure 1; however, there are apparently other factors that are of importance too. Each of the factors under investigation contributes substantive and significant increases in a program's effectiveness.

Figure 2 displays the impacts of the differing factors for nonresidential programs. Note that in general the nonresidential programs are not as effective as the residential programs ;although, based on the WLS model, neither type of program was associated with a reduction in recidivism on average. However, it is again the case that with nonresidential programs, effectiveness

Figure 2
Predicted *r* Value Based on Adherence to Risk Principle and Treatment Type for Nonresidential Programs



Note: Non Res Prog = Nonresidential Program, Cog = Cognitive Behavioral or Behavioral Program, Higher risk = 66% or more of sample higher risk, RP Tx = more services for offenders who are higher risk, RP Sup = Longer or equal length of stay (LOS) for offenders who were higher risk.

in reducing recidivism was achieved when the factors implicated by the risk principle were followed.

The results of these analyses, taken together, show a consistent pattern. The correctional programs included in these analyses, whether residential or nonresidential showed increases in recidivism rates unless offenders who were higher risk were targeted and provided more services for a longer period of time.

Discussion

Traditionally, outcome studies of correctional interventions and programs provided limited direction for correctional practitioners. However, more recent research by Andrews (1999), Gendreau (1996), and others (Lipsey, 1992, 1999a, 1999b; Palmer, 1995; S. J. Wilson et al., 2003) have led to the formulation of some important principles, one of which is the risk principle. This principle states that our most intensive correctional treatment and inter-

vention programs should be reserved for offenders who are higher risk (Andrews et al., 1990). *Risk* in this context refers to those offenders with a higher probability of recidivating. Placing offenders who were lower risk in structured programs (whether treatment or supervision oriented) clearly demonstrates that recidivism can actually be increased (Andrews & Dowden, 1999; Bonta, Wallace-Capretta, & Rooney, 2000; Hanley, 2002; Lowenkamp & Latessa, 2002). There are several possible reasons for this.

First, placing offenders who are lower risk with offenders who are higher risk provides an environment in which individuals who are lower risk learn antisocial behavior that is modeled for them, and form new peer associates, many of whom are more likely to support and reinforce criminal behavior. Second, placing offenders who are lower risk in these programs also tends to disrupt their prosocial networks; in other words, the very attributes that make them lower risk become interrupted, such as school, friendships, employment, family, and so on (Lowenkamp & Latessa, 2004). Third, increased supervision, along with more stringent conditions (such as frequent drug testing), increases the likelihood that violations will occur.

The results from this study indicated that even when some form of CBT is provided it is not sufficient. Offenders who are higher risk must also be provided more services and kept in programming longer to have appreciable effects on outcome. Based on these findings the following recommendations are in order:

- Correctional programs need to utilize objective and standardized assessment tools to identify appropriate offenders for highly structured programs. Although we did not investigate the relationship that this practice has with program effectiveness in the current study, prior research indicates that standardized and actuarial assessments are the best method to use for accurate prediction of offender risk (Bonta, Law, & Hanson, 1998; Grove, Zald, Lebow, Snitz, & Nelson, 1995; Hanson & Bussiere, 1998). Without such assessments, programs would likely target the wrong offenders.
- Length of programming and supervision needs to be clearly tied to levels of risk. Offenders who are lower risk are best served with more traditional levels of supervision, while offenders who are higher risk should be kept in programming longer to address their risk factor and needs. While this concept seems straightforward, very few programs in this study met this principle. Further, unpublished data (Latessa, 2005) on 362 CPAI assessments indicates that only 7% of the programs assessed vary the intensity of programming by risk level and only 2% vary duration by risk level.
- Offenders are not higher risk because they have a particular risk factor, but rather because they have a multitude of risk factors. Accordingly, a range of services and interventions should be provided that target the specific crime-

producing needs of the offenders who are higher risk. Multiple services are required for offenders who are higher risk.

- Obviously there are a number of factors that should be considered when sentencing offenders, including severity of offense, harm to the victim, and other mitigating and aggravating circumstances. However, this research has some clear implications for sentencing, especially when judges are considering conditions for supervision. To have the greatest impact on recidivism, length of supervision and services provided should be clearly tied to an offender's risk level. Sentencing guidelines may often provide difficulty in implementing any number of effective correctional practices.
- To tie sentencing and related decisions to risk level, judges (and postsentencing agencies) need to utilize a validated risk assessment method that meaningfully differentiates between offenders who are high risk and low risk. As important, sentencing judges need to have at their disposal correctional intervention options that are appropriate for the risk level of the offenders being processed. In turn, correctional agencies (those that are strictly supervisory and/or control oriented, and those that offer rehabilitative services) will benefit from internally incorporating the risk principle whenever possible. The results of the research presented above demonstrate the increased effectiveness of programs and agencies that ensure those that need the most, receive the most. Although sentencing takes a multitude of factors into account, benefits may be gleaned by ensuring that judges have a variety of valid assessment information at their disposal, and know how to fully utilize it.

As with any research study, there are limitations to this research. First, the programs investigated include only programs from Ohio and only programs that serve adult offenders. Second, the outcome measure used, for the majority of programs, was limited to return to prison for any reason. Third, the data come from studies that used quasi-experimental designs. Fourth, the follow-up period for recidivism was limited to 2 years. Even with these limitations, this research provides important information that (a) confirms the fact that very few correctional programs are meeting the risk principle when assessing adherence with data on services provided and LOS, (b) indicates that programs that do adhere to the risk principle are apparently more effective than those that do not, (c) directs future researchers interested in assessing the importance of the risk principle, and (d) helps correctional programs in making changes that might increase their effectiveness in reducing the recidivism of offenders they serve.

Notes

1. This score included the following factors: arrest history, felony arrest history, incarceration history, violent offense history, sex offense history, drug problems, alcohol problems, employment status at arrest, age, marital status, current offense type, current offense level, and history of or current community supervision violations. The scoring of this measure, the cutoff scores for the risk categories, and the recidivism rates for those categories are contained in Lowenkamp and Latessa (2005b). In summary, however, the risk score was composed of 13 factors with a range of 0 to 15 with a mean of 7.4. The correlations between the risk score and any incarceration and any arrest was .35 and .31, respectively.

2. Formula 1 (r to Zr): $Z_r = .5 \ln\left[\frac{1+r}{1-r}\right]$ Where r = the correlation coefficient and ln = the natural logarithm (e).

Formula 2 (Z_r to r): $r = \frac{e^{2Z_r} - 1}{e^{2Z_r} + 1}$ Where Z_r = the Fisher transformed value of r and e = approximately 2.718.

Formula 3 (calculation of standard error): $se = \sqrt{\frac{1}{n-3}}$ Where N equals the total number of cases.

Formula 4 (calculation of weight for analyses): $w = \frac{1}{se^2}$

3. The WLS models were estimated using SPSS syntax developed and presented by Lipsey and Wilson (2001).

References

Andrews, D. A. (1999). Assessing program elements for risk reduction: The Correctional Programme Assessment Inventory (CPAI). In P. M. Harris (Ed.), *Research to results: Effective community corrections. Proceedings of the 1995 and 1996 Conferences of the International Community Corrections Association (ICCA)* (pp. 151-170). Lanham, MD: American Correction Association.

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

Andrews, D. A., & Dowden, C. (1999). A meta-analytic investigation into effective correctional intervention for female offenders. *Forum on Corrections Research, 11*, 18-21.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology, 8*, 369-404.

Bonta, J., Law, M., & Hanson, R. K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin, 39*, 127-144.

Bonta, J., Wallace-Capretta, S., & Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior, 27*, 312-329.

Dowden, C., & Andrews, D. A. (1999a). What works for female offenders: A meta-analytic review. *Crime & Delinquency, 45*, 438-452.

Dowden, C., & Andrews, D. A. (1999b). What works in young offender treatment: A meta-analysis. *Forum on Corrections Research, 11*, 21-24.

Dowden, C., & Andrews, D. A. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology, 42*(4), 449-467.

Gendreau, P. (1996). The principles of effective interventions with offenders. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 117-130). Thousand Oaks, CA: Sage.

Gendreau, P., & Andrews, D. A. (1994). *The Correctional Program Assessment Inventory.* St. John, Canada: University of New Brunswick and Carleton University.

Gendreau, P., & Goggin, C. (1996). Principles of effective correctional programming. *Forum on Corrections, 8*, 38-41.

Gendreau, P., & Goggin, C. (2000). Correctional treatment: Accomplishments and realities. In P. Van Voorhis, M. Braswell, & D. Lester (Eds.), *Correctional counseling and rehabilitation* (4th ed., pp. 271-280). Cincinnati, OH: Anderson.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (1995). *Clinical vs. mechanical prediction: A meta-analysis.* Unpublished manuscript.

Hanley, D. (2002). *Risk differentiation and intensive supervision: A meaningful union?* Unpublished doctoral dissertation, University of Cincinnati.

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

Hoffman, P. B. (1983). Screening for risk: A revised Salient Factor Score (SFS-81). *Journal of Criminal Justice, 11*, 539-547.

Hoffman, P. B. (1994). Twenty years of operation use of a risk prediction instrument: The United States Parole Commission's Salient Factor Score. *Journal of Criminal Justice*, 22, 477-494.

Hoffman, P. B., & Beck, J. L. (1974). Parole decision-making: A Salient Factor Score. *Journal of Criminal Justice*, 2, 195-206.

Hoffman, P. B., & Beck, J. L. (1985). Recidivism among released federal prisoners: Salient Factor Score and five year follow-up. *Criminal Justice and Behavior, 12*(4), 501-507.

Hoge, R., Leschied, A., & Andrews, D. A. (1993). *An investigation of young offender service in the province of Ontario: A report of the repeat offender project.* Toronto, Canada: Ministry of Community and Social Services.

Latessa, E. J. (2005). *Unpublished data from 362 CPAI Assessments.* Cincinnati, OH: University of Cincinnati, Division of Criminal Justice.

Latessa, E. J., & Holsinger, A. M. (1999). The importance of evaluating correctional programs: Assessing outcome and quality. *Corrections Management Quarterly*, 2, 22-29.

Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variable of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, et al. (Eds.), *Meta-analysis for explanation: A casebook* (pp. 83-127). New York: Russell Sage.

Lipsey, M. W. (1999a). Can intervention rehabilitate serious delinquents? *Annals of the American Academy of Political and Social Science, 564*, 142-166.

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

Lipsey, M. W., & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Lowenkamp, C. T. (2004). *Correctional program integrity and treatment effectiveness: A multi-site, program-level analysis*. Unpublished doctoral dissertation, University of Cincinnati.

Lowenkamp, C. T., & Latessa, E. J. (2002). *Evaluation of Ohio's community based correctional facilities and halfway house programs* [Technical report]. Cincinnati, OH: University of Cincinnati.

Lowenkamp, C. T., & Latessa, E. J. (2004). Understanding the risk principle: How and why correctional interventions can harm low-risk offenders [Technical Report]. *Topics in community corrections, 2004* (pp. 3-8). Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Lowenkamp, C. T., & Latessa, E. J. (2005a). *Evaluation of Ohio's CCA funded programs* [Technical report]. Cincinnati, OH: University of Cincinnati.

Lowenkamp, C. T., & Latessa, E. J. (2005b). Increasing the effectiveness of correctional programming through the risk principle: Identifying offenders for residential placement. *Criminology and Public Policy, 4*(2), 263-290.

Matthews, B., Jones Hubbard, D., & Latessa, E. J. (2001). Making the next step: Using evaluability assessment to improve correctional programming. *The Prison Journal, 81*, 454-472.

Palmer, T. (1995). Programmatic and nonprogrammatic aspects of successful intervention: New directions for research. *Crime & Delinquency, 41*, 100-131.

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

Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: A meta-analysis. *Journal of Quantitative Criminology, 17*(3), 247-272.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive and disruptive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology, 71*(1), 136-149.

Christopher T. Lowenkamp is an assistant research professor at The University of Cincinnati, Division of Criminal Justice. His research interests include risk assessment and the evaluation of correctional interventions.

Edward J. Latessa is a professor and head of the Division of Criminal Justice at the University of Cincinnati. He has published more than 75 works in the area of criminal justice, corrections, and juvenile justice. He is coauthor of seven books including *Corrections in the Community*, which is now in its 3rd edition. He has directed more than 60 funded research projects including studies of day reporting centers, therapeutic communities, juvenile justice programs, drug courts, intensive supervision programs, halfway houses, and drug programs.

Alexander M. Holsinger currently serves as associate professor of criminal justice and criminology at the University of Missouri-Kansas City. His research interests include correctional treatment effectiveness, offender assessment and classification, correctional program evaluation, and violent crime reduction. His written work has appeared in several academic criminal justice journals including *Criminal Justice & Behavior*, *Journal of Criminal Justice*, *Crime & Delinquency*, *Journal of Crime and Justice*, and several professional journals as well.

[running head: TRAINING TO SEE RISK]

Training to See Risk: Measuring the Accuracy of Clinical and Actuarial Risk Assessments among Federal Probation Officers

J.C. Oleson
University of Auckland

Scott VanBenschoten
Charles Robinson
Christopher T. Lowenkamp
Administrative Office of the United States Courts

THE PREDICTION OF RISK is ubiquitous in modern society (Beck, 1992). Physicians consider risk when treating patients, financiers consider risk when making investments, and psychologists consider risk when working with clients. Within the criminal justice system, predictions of risk guide discretion at all points (Gottfredson & Tonry, 1987). When police officers choose between formal citations and verbal warnings, they evaluate risk; when judges impose sentences upon defendants, *they* evaluate risk; and when community corrections officers monitor the conditions of pretrial defendants, parolees, and probationers, they, too, evaluate risk.

Over time, research suggests, professionals within the justice community develop the ability to distinguish high-risk offenders from those who present little risk of reoffending (Fong, et al., 1990; Mossman, 1994). They do so by drawing upon their own personal experiences, using heuristics and other mental shortcuts to simplify complex calculations (Nisbett & Ross, 1980). But this kind of professional (or clinical) judgment is limited to the experience of the decision maker and is subject to a host of faults: unreliable evaluations, discretionary decisions based upon biases and stereotypes, and politicized administration (Walker, 1993). An alternative approach is to use statistically-derived instruments to predict actuarial risks of violence, dangerousness, reoffending, rearrest, or reconviction.

The statistical prediction of recidivism risk has an 80-year history, and can be traced at least as far back as the 1928 parole prediction instrument developed by Ernest Burgess (Burgess, 1928). Early attempts to use actuarial risk assessment in the justice system were often controversial, particularly

given high rates of false positives (Selective Incapacitation, 1982). Evaluators identifying subjects as dangerous were wrong twice as often as they were right (Monahan, 1981). Nevertheless, despite these flaws, research suggested that actuarial prediction outperformed the clinical judgment of even trained professionals across an array of disciplines (e.g., Meehl, 1954). The superiority of actuarial assessment over unstructured clinical judgment is a finding that has been replicated by many researchers (Grove & Meehl, 1996; Harris, 2006). One meta-analysis of 136 studies concluded that statistical predictions were 10 percent more accurate than clinical judgments and were dramatically more accurate one third of the time (Grove, et al., 2000). The accuracy of assessment instruments also appears to have improved (Hilton, et al., 2006). A more recent meta-analysis of 67 studies concluded that actuarial assessment generally is 13 percent more accurate than clinical judgment and is 17 percent more accurate in predictions of future violent or criminal behavior (Ægisdóttir, et al., 2006).

Today, the academic debate is no longer about whether actuarial assessments out-predict clinical judgments; *that* debate is long since over (Monahan, et al., 2001). Even the skeptics of actuarial risk prediction now acknowledge a consensus that actuarial judgments consistently outperform clinical ones (Harcourt, 2007; Litwack, 2001). Instead, the current debate is about whether there is any place in risk assessment for *clinical* judgment (Hanson, 2009). Some researchers argue for a synthetic approach, combining actuarial and clinical techniques (e.g., Gottfredson & Moriarty, 2006; Sjöstedt & Grann, 2002, Sreenivasan, et al., 2000). After all, for all its strengths, actuarial prediction is not particularly good at accounting for exceptional circumstances, predicting rare events, or predicting risk for young people (for whom three is less historical information available) (Bullock, 2011). Other researchers, however, argue for an actuarial-only approach (e.g., Grove & Meehl, 1996; Quinsey, et al., 1998). They claim that the introduction of clinical judgment only reduces the accuracy of the instrument. And after all, "[e]ven if actuarial methods merely equal the accuracy of clinical approaches, they may save considerable time and expense" (Dawes, et al., 1989: 1673).

Numerous commercial risk assessment instruments are available, all predicting recidivism about equally, all more accurate than unstructured clinical judgment (Yang, et al, 2010). Many jurisdictions use commercial instruments such as the PCL-R, CAIS, COMPAS, or the LSI-R. Other jurisdictions have adapted off-the-shelf instruments to fit their specific needs or have developed their own in-house assessment tools. Used effectively, these assessment tools allow probation officers to accurately assess risk, a requisite first step in employing evidenc e-based practices (Harris, 2006; VanBenschoten, 2008).

Yet despite the lengthy history of statistical risk assessment and despite a substantial body of research demonstrating that actuarial predictions outperform unstructured clinical judgment, probation officers – both in the United States and abroad – have exhibited skepticism, ambivalence, and outright hostility toward actuarial assessment devices. Irish probation officers have cultivated an attitude of "resistance" to assessment instruments (Fitzgibbon, et al., 2010). In England, Horsefield suggested that, using their clinical judgment, "it is not difficult for probation service staff to identify who is likely to commit further offences" (2003: 377), and argued that the *real* value of using actuarial risk instruments lies in justifying the operations within the probation service, competing for resources, and regulating staff behavior. In the United States, Schneider and her colleagues (1996) reported similar attitudes among Oklahoma probation officers. Officers held negative-to-neutral views about risk instruments (e.g., only 15 percent thought risk instruments are more accurate than officer judgment) but thought actuarial tools *were* useful in justifying supervision levels to the public and legislature. Lynch (1998) reported that California parole officers deliberately subverted directives issued by their actuarial risk managers. But even managers appear to express reservations about the value of risk assessment instruments. In a 2003 national survey of community corrections agencies, 61 percent of respondents described themselves as satisfied or very satisfied with the risk instruments used in their departments, but a full 39 percent described themselves as neutral, uncertain, or dissatisfied (Clem, 2003: 22).

Background and Research Question

The tension between professional judgment and actuarial risk assessment affects the federal probation and pretrial services system as well. Risk assessment is not new to the federal courts. The district court for the District of Columbia began using a risk prediction scale, the "U.S.D.C. 75," in 1970 (Hemple, et al., 1976). This instrument was renamed the Risk Prediction Scale 80 (RPS 80) and adopted for use throughout the probation system in January of 1981 (Eaglin & Lombard, 1981). In September of 1997, the RPS 80 was replaced by the Risk Prediction Index (RPI), an eight-question, second-generation risk assessment tool (Lombard & Hooper, 1998). But many probation officers did not use the RPI scores they calculated (VanBenschoten, 2008), and did not always link supervision practices to risk levels (Lowenkamp, et al., 2006).

Responding to the Criminal Law Committee's endorsement of evidence-based practices in the supervision of defendants and offenders (Judicial Conference, 2006), probation staff at the Administrative Office of the U.S. Courts have developed a new, third-generation risk assessment instrument, the Federal Post Conviction Risk Assessment (PCRA). The PCRA was validated on a large sample of federal probation cases (**see** article in this issue by **Johnson et al.)**

We were interested in whether use of the PCRA would improve the ability of federal probation officers to accurately assess risk. On the one hand, 50 years of research suggests that actuarial prediction consistently outperforms unstructured professional judgment (e.g., Ægisdóttir,et al., 2006; Grove & Meehl, 1996; Grove, et al., 2000; Monahan, et al., 2001); on the other hand, federal probation officers are considered to be the "crème de la crème" of community corrections officers (Buddress, 1997: 6). They are well educated, well trained, and often come to the federal system with substantial practical experience. Would the use of the PCRA allow even federal officers to improve their ability to assess risk?

Methods

The question of whether the use of the PCRA would improve the risk assessment skills of federal probation officers was investigated during four regional training meetings convened during 2010 and 2011. Federal probation officers from districts in the greater Washington, DC metropolitan region gathered in Washington, DC to participate in PCRA training; officers from districts in the eastern United States gathered in Charlotte, NC; officers from districts in the middle of the country gathered in Detroit, MI; and officers from districts in the western United States, including Pacific islands, gathered in Salt Lake City, UT. Approximately 150-350 officers attended each of the training meetings.

Prior to the training session each officer was asked to complete an eight hour online training program that reviewed the fundamentals of risk, need and responsivity (Andrews, et al., 1990). At each session, trainers explained to the participating officers that they would be asked to assess an offender's risk based on a videotaped mock intake interview and supplementary written documentation. Specifically, they were told that they would be asked to place the offender in the case vignette in one of four risk categories (, low, low/moderate, moderate, or high) and to identify the offender's three most important criminogenic needs (in rank order). The description of the risk levels were not defined, thus the officers needed to define for themselves what each risk level meant. Although the probation officers were in a large group setting, the trainers emphasized that officers were not to discuss their rankings of risk or identification of criminogenic needs until they submitted their data collection form.

The case vignette consisted of a 24-minute mock intake interview (based upon an actual case, with identifiers and key case details modified in order to protect the offender's anonymity). The probation officer in the vignette asked the offender – a man in his fifties with a long history of methamphetamine addiction and firearms charges – a series of questions about the offender's criminal behavior, employment, social networks, cognitions, substance abuse, time in custody, and current

accommodations. Supplemental written materials included a presentence report and release paperwork from the Federal Bureau of Prisons.

The offender in the vignette was working and lived in a stable residence. He participated in treatment, remained free of drug use, and could articulate a relapse prevention plan. He did not associate with anti-social peers and was in the process of developing a pro-social network. The correct score, according to the PCRA, was low/moderate risk. Specifically, the numerical score was 6 (PCRA scores range between 0 and 18).

After the video concluded, officers were given as much time as needed to identify the risk level and three top criminogenic needs. Officers typically took between five and ten minutes to review the supplementary material and submit a complete data collection form. They were not provided with the correct score after this first exercise.

On the second day of the training, after learning the scoring rules of the PCRA and practicing on several scenarios, probation officers viewed the training vignette for a second time. Instead of using their professional judgment to identify the offender's risk level and criminogenic needs, they were asked to use the PCRA and identify a risk score. The officers were shown the same video and were provided with the same written supplementary materials. Once again, they were asked to score the case independently and to provide their answers to the trainers. These actuarial (PCRA) risk assessments were collected and compared with the risk assessments made with clinical judgments.

Risk category (, low, low/moderate, moderate, or high) is an ordinal variable. As such, typical measures of central tendency and measures of dispersion do not apply. We were, however, interested in whether officers can accurately assign the offender into the correct risk category unaided by actuarial risk assessment and if the PCRA increases the reliability of the assessment of risk and thereby risk classification. To evaluate the effect of the PCRA on reliability of risk assessments, we used the consensus measure (Tastle, et al., 2005) to measure dispersion. The consensus measure is a measure

that ranges in value from 0 to 1, with 1 representing complete agreement among those that ranked an item (in our case risk category) regardless of the category chosen. A value of 0 on the consensus measure (complete dissention) would be the result when two equal groups of participants rank a case at the far ends of the scale. This characteristic of the consensus measure is important, as it allowed us to determine whether the officers' categorization of risk was consistent, regardless of whether or not their assessments agreed with the results of the PCRA.

Results

A total of 1,087 officers identified a risk category for the case vignette when asked to do so without administering the PCRA. A total of 1,049 officers provided a risk categorization for the case vignette using the PCRA. The distributions of these ratings are presented in the following two figures.

Figure 1 displays the frequencies (percentages are in parentheses) of risk categories identified by the officers using clinical judgment (without the use of the PCRA). As indicated in Figure 1, it is clear that the largest identified category of risk for the case vignette is moderate risk. Just over 50 percent of the officers indicated, based on the information provided, that the offender was moderate risk. Thirty percent of the officers identified the offender's risk level as low/moderate, while 17 percent identified the risk level as high. A much smaller percentage (2 percent) identified the offender's risk level as low. Given this distribution of scores, a calculation of the consensus measure (Cns) yielded a value of 0.66.

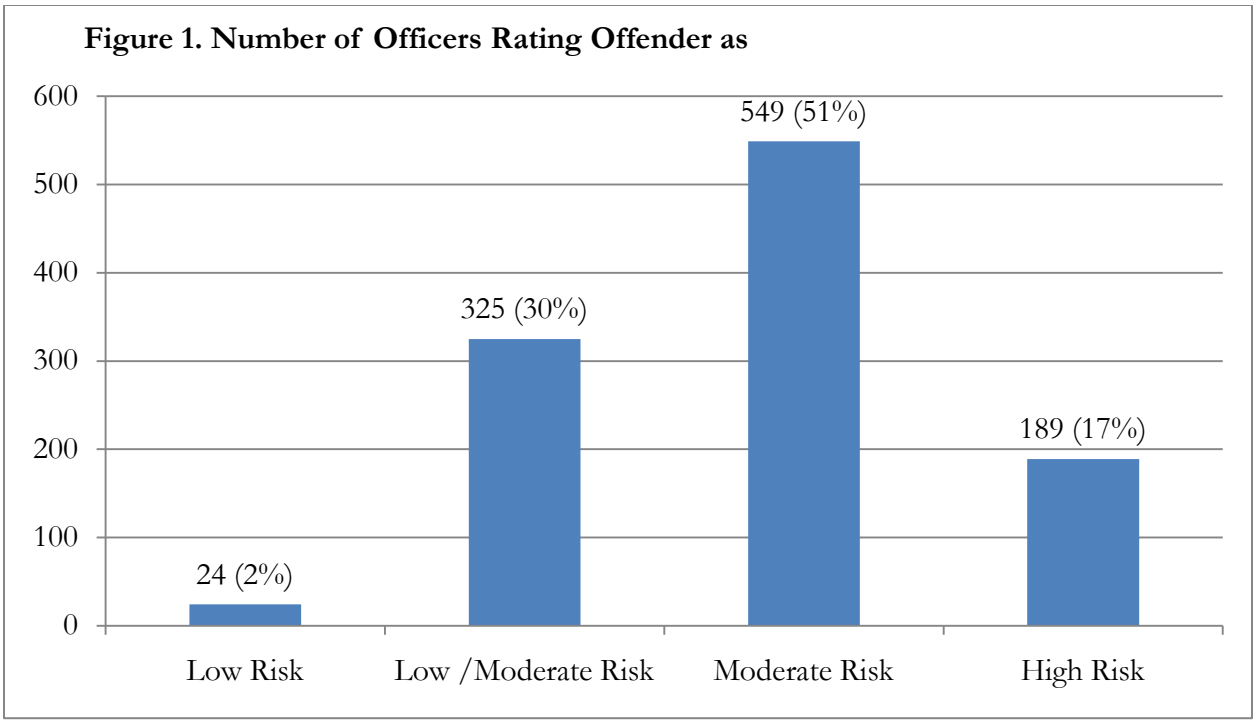
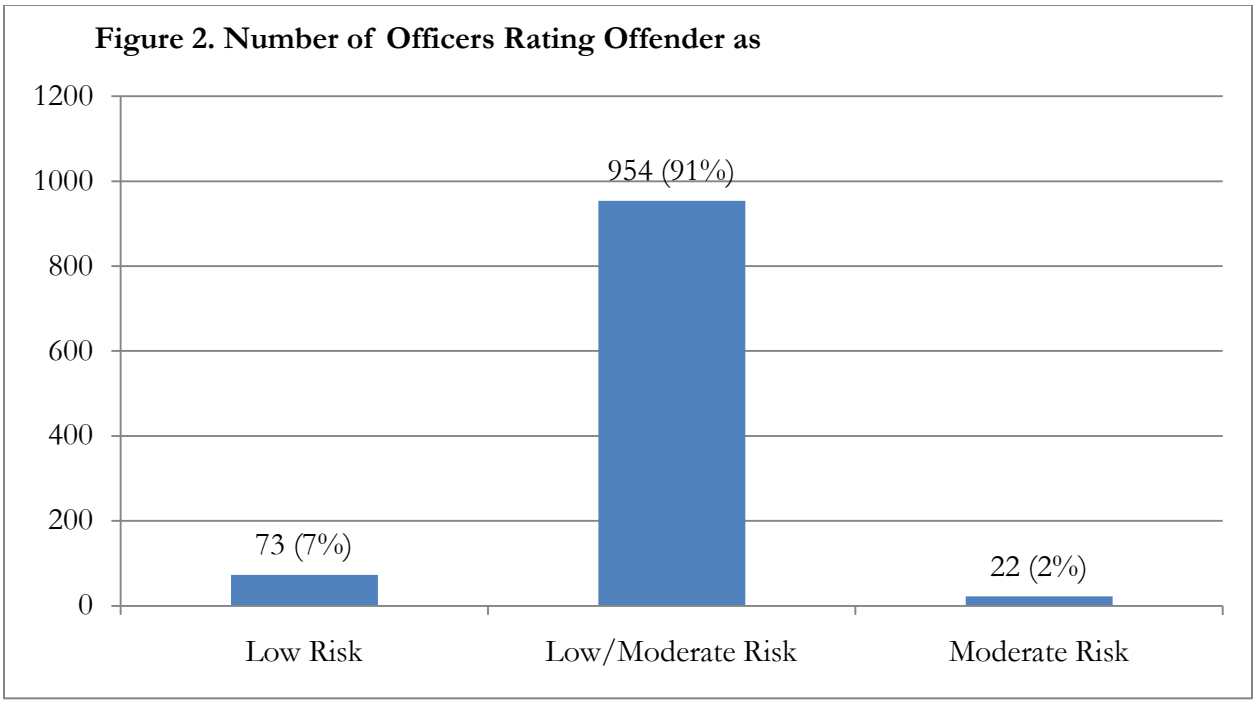


Figure 2 displays the distribution of risk categories assigned by officers when using the PCRA to guide their determination of risk. Note that in Figure 2, only three bars indicate the estimation of risk. No officers identified the offender's risk level as high when using the PCRA. A second noteworthy feature of Figure 2 is that the largest category of risk identified by the officers accounts for ratings from 954, or 91 percent, of the officers. The consensus measure based on the distribution of these ratings yielded a Cns value of .93, or about 1.4 times as great as the Cns measure yielded from the distribution of ratings in Figure 1. In addition, the officers selected the proper risk category, according to the PCRA, 91 percent of the time. Given that this was only these probation officers' first or second administration of the PCRA, these results are encouraging.



Discussion

Federal probation officers made more consistent and more accurate assessments of offender risk when using the PCRA than when using unstructured clinical judgment. Assessments made with the PCRA were more accurate (e.g., more officers correctly identified the risk level) and had greater consensus (e.g., even officers who did not correctly identify the risk level selected categories adjacent to the actual risk level). These findings support the view that, in assigning offenders to the correct risk category, actuarial prediction outperforms unstructured clinical judgment. Our findings are consistent with a robust body of work, collected over many decades (e.g., Ægisdóttir, et al., 2006; Grove & Meehl, 1996; Grove, et al., 2000; Meehl, 1954; Monahan, et al., 2001). But they are still remarkable. It is important to note that Federal Probation officers have very high. They must meet medical standards, pass regular background investigations, possess at least a bachelor's degree from an accredited university, and complete the six-week training program at the Federal Probation and Pretrial Services Training Academy. Typically these officers have prior probation experience from other jurisdictions. Additionally, these highly-skilled

professionals are part of a single system with one set of national policies (with local variation) and a uniform training academy. Despite this, the federal probation officers produced a more consistent risk level assignment with the use of a actuarial tool

The research also indicates that clinical judgments tended to overestimate risk. It is not difficult to understand why. Ansbro notes that probation officers "face the mutually-exclusive targets of high accuracy and high throughput, and exist in a climate where failings in practice will be hunted for if an offender commits a serious offence whilst on supervision" (2010: 266). A signal detection analysis lies beyond the scope of this article, but in a situation where there are dire consequences to missing a true positive (i.e., not identifying a high-risk offender as such) and few direct costs to officers when making false positives (i.e., wrongly identifying a low-risk offender as high-risk), it is easy to see why officers would yield to the so-called precautionary principle identified by Kemshall (1998). Of course, over-supervising low-risk offenders is expensive, and diverts resources away from the high-risk offenders who need them. Austin analogizes this to a "hospital that decides to provide intensive care for patients who have a cold – the treatment is not only unnecessary but expensive" (2006: 63). There is also research suggesting that over-supervising low-risk offenders can make them *worse*, affirmatively increasing their likelihood of recidivism (Lowenkamp & Latessa, 2004). Actuarial tools can serve as a valuable check against the precautionary principle. They can provide a means of engaging in professional triage, ensuring that resources are allocated where they should be, maximizing community safety while allowing for offender rehabilitation (Flores, et al., 2006).

It is also important to note that specific descriptions of the risk terms were not defined for the officers. This may have caused some of the risk category assignment variation. What "low risk" means to one officer may mean something different to another. This means that the variation in risk assignment may be due to how the case is seen and understood by an officer, but equally concerning is

that the difference may also be due to various definitions of language that officers and national policy use related to risk.

In a landmark article, Feeley and Simon suggested that the rise of risk assessment was symptomatic of a shift to a new penology: "[T]he new penology is markedly less concerned with responsibility, fault, moral sensibility, diagnosis or intervention and treatment of the individual offender. Rather, it is concerned with techniques to identify, classify, and manage groupings sorted by dangerousness. The task is managerial, not transformative" (1992: 452). Without question, the use of risk assessment instruments in community corrections has exploded since Feeley and Simon published their article, and its ascendance has been criticized by many thoughtful critics (e.g., Hannah-Moffat, et al., 2009; O'Malley, 2004; Wandall, 2006). Indeed, Harcourt (2007) demonstrates that risk-based justice may actually *increase* the overall amount of crime in society. In jurisdictions around the world, probation and parole officers have resisted the tyranny of risk and rejected managers' instructions to manage offenders under their supervision by risk score (Fitzgibbon, et al., 2010; Lynch, 1998). But this view of risk assessment may be too dystopian. Other commentators have realized that the consequences of risk assessment are far more nuanced than its critics suggest. For example, Robinson (2002) notes that actuarialism's focus on outcomes actually underlies the new rehabilitation of "what works" (see Petersilia, 2004; Taxman, et al., 2004). To be sure, this is a form of rehabilitation that takes public safety as its ultimate object – not the transformation of every individual offender (Robinson, 2002). But instead of contributing to an inexorable increase in prison populations and persons under supervision – a population that exceeded five million, or 1 in 31 U.S. citizens, during 2009 (Pew Center, 2009) – risk assessment can *reduce* prison and community corrections populations (Bonta, 2008). By operating as a check against the precautionary principle and reducing over-classification, actuarial risk assessment can reduce recidivism among low-risk offenders by ensuring that they are not over-supervised. It can simultaneously reduce recidivism among high-risk offenders by ensuring that these

individuals are carefully supervised and provided with interventions that correspond to their criminogenic needs. Instead of stripping the humanity from probation work (Wandall, 2006), actuarial risk assessment with the PCRA can allow federal probation officers to be far more effective in facilitating real transformative change in the lives of offenders.

References

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *Counseling Psychologist*, 34(3), 341-382.

Andrews, D. A., Bonta, J., & Hoge, R. D. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, *17*, 19–52.

Ansbro, M. (2010). The nuts and bolts of risk assessment: When the clinical and actuarial conflict. *Howard Journal*, *49*(3), 252-68.

Austin, J. (2006). How much risk can we take? The misuse of risk assessment in corrections. *Federal Probation*, *70*(2), 58-63.

Beck, U. (1992). *Risk society: Towards a new modernity*. Thousand Oaks, CA: SAGE Publications.

Bonta, J. (2008). Offender risk assessment and sentencing. *Canadian Journal of Criminology and Criminal Justice*, 49(4), 519-529.

Buddress, L.A.N. (1997). Federal probation and pretrial services – a cost-effective and successful community corrections system. *Federal Probation*, *61*(1), 5-12.

Bullock, K. (2011). The construction and management of risk management technologies in contemporary probation practice. *British Journal of Criminology*, *51*, 120-135.

Burgess, E.W. (1928). Factors determining success or failure on parole. In Bruce, A. A., Harno, A. J., Burgess, E. W., & Landesco, J. (eds.). *The workings of the indeterminate-sentence law and parole system in Illinois* (pp. 221-234). Springfield, IL: State Board of Parole.

Clem, C. (2003). *Topics in community corrections: Annual issue 2003: Offender assessment*. Washington, DC: National Institute of Corrections.

Dawes, R. M., Faust, D. & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, New Series*, *243*(4899), 1668-1674.

Eaglin, J. B. & Lombard, P. A. (1981). Statistical risk prediction as an aid to probation caseload classification. *Federal Probation*, *45*, 25-32.

Feeley, M. & Simon, J. (1992). The new penology: Notes on the emerging strategy of corrections and its implications. *Criminology*, *30*, 449-474.

Fitzgibbon, W., Hamilton, C., & Richardson, M. (2010). A risky business: An examination of Irish probation officers' attitudes towards risk assessment. *Probation Journal*, *57*(2), 163-174.

Flores, A.W., Lowenkamp, C.T., Smith, P., & Latessa, E.J. (2006). Validating the Level of Service Inventory—Revised on a sample of federal probationers. *Federal Probation*, *70*(2), 44-48.

Fong, G.T., Lurigo, A.J., & Stalans, L.J. (1990). Improving probation decisions through statistical training. *Criminal Justice and Behavior*, *17*(3), 370-388.

Gottfredson, M.R., & Tonry, M. (1987). *Prediction and classification: Criminal justice decision making*. Chicago: University of Chicago Press.

Gottfredson, S.D. & Moriarty, L.J. (2006). Clinical versus actuarial judgments in criminal justice decisions: Should one replace the other? *Federal Probation*, *70*(2), 15-18.

Grove, W.M., & Meehl, P.E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, *2*, 293-323.

Grove, W.M., Zald, D.H., Lebow, B.S., Snitz, B.E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, *12*(1), 19-30.

Hannah-Moffat, K., Maurutto, P., & Turnbull, S. (2009). Negotiated risk: Actuarial illusions and discretion in probation. *Canadian Journal of Law and Society*, *24*(3), 391-409.

Hanson, R. K. (2009). The psychological assessment of risk for crime and violence. *Canadian Psychology*, *50*(3), 172-182.

Harcourt, B. (2007). *Against prediction: Profiling, policing, and punishing in an actuarial age*. Chicago, IL: University of Chicago Press.

Harris, P. (2006). What community supervision officers need to know about actuarial risk assessment and clinical judgment. *Federal Probation*, *70*(2), 8-14.

Hemple, W.E., Webb, W.H. Jr., & Reynolds, S.W. (1976). Researching prediction scales for probation. *Federal Probation*, *40*(1), 33-37.

Hilton, N.Z., Harris, G.T., & Rice, M.E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *Counseling Psychologist*, *34*(3), 400-409.

Horsefield, A. (2003). Risk assessment: Who needs it? *Probation Journal*, *50*(4), 374-379.

Judicial Conference (2006). *September proceedings of the Judicial Conference of the United States*. Washington, DC: Administrative Office of the United States Courts.

Kemshall, H. (1998). Defensible decisions for risk: Or "it's the doers wot get the blame," *Probation Journal*, *45*(2), 67–72.

Litwack, T.R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law*, *7*, 409-443.

Lombard, P., & Hooper, L. (1998). *RPI facts bulletin*. Washington, DC: Federal Judicial Center.

Lowenkamp, C.T., Latessa, E.J., & Holsinger, A.M. (2006). The risk principle in action: What have we learned from 13,676 offenders and 97 correctional programs? *Crime & Delinquency*, *52*, 77-93.

Lowenkamp, C.T., & Latessa , E.J. (2004). Understanding the risk principle: How and why correctional interventions can harm low-risk offenders, in *Topics in community corrections* (pp. 3-8). Washington, DC: National Institute of Corrections.

Lynch, M. (1998). Waste managers? The new penology, crime fighting, and parole agent identity. *Law and Society Review,* 32(4), 839–869.

Meehl, P. E. (1954). *Clinical versus mechanical prediction*. Minneapolis, MN: University of Minnesota Press.

Monahan, J. (1981). *Predicting violent behavior*. Beverly Hills, CA: SAGE.

Monahan, J., Steadman, H.J., Silver, E., Appelbaum, P.S., Clark Robbins, P., Mulvey, E.P., Roth, L.H., Grisso, T., & Banks, S. (2001). *Rethinking risk assessment*. New York: Oxford University Press.

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

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

O'Malley, P. (2004). The uncertain promise of risk. *Australian and New Zealand Journal of Criminology*, *37*(3), 323-343.

Petersilia, J. (2004). What works in prisoner reentry? Reviewing and questioning the evidence. *Federal Probation*, *68*(2), 4-8.

Pew Center on the States. (2009). *One in 31: The long reach of American corrections*. Washington, DC: The Pew Charitable Trusts.

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

Robinson, G. (2002). Exploring risk management in probation practice. *Punishment and Society*, *4*(1), 5-25.

Schneider, A.L., Ervin, L., Snyder-Joy, Z. (1996). Further exploration of the flight from discretion: The role of risk/need instruments in probation supervision decisions. *Journal of Criminal Justice*, *24*(2), 109-121.

Selective incapacitation: Reducing crime through predictions of recidivism (note). (1982). *Harvard Law Review*, 96, 511-533.

Sjöstedt, G., & Grann, M. (2002). Risk assessment: What is being predicted by actuarial prediction instruments? *International Journal of Forensic Mental Health*, *1*(2), 179-183.

Sreenivasan,S., Kirkish, P., Garrick, T., Weinberger, L E., & Phenix, A. (2000). Actuarial risk assessment models: A review of critical issues related to violence and sex-offender recidivism assessments. *Journal of the American Academy of Psychiatry and the Law*, *28*, 438-448.

Tastle, W J., Wierman, M.J., & Dumdum, U.R. (2005). Ranking ordinal scales using the consensus measure. *Issues in Information Systems*, *6*(2), 96-102.

Taxman, F.S., Shepardson, E.S., Byrne, J.M., Gelb, A., & Gornik, M. (2004). *Tools of the trade: A guide to incorporating science into practice*. Washington, DC: National Institute of Corrections.

VanBenschoten, S. (2008). Risk/needs assessment: Is this the best we can do? *Federal Probation*, *72*(2), 38-42.

Walker, S. (1993). *Taming the system: The control of discretion in criminal justice, 1950-1990*. New York: Oxford University Press.

Wandall, R.H. (2006). Actuarial risk assessment: The loss of recognition of the individual offender. *Law, Probability, and Risk*, *5*(3-4), 175-200.

Yang, M., Wong, S.C.P., & Coid, J. (2010). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, *136*(5), 740-767.

THE DEVELOPMENT AND VALIDATION OF THE FEDERAL POST CONVICTION RISK ASSESSMENT (PCRA)

JAMES L. JOHNSON, PH.D.[1]
CHRISTOPHER T. LOWNEKAMP, PH.D.[2]
SCOTT W. VANBENSCHOTEN, M.S.W[3]
CHARLES R. ROBINSON

ADMINISTRATIVE OFFICE OF THE US COURTS
OFFICE OF PROBATION & PRETRIAL SERVICES
ONE COLUMBUS CIRCLE NORTHEAST
WASHINGTON, D.C.

[1] Corresponding author
[2] For questions related to analysis
[3] For questions related to automation, certification, training and implementation

CONSTRUCTION AND VALIDATION OF THE POST-CONVICTION RISK ASSESSMENT INSTRUMENT

INTRODUCTION

The United States probation system was created in 1925 by the Federal Probation Act. This Act gave the U.S. Courts the power to appoint Federal Probation Officers and the authority to sentence defendants to probation instead of a prison term. One of the primary functions of federal probation is to supervise convicted offenders who are sentenced to a term of probation or a term of supervised release following a period of imprisonment, and offenders released early from prison on parole or mandatory release by the U.S. Parole Commission or military authorities.

The federal probation and pretrial services system is organized into 94 districts within 11 regional circuits and operates under a decentralized management structure. As a result of being decentralized, each district operates with a great deal of autonomy and control over their respective district. Despite this autonomy, the system maintains cohesion through the Administrative Office of the U. S. Courts (AO). The AO serves as the administrative headquarters for this decentralized system and develops national policies that help districts in their efforts to protect the community and reduce recidivism.

During the past two decades, advancements in social science research, the need to use resources more efficiently and effectively, and increased expectations to reduce recidivism have sparked a major philosophical shift in the field of probation. Although probation officers are still required to monitor offender behavior and report non-compliance to the court, the general focus has shifted to reducing future criminal behavior (Alexander & VanBenschoten, 2008). Arguably, the best chances for reducing recidivism occurs not only when officers have a reliable way of distinguishing high risk offenders from low risk offenders but also when officers are able to intervene in the criminogenic (crime supporting) needs of high risk offenders (Andrews et al., 1990; Lowenkamp & Latessa, 2004; Bonta & Andrews, 2007; Campbell, French & Gendreau, 2007). For federal probation, this has meant looking for more effective ways to manage offenders by predicting their potential to reoffend and/or their potential dangerousness to the community (Walklate, 1999).

The purpose of this article is to share the process the AO used to develop a risk assessment instrument for use with its post-conviction supervision population. In this article, we provide a brief overview of the principles of effective classification, a summary of the evolution of risk assessments, and explain why the AO chose to create its own risk assessment instrument rather than use an existing instrument. However, the primary purpose of the article is twofold: (1) to present the methodology and results produced in the development of the Post-Conviction Risk Assessment (PCRA) tool, and (2) to discuss limitations of the PCRA as well as future developments.

Principles of Effective Risk Classification

In general terms, the principles of effective risk classification refer to the prediction or identification of offenders most likely to violate the law or conditions of supervision during a period of criminal justice supervision, the identification of factors that can be influenced to change the likelihood of recidivism, and the acknowledgement of factors that might influence the benefits of a particular service (Van Voorhis & Brown, 1996). Risk of recidivism, criminogenic need, and general responsivity are three of the primary principles of effective classification (Andrews et al. 1990). The fourth principle, professional discretion, targets the professional's ability to look beyond the application of the first three principles when circumstances indicate a need to do so (Gottfredson, 1987).

The principles of effective risk classification suggest that agencies should use actuarial assessment tools to identify dynamic risk factors, especially in high risk offenders, while also identifying potential barriers to treatment (Bonta & Andrews, 2007; Latessa et al., 2010). Actuarial risk assessments rest on three factors: (1) certain individual characteristics and behaviors are statistically predictive of future involvement in criminal behavior; (2) the more risk factors an offender has the greater the likelihood of future criminal behavior; and (3) when properly validated and administered, actuarial risk predictions are more accurate than clinical predictions (Meehl, 1954; Sawyer, 1966; Gottfredson, 1987; Andrews and Bonta, 1994). Andrews and Bonta (1998) argue that it is the combined assessment of risk and need that improves the ability to predict who is likely to offend and outlines what interventions should take place to reduce risk and subsequently recidivism.

BRIEF HISTORY OF RISK ASSESSMENT TOOLS

Purpose of a Risk Assessment Tool

The assessment of offenders has long been acknowledged as a necessary component for criminal justice practitioners who are responsible for assessing and managing offenders. In the field of probation, the primary purpose for using a risk assessment tool is to help keep communities safe from offenders who are most likely to reoffend. Although security was the primary reason for the development of risk assessment instruments, the ability to classify offenders at the appropriate risk level is also beneficial. Consequently, risk assessment tools help probation officers identify which offenders need intensive interventions and what needs should be targeted by the interventions.

Evolution of Risk Assessment Instruments

The evolution of risk assessment is described as following a generational path that started with the most basic form of assessment and has progressed to a more complex form of risk assessment (Bonta & Wormith, 2008). Each generation utilized the best available methods to predict the risk of recidivism and then applied the results of the assessment to supervision strategies. This tradition continues today, with researchers continually refining their understanding of criminal behavior and the associated enhancements to risk/needs prediction tools (VanBenschoten, 2008).

First generation

For most of the 20th century, professional judgment or intuition was the most common method used to predict criminal behavior. This form of assessment involved an unstructured interview with the offender and a review of official documentation (Bonta, 1996; Van Voorhis & Brown, 1996; Andrews & Bonta, 2006; Connolly, 2003). Guided by their own professional training and experience, probation officers and clinical professionals would make judgments as to who required enhanced supervision or correctional programming (Bonta & Andrews, 2007). One of the inherent weaknesses of such an unstructured process is there is no quantitative way to determine how decisions are reached which causes a lack of consistency and agreement resulting in low inter-rater reliability (O'Rouke, 2008). In other words, the same interview conducted by different interviewers could net dramatically different results; therefore, the conclusions and recommendations regarding the offender could vary depending on the interviewer (Wardlaw & Millier, 1978; Monahan, 1981; Van Voorhis & Brown, 1996).

Second generation

Although second generation risk tools have been available since the late 1920s, it was not until the 1970s that the assessment of risk began to depend more upon actuarial, evidence-based science and less on professional judgment and intuition. Second generation risk assessments are often referred to as actuarial methods (O'Rouke, 2008). Actuarial risk assessments consider individual items (e.g., history of substance abuse) that have been demonstrated to increase the risk of reoffending and assign these items quantitative scores (Bonta & Andrews, 2007). Burgess (1928) established the first of these models. In the Burgess method, each variable in the model can be scored as a "point," and the prediction is based on the aggregate number of points assigned to an offender (Connolly, 2003). For example, the presence of a risk factor may receive a score of one and its absence a score of zero. The scores on the items can then be summed – the higher the score, the higher the risk that the offender will reoffend (Bonta & Andrews, 2007) This technique gives equal weight to all predictors, even though there may be unequal effects. There is little research, if any, indicating that more complex (i.e., weighted) scoring methods produce better prediction than simple (i.e., unweighted) methods (Gottfredson 1987).

Third generation

Recognizing the limitations of second generation risk assessment, research began to develop in the late 1970s and early 1980s on assessment instruments that included dynamic risk factors (Bonta & Wormith, 2007). The third generation of assessment is commonly referred to as risk-need assessments (Andrews & Bonta, 1995; Bonta & Andrews, 2007). These instruments combined the static predictor variables of the second-generation instruments

with dynamic criminogenic need items (e.g., present employment, criminal friends, and family relationships) that were sensitive to changes in an offender's circumstances (Connolly, 2003; Bonta & Andrews, 2007). Third generation risk assessment tools exceed statistical risk prediction by adding the element of need identification. As previous instruments assisted in decision-making regarding supervision conditions, third generation assessments help identify areas that require intervention to mitigate recidivism risk while under supervision (Van Voorhis & Brown, 1996).

Fourth generation

The last few years has seen the introduction of fourth generation risk assessment instruments. These new risk assessment instruments go beyond the third generation risk-need assessments. Not only do fourth generation instruments include risk-need assessments, they also assess a broader range of risk factors along with responsivity factors important to treatment for integration into the case management plan (Bonta & Andrews, 2007; Bonta & Wormith, 2007). Some examples of responsivity factors include reading and cognitive abilities, race, gender motivation to change as well as external factors such as treatment setting and counselor characteristics (Andrews et al., 1990; Bonta & Wormith, 2007). One other aspect of fourth generation risk assessments is the attempt to explicitly link identified needs with supervision and treatment services (Bonta & Wormith, 2007).

POST CONVICTION RISK ASSESSMENT TOOL

Actuarial risk assessments are not new to the federal probation system; in fact, they have been part of the supervision process since the early 1980s. To better assist probation officers in identifying high risk offenders and intervening in their criminogenic needs, the AO chose to develop a risk assessment instrument tailored specifically to its population of offenders. The Post-Conviction Risk Assessment (PCRA) is an actuarial risk and needs assessment tool developed from data collected on federal offenders who started a term of supervision between October 1, 2005 and August 13, 2009. This tool is designed to target treatment interventions prioritized by risk, need, and responsivity.

How the PCRA Came into Existence

In the *Strategic Assessment of the Federal Probation and Pretrial Services System* (hereafter referred to as the Strategic Assessment), the authors identified shortcomings with the AO's use of the Risk Prediction Index (RPI).[4] One of the concerns expressed by the authors was the RPI's static nature which causes a disconnection between the risk score and case management (IBM, 2004). Put another way, if an offender's risk to recidivate changes during the course of supervision, the RPI would not reflect this change; therefore, officers would not be able to consistently and effectively interpret those changes and provide the proper supervision response.

As a means to address the RPI's shortcomings, the Strategic Assessment recommended the AO research other data-driven supervision tools (IBM, 2004). The desire to meet the Strategic Assessment recommendation coupled with emerging criminal justice literature about more advanced risk assessment tools influenced the AO to develop its own Research to Results (R2R) effort. During the R2R effort, 16 of the 94 federal probation districts were awarded funding to implement evidence based practices[5] into their district. Of those 16 districts, five districts chose to use a commercially available risk and needs tool to conduct risk assessments. In addition, AO staff members met with developers of three commonly used off-the-shelf risk/needs tools (LS/CMI, COMPAS, RMS)[6] to better understand the advantages and disadvantage of each tool.

[4] The RPI uses 8 largely static questions to determine the risk that an offender will recidivate during his or her term of supervision and the results are intended to assist officers in creating the offender's initial supervision case plan.

[5] Districts were required to submit a proposal, which included a budget, outlining an area of evidence based practices (EBP) they wanted to implement. The areas of EBP available were risk assessment, cognitive behavioral interventions, motivational interviewing, and other. The "other" category was open and districts that chose this option tended to use it for drug courts and workforce development.

[6] LSI (Level of Service Inventory), COMPAS (Correctional Offender Management Profiling for Alterative Sanctions), RMS (Risk Management Services)

Since the federal criminal justice system represents a distinctive population and specific trailer assessments for special needs populations (e.g., sex offenders) are also required, it became obvious that more flexibility would be needed. At the conclusion of the experimentation and information gathering stage, the AO assembled a panel of experts to examine the options of purchasing a commercially available tool or building a new tool. After much discussion, the consensus of the group was to build a new tool with data specific to federal probation.

CONSTRUCTION AND VALIDATION OF THE PCRA

METHODS

Data used to construct and validate the PCRA came from federal presentence reports (PSR), existing risk assessments, criminal history records, and PACTS.[7] Criminal history records or rap sheets were used to identify any new arrest after the start of supervision. The five R2R districts that were using a commercially available risk assessment tool were asked by the AO to provide data to assist in the development of the PCRA.[8] Each district provided a list of offenders who had received an assessment using an off-the-shelf risk prediction instrument and who also had a completed PSR. In total, the five districts submitted a list of 4,746 offenders of which 479 cases were randomly selected.[9] Districts were then asked to provide rap sheets on the randomly selected cases. PACTS was the main source of data for scored elements on the PCRA and it included data on roughly 100,000 offenders.

Data Elements

There are two sets of items included on the PCRA: scored and not scored. The first set of items are rated and scored and thus contribute to an offender's risk score. Rated and scored items used to develop the PCRA were based on prior research in the area of predicting criminal behavior (for example Gendreau, Little, & Goggin, 1996; Simourd & Andrews, 1994; Hubbard & Pratt, 2002; Andrews & Bonta, 2006) that were also available in PACTS. Based on a review of extant research, data elements related to criminal history, peer associations, family, employment, substance abuse, and attitudes were selected from PACTS. As a result of bivariate analyses, some interval and ratio variables (e.g., age, prior arrests, education, and drug and alcohol problems) were collapsed into ordinal measures. Multivariate models and completeness of data were used to identify the most predictive and practical data elements to be included on the instrument. Variables included on the PCRA had a significance level of .10 or below (see Table 1).

The second set of data elements are rated but not scored and do not contribute to an offender's risk score. These items were identified as potentially predictive in a smaller sample of offenders from five of the R2R districts. With the exception of peer relationships, which came from the COMPAS and RMS, data elements came from the PSR. A total of 104 elements were collected from the PSR, however, four of those elements were personal identifiers (i.e., first name, last name, middle initial, and PACTS ID number). Additional rated but not scored items were added based on probation officers' input on what data they need to supervise a case. A total of 29 factors were identified as potential predictors and included on the assessment (see Appendix 1). These potential predictors were included as "test items" and future analysis will determine whether these items will become rated and scored PCRA items.[10]

Sample

[7] PACTS (**P**robation/Pretrial Services **A**utomated **C**ase **T**racking **S**ystem) is an electronic case management tool used by probation and pretrial services officers in all 94 federal districts to track federal defendants and offenders. At the end of each month, districts submit case data into a national repository that is accessible to the Administrative Office of the U.S. Courts (AO), Office of Probation and Pretrial Services.

[8] One district was not an R2R district but had been using a commercially available risk assessment tool (RMS) longer than the other four R2R districts.

[9] Districts were initially informed that 100 cases from each district would be randomly selected, but one district only permitted 10 percent of their cases to be selected, which limited their sample to 64 cases.

[10] Due to ongoing data collection, the test items have yet to be analyzed. Decisions to include or omit test items will be determined by statistical significance and by how a test item impacts the predictive accuracy of the PCRA.

In order to construct and validate the PCRA, the researchers devised three sample groups. A construction group was created for the construction of the instrument, and two validation groups were created for the validation of the instrument.[11] These groups were created using an existing analysis file from PACTS data that contained 185,297 offenders on probation or supervised release.[12] The construction group was created from data obtained from the initial case plan.[13] Using a near 50/50 split, data from the first case plan was divided into two sample groups, one became the construction sample and the other became the first validation group. One validation group (Validation) was taken from the initial case plan the offender receives during his/her term of supervision and the second validation group was taken from subsequent case plans, hence the name Subsequent Case Plan. Both the construction (N=51,428) and validation (N=51,643) groups comprised offenders who started a term of supervised release or probation on or after October 01, 2005. The subsequent case plan group was comprised of 193,586 case plan periods.

Analysis

A fairly straightforward and traditional approach was used in the development of the PCRA. Multivariate logistic regression models[14] were used to determine which items were superfluous. As a result, the total number of items included in the multivariate model was reduced to ensure that statistical significance and direction of the relationship were maintained. Once the multivariate model was finalized, bivariate cross tabulations were used to assign appropriate weights. This method was chosen due to its transparency and, to date, there is little research indicating the superiority of complex weighting structures over dichotomous coding risk factors (see Gottfredson & Gottfredson, 1980; Silver, Smith, & Banks, 2000; Gottfredson & Synder, 2005; Harcourt, 2007).[15] The bivariate cross tabulations are presented in Appendices 2-4.

Once the final scoring algorithm was determined, a score was calculated with a cutoff score developed by visual inspection of the data. Although the data cutoffs were fairly evident in the data, alternate cutoffs were tested with confirmation of best fit as determined through the use of chi-square statistics. A final set of analyses were conducted to determine how changes or stability in risk category from the beginning to the end of supervision was correlated with change in the probability of a new arrest.

FINDINGS

Table 1 displays the results of a multivariate model predicting arrest during the initial case plan period using a split sample from the construction sample. As Table 1 shows, many of the variables included in the multivariate model were statistically significant at the .001 level. Odds ratios in the model also appear to be consistent with existing research that support well accepted beliefs that alcohol and drug problems, unemployment, poor attitude (not

[11] Two validation samples were developed in order to test the robustness of the instrument.

[12] Data from the analysis file was assembled from PACTS and matched with data from the Federal Bureau of the Prisons (BOP), the U.S. Sentencing Commission (USSC), and the Census Bureau. Arrest data came from ATLAS (Access to Law Enforcement System) and from the FBI's Computerized Criminal History (CCH) database. Arrest data are current through August13, 2009. Offenders in the analysis file began active post-conviction supervision between October 1, 2004 and August 13, 2009 (see Baber, 2010). Of the 185,297 offenders in the analysis file, only 103,071 had criminal histories and other relevant items used to construct the PCRA.

[13] As outlined in *Monograph 109*, case plans are to be submitted within 30-60 days of the start of the offender's supervision term. This plan is formally evaluated and modified during the sixth month of supervision and updated annually for the duration of the supervision term.

[14] When the outcome variable is composed of only two values (e.g., arrest or no arrest), which is typical for risk classification in probation, logistic regression is usually the best approach to use. The main advantage of logistic regression is that few statistical assumptions are required for its use. In addition, it generates probability values that are constrained between zero and one. Logistic regression calculates the probability of an event occurring or not occurring (e.g., getting arrested or not getting arrested) and presents the results in the form of an odds ratio (Exp(B)). For the purposes of this article, the odds ratio is the number by which you multiply the odds of getting re-arrested for each one-unit increase in the independent variable (i.e., a variable in the equation). An odds ratio greater than 1 indicates that the odds of getting re-arrested increase when the independent variable increases; an odds ratio less than 1 indicates that the odds of getting re-arrested decrease when the independent variable increases (Menard, 2002).

[15] While the iterative classification processes seem to rate higher on some measures of utility, they also tend to have higher degrees of predictive shrinkage (see Banks et al., 2000).

motivated to change), criminal history, and lack of social support increase an offender's chances of getting re-arrested. Females appear to have a decreasing effect on the likelihood of re-arrest which is also consistent with much of the existing research on gender and crime (Gendreau et al., 1996).

From the multivariate analysis, variables were selected for inclusion on the risk assessment instrument (see Appendix 5). To gain a better understanding of the bivariate relationships between the significant predictors in the multivariate model, a series of cross tabulations were conducted. Those results are reported in Appendices 2-4. In general, the bivariate cross-tabulations allowed us to assign 1 or 2 points to each of the factors. Although this approach may seem counter to prevailing wisdom on the development of weights for risk assessment, there is evidence that suggests that this approach produces an instrument that still outperforms clinical approaches to prediction (Dawes, 1979) and is more robust across time and sample variations (Gottfredson & Snyder, 2005; McEwan, Mullen, & McKenzie, 2009).

Table 1. Multivariate Model Predicting Arrest During Initial Case Plan Period (Split Sample Construction Only)

Variable	B	SE	Wald	df	Sig	Exp(B)
Community Supervision Violation	.343	.052	43.551	1	.000	1.410
Varied Offending Pattern	.226	.049	21.416	1	.000	1.253
Institutional Adjustment	.227	.103	4.848	1	.028	1.255
Violent Offending	.320	.079	16.312	1	.000	1.378
Unemployed	.368	.045	66.248	1	.000	1.445
Poor Work Outlook	.322	.061	27.495	1	.000	1.380
Alcohol Problems	.479	.102	22.079	1	.000	1.615
Lacks Social Support	.267	.048	30.673	1	.000	1.306
Family Problems	.191	.051	14.278	1	.000	1.210
Single	.097	.054	3.175	1	.075	1.102
Not Motivated to Change	.383	.050	59.803	1	.000	1.467
Drug Problems	.710	.062	132.195	1	.000	2.033
Arrest History	.149	.021	50.543	1	.000	1.160
Age	.383	.033	136.614	1	.000	1.467
Educational Attainment	.234	.045	27.195	1	.000	1.264
Mental Health Problems	.068	.049	1.920	1	.166	1.070
Gambling Addiction	-.395	.283	1.945	1	.163	.674
Criminal Associates	-.080	.050	2.529	1	.112	.923
Weapon Concerns	-.086	.064	1.789	1	.181	.917
Financial Problems	-.070	.078	.806	1	.369	.932
Life Skills Deficiencies	-.019	.060	.103	1	.748	.981
Female	-.215	.058	13.586	1	.000	.807
Race			3.106	4	.540	
Asian	.613	.490	1.568	1	.211	1.846
Black	.638	.467	1.866	1	.172	1.892
Native American/Eskimo	.668	.475	1.977	1	.160	1.951
White	.683	.466	2.145	1	.143	1.980
Constant	-4.540	.472	92.691	1	.000	.011

Model $\chi^2(26) = 1503.78$, $p < .000$; -2LL = 15868.80; Nagelkerke $R^2 = .119$

Table 2 presents the descriptive statistics on the risk assessment score which can theoretically range from 0 to 19. There are 15 scored items. The scoring for each of the 15 items is displayed in detail in Appendix 5. Table 2 presents the number of cases in each sample, minimum and maximum values, mean, and standard deviation of the linear risk score. There are no significant differences in the length of the prediction period or average risk score for the construction sample and first validation sample (6.46 and 6.43, respectively). However, there are differences in the mean risk score between the subsequent case plan sample and construction sample and subsequent case plan sample and first validation sample. The difference in prediction periods are a matter of policy as the first case plan period is approximately 6 months while the third case plan is completed 12 months after the second case plan or 18 months after the beginning of supervision. The lower mean risk score might simply be a function of lower-risk offenders surviving supervision to the third and subsequent case plan periods. At any rate, there could be some debate that the difference in risk scores is not practically significant and this argument might be valid since all three mean scores fall into the low risk category.

Table 2. Descriptive Statistics

Sample Group	N	Minimum	Maximum	Mean	Std. Deviation
Construction	51,428	0	16	6.4634	2.83052
Validation	51,643	0	16	6.4272	2.80699
Subsequent Case Plan	193,586	0	17	6.0320	2.73192

Table 3 presents the distribution of risk categories by the type of sample used. In all three samples, low and low-moderate risk offenders accounted for at least 85 percent of the cases; whereas high-risk offenders accounted for only 1 percent. There was no statistically significant difference between the construction sample and the validation sample at an alpha level of .01. However, there was a significant difference between the second validation sample (subsequent case plan) and the construction sample as well as between the second validation and the first validation sample. This is likely due to higher-risk offenders having a greater likelihood of revocation and thereby failing to survive to the second and subsequent case plan periods. This finding, as was the case with the linear risk score, might be more an issue of sample size rather than practical significance. The change in the percentage of low-risk cases is what seems to be driving the overall significant chi-square test.

Table 3. Distribution Across Risk Categories

	Sample					
	Construction		Validation		Subsequent Case Plan	
Risk Category	N	%	N	%	N	%
Low	19,080	37%	19,175	37%	83,037	43%
Low-Moderate	24,751	48%	25,175	49%	90,003	47%
Moderate	7,019	14%	6,748	13%	19,244	10%
High	578	1%	545	1%	1,302	1%

The next set of analyses focused on assessing the PCRA's predictive ability. AUC-ROC (Area of the Curve-Receiver Operating Characteristics)[16] was chosen as the measure to assess prediction in large part because it is not impacted by base rates. Another convenient property of the AUC-ROC, over a correlation coefficient is that AUC-ROC is a singular measure and does not have differing calculations depending on level of measurement of the variables being evaluated (Harris & Rice, 2005). Table 4 displays the AUC-ROC between risk scores and re-arrests.

[16] The AUC measures the probability that a score drawn at random from one sample or population (e.g., offenders with a re-arrest) is higher than that drawn at random from a second sample or population (e.g., offenders with no re-arrest). The AUC can range from .0 to 1.0 with .5 representing the value associated with chance prediction. Values equal to or greater than .75 are considered good.

A fourth sample (long-term follow-up) that includes initial case plan data on all offenders placed on supervision between September 30, 2005 and September 30, 2006 is introduced in Table 4. The data therefore allow for a follow-up period between 3 and 4 years. As Table 4 shows, the AUC for each of the four sample groups is close to or exceeds the AUC-ROC value associated with large effect sizes (Harris & Rice, 2005). The AUC for the second validation sample rose to .73 while the AUC for the long term follow-up sample rose even higher to .78. Based on these results, the PCRA appears to have very good predictive validity in terms of accurately classifying offenders' risk levels.

Table 4. AUC-ROC Between Risk Score and Re-arrest[17]

Sample	AUC	Lower 95% CI	Upper 95% CI	Significance
Construction	.709	.699	.719	.000
Validation	.712	.702	.721	.000
Subsequent Case Plan	.734	.729	.739	.000
Long-term Follow-up	.783	.778	.789	.000

To put the AUC values into practical terms,[18] we calculated the failure[19] rates by each category of risk for each sample. These results are presented in Table 5 below. With the exception of the long-term follow-up sample, the failure rates were relatively unchanged for a risk category across samples. For example, low-moderate risk offenders failed at a rate of 13 percent in both the construction and initial validation samples, and at 12 percent in the subsequent case plan sample. However, in the long-term follow-up sample, the low-moderate risk group's failure rate increased significantly to 42 percent. Overall, the failure rate for the long-term follow-up group was 44 percent, but the failure rate was significantly higher for high risk offenders in this same group. Moderate risk offenders failed at a rate of 71 percent and high risk offenders had an 83 percent failure rate. The uniform increase in failure rates across categories of risk and across the various samples continue to provide support for the validity of the PCRA.

Table 5. Cross-tabulation between Risk Categories and Re-arrest

	Sample			
Risk Category	**Construction**	**Validation**	**Subsequent Case Plan**	**Long-term Follow-up***
Low	5%	5%	4%	11%
Low-Moderate	13%	13%	12%	42%
Moderate	27%	28%	27%	71%
High	39%	42%	41%	83%
$\chi 2$	1354.76	1444.74	6761.77	4997.40

*Outcome measure is arrest for new criminal behavior only.

[17] Analyses based on TSR versus probation supervision were estimated. AUC-ROC values for the probation sub-samples were .65 (construction), .64 (validation), .72 (subsequent case plan), and .76 (long-term follow-up). While AUC-ROC values for the construction and validation samples were somewhat smaller than those generated for the overall sample, the AUC-ROC values for the subsequent case plan and long-term follow-up probation sub-samples were very similar to those generated for the overall sample.

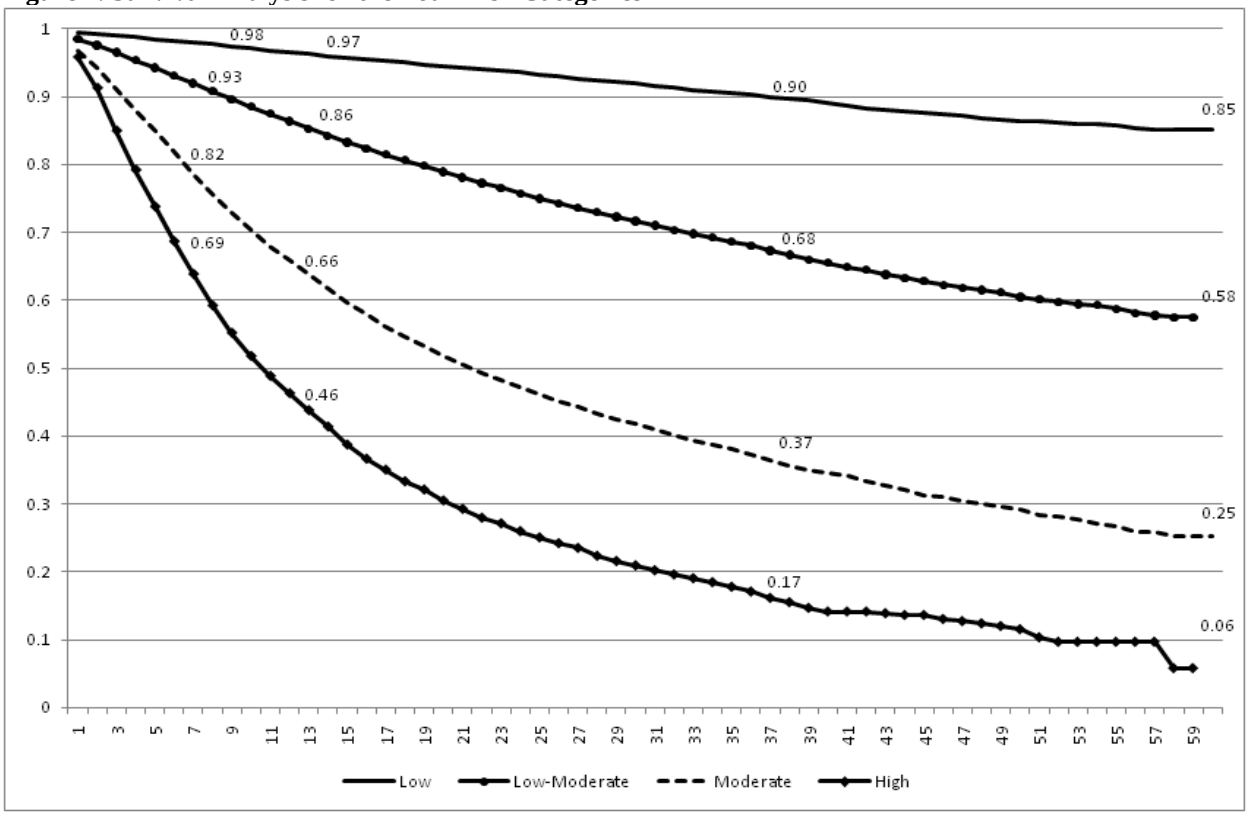
[18] Harris and Rice indicate that the AUC holds the same meaning as the common language effect size indicator. That is, the probability that the PCRA score for a randomly selected recidivist is higher than the PCRA score for a randomly selected non-recidivist. For example, using the long term follow-up data (AUC = .78), if you randomly select a recidivist and a non-recidivist, the recidivist's PCRA score should be higher than the non-recidivist's score 78 percent of the time.

[19] Failure is defined as any new arrest during a term of supervision.

Survival analysis was conducted for each risk category and the survival curves associated with those analyses are displayed in Figure 1. All possible data points, regardless of follow-up time, were used in the analysis.[20] The follow-up period ranged from 0 to 60 months. Survival rates for each risk category are displayed at 6 months, 12 months, 36 months, and 60 months. As Figure 1 shows, high-risk offenders have a very steep decrease in survival as only 69 percent survived the first 6 months of supervision. As time passes, survival rates continue to drop rapidly for high-risk offenders as only 46 percent survived at 12 months and only 17 percent at 36 months. After 60 months of supervision, a mere 6 percent of the high-risk offenders remain. In contrast to high-risk offenders, low-risk offenders have a significantly different experience on supervision. For example, while the survival rate for high-risk offenders was only 17 percent at 36 months, 90 percent of the low-risk offenders survived at this time period. Moreover, the survival rate for low-risk offenders decreased only 5 percentage points through 60 months to 85 percent.

Low-moderate risk offenders have a survival curve that is almost literally between the survival curves of the low- and moderate-risk cases. Interestingly, the survival curve for the moderate-risk offenders seems to follow a form that is closer to the high-risk offenders than to the lower-risk offenders. Note that the survival rates continue to grow throughout the follow-up period for each group and each curve, with the exception of low-risk offenders, shows little sign of leveling off.

Figure 1. Survival Analysis for the Four Risk Categories



One of the major benefits of third and fourth generation risk assessment is their ability to measure change in risk over time. While many of the risk factors on the PCRA would be considered stable, some would also be considered acute (for a full discussion see Serin, 2009). As such, analyses were conducted that compared actual failure rates based on changes in initial and subsequent PCRA assessments. Table 6 outlines changes in failure rates based on first and last case plan assessment categories. The failure rates are based on the risk category for the last case plan

[20] STATA adjusts for cases that were lost during follow-up when calculating survival tables.

period of the offender's supervision term; therefore, to be included in this table the offender had to have at least two case plan periods that allowed for the scoring of the PCRA. According to the results presented in Table 6, not surprisingly, offenders in the higher risk categories (moderate and high) failed at a higher rate than offenders in the lower risk categories (low and low-moderate). However, offenders whose risk rating increased while under supervision appear to fail at a higher rate than offenders who maintained their initial rating through to their last assessment. For example, low-moderate risk offenders whose risk category increased to moderate had a failure rate of 41 percent, whereas low-moderate risk offenders who remained low-moderate risk or were reassessed as low risk had a failure rate between 16 and 18 percent. Similarly, moderate risk offenders who continued to be moderate risk had a 38 percent failure rate while those who were reassessed as low-moderate had an 18 percent failure rate and moderate risk offenders reassessed as high risk had a 61 percent failure rate.

Table 6. Changes in Failure Rates Based on First and Last Case Plan Assessment Categories

	Last Case Plan Assessment Category			
Initial Case Plan Assessment Category	**Low**	**Low-Moderate**	**Moderate**	**High**
Low (n = 13,589)	4%	18%	--	--
Low-Moderate (n = 15,660)	5%	16%	41%	--
Moderate (n = 3,581)	--	18%	38%	61%
High (n = 233)	--	--	37%	53%
$\chi 2$	237.65	396.23	162.85	10.54

DISCUSSION

As previously stated, the purpose of this article is twofold: (1) To present the methodology and results produced in the development of the PCRA, and (2) to discuss limitations of the PCRA as well as future developments. This article has provided details on the methods, measures, and sample used in the development of the PCRA. A fairly traditional model was followed in the development of the PCRA. Our efforts were supported by a relatively large dataset and fairly complete data. The sample was fairly representative of the population served and allowed for a construction and two validation samples. The overall results have demonstrated that the PCRA provides adequate predictive validity both in the short term (6-12 months) as well as in longer follow-up periods (up to 48 months).

Multivariate analysis (see Table 1) of proposed predictors revealed that 15 factors were significantly related to the outcome of interest (new arrest). Seven additional factors tested were determined to be unrelated to a prediction of new arrest once the effects of the other factors were controlled. One additional measure, being female, was found to be significantly related to a new arrest. Subsequent models, not reported here, indicated that the addition of gender to the models yielded no increase in the predictive validity of the model. In addition, non-significant differences were noted in the AUCs between males and females for each sample (i.e., construction, validation, subsequent case plan, and long-term follow-up). As such, we concluded that the instrument performs equally well for males and females even though the failure rates for males might be slightly higher than for females with similar risk scores.

The creation of the risk score and categories allowed for the identification of four risk categories: low, low-moderate, moderate, and high. Approximately 80 percent of each sample was made up of low and low-moderate risk offenders. Much smaller percentages were identified in each sample as moderate and high risk (approximately 12 percent and 1 percent, respectively). Due to the distribution of risk categories being heavily skewed toward lower risk, the validity of the instrument may be brought into question. However, it should be noted that a current validated risk prediction instrument used in the federal system (RPI) yields a similarly skewed distribution. Analysis of failure rates by risk score and category using the PCRA yielded AUC-ROC values over the traditionally accepted value of .70 and an AUC value for the long term follow-up over .78. All of the AUC-ROC values were close to or exceeded the value associated with large effect sizes. Practically speaking, the instrument provided categorizations that are associated with the group failure rates that are differentiated and meaningful for meeting the risk principle (see Tables 4 and 5).

The final analysis conducted in this study related to the dynamic nature of the PCRA. Recall from Table 6 that changes in actual failure rates where associated with changes in risk category from the initial assessment to the last assessment. This finding is rather important as it provides the opportunity to track meaningful changes in risk that occur throughout the supervision process. Moreover, Table 6 confirmed that the PCRA identifies and measures dynamic risk factors that, apparently, when changed through supervision, services, or some other unmeasured process (i.e. natural desistance) lead to commensurate reductions in actual failure rates. The dynamic nature of the PCRA adds to its usefulness in developing case plans throughout the life of the supervision term.

Limitations and Future Research

Although this study was fairly comprehensive in scope and the dataset used was large and representative of the population served, there are a number of limitations and areas for future research that deserve mention. Firstly, while the dataset was large and comprehensive, we have not investigated how scoring algorithms might be adjusted for each district. As with any measure, there is a distribution of AUC values when that test is calculated for each district. Data from 17 districts generated AUC values below .70; however, only three districts had 95% confidence intervals that failed to cross the .70 threshold. While this finding may have been due to small samples in some districts, subsequent analysis should focus on bringing AUC values between risk scores and re-arrests up to larger values.

A second limitation is that the data used in this research came from an administrative dataset. While it proved useful for our initial task of creating and validating a risk assessment instrument, it will be important to conduct similar validation analyses once we have an ample sample of offenders that were actually assessed using the assessment protocol.

The third limitation involves the nature of the outcome measure being predicted. In this research we focused exclusively on the likelihood of a re-arrest and not the severity of the offense. We found it important to assess and determine the likelihoods of a re-arrest as a first step in the assessment process. Because we do recognize that there is more than one dimension to an assessment in the criminal justice system, future analysis will focus on predicting the dangerousness of an offender.

Fourthly, while the PCRA is apparently dynamic with changes in risk associated with changes in actual failure rates, it may not be sensitive enough for use on a monthly or shorter schedule. Due to the high value associated with a dynamic risk assessment, it will be necessary to make the PCRA more sensitive to change, or supplement it with a more sensitive trailer assessment that increases its utility as a guide to service allocation.

Finally, because rated but not scored items outnumber scored items on the assessment, future analysis will include reviewing the impact of rated but not scored items. For example, the PCRA currently only has one scored item in the area of cognitions. As a result of current testing on 80 self-report items that relate to criminal thinking styles, the number of scored items in the area of cognitions will likely increase. Continued analyses on rated but not scored items will also increase the understanding of the impact of self-reported attitudes, as well as guide adjustments to algorithms based on district, gender, and race differences, if relevant.

Policy Implications

Notwithstanding the limitations discussed above, there are two major policy implications that stem from this research. First, the federal probation system now has a dynamic fourth generation risk assessment for use on offenders under its jurisdiction. The instrument can be used to identify higher risk offenders for enhanced services (see Andrews et al., 1990) and can also be used to identify targets for change via external service providers. The second major policy implication is the apparent necessity for ongoing re-assessment. Data analyzed in this study indicate that changes in levels of risk are associated with changes in actual failure rates. With that in mind, it is incumbent upon officers to monitor risk in a standardized way to ensure that supervision and services are having intended impacts. If intended impacts are not being achieved, then officers will be able to modify supervision services to reduce the risk of recidivism.

APPENDIX

Appendix 1. Rated Test Items

Domain	Factor	Rating
Criminal History	Arrested Under Age 18	Yes/No
Employment	Number Of Jobs In Past 12 Months	None/One/More than One
Employment	Employed Less than 50% Of The Last 24 Months	Yes/No
Substance Abuse	Disruption At Work, Home, School	Yes/No
Substance Abuse	Use When Physically Hazardous	Yes/No
Substance Abuse	Legal Problems Related To Use	Yes/No
Substance Abuse	Continued Use Despite Social/Interpersonal Problems	Yes/No
Social Networks	Lives With Spouse And/Or Children	Yes/No
Social Networks	Lack of Family Support	Yes/No
Social Networks	Companions	Good Support And Influence/Occasional Association with Negative Peers/More Than Occasional Association With Negative Peers/No Friends
Attitudes	Antisocial Attitudes	Yes/No
Attitudes	General Criminal Thinking (PICTS)	Scale Scores
Other	No Or Unstable Home	One Address In Past 12 Months/More Than One Address In Past 12 Months or No Permanent Address
Other	Risk Influence At Home	No Criminal Risks Present/Criminal Risks At Home
Other	Financial Stressors	Adequate Income to Manage Debts/No Plan in Place to Meet Financial Debts, Expenses Exceed Income
Other	Pro Social Recreation	Engages In Prosocial Activities/Has No Interests, Does Not Engage In Them, or Recreation Presents Criminal Risk
Responsivity	Low Intelligence	Check Box
Responsivity	Physical Handicap	Check Box
Responsivity	Reading And Writing Limitations	Check Box
Responsivity	Mental Health Issues	Check Box
Responsivity	No Desire to Change/Participate In Programs	Check Box
Responsivity	Homeless	Check Box
Responsivity	Transportation	Check Box
Responsivity	Child Care	Check Box
Responsivity	Language	Check Box
Responsivity	Ethnic Or Cultural	Check Box
Responsivity	History Of Abuse Or Neglect	Check Box
Responsivity	Interpersonal Anxiety	Check Box
Responsivity	Social Security Card, Driver’s License, ID	Check Box

Appendix 2. Cross Tabulations between Risk Factors and Re-arrest for Construction Sample

Domain	Variable	Arrest Rate	$\chi 2$	P
Criminal History	Prior Arrests			
	0 = No prior arrests	9%		
	1 = 1-2 prior arrests	12%	618.33	.000
	2 = 3-6 prior arrests	13%		
	3 = 7 or more prior arrests	20%		
Criminal History	Community Supervision Violations			
	0 = No prior CS violations	11%	423.49	.000
	1 = 1 or more CS violations	20%		
Criminal History	Varied Offending Pattern			
	0 = 1 type of offending	14%	209.81	.000
	1 = 2 or more types of offending	20%		
Criminal History	Institutional Adjustment			
	0 = No adjustment problems	12%	98.57	.000
	1 = Adjustment problems	22%		
Criminal History	Violent Offense			
	0 = No history or current violence	15%	50.405	.000
	1 = History or current violence	19%		
Criminal History	Age			
	0 = 41+	11%	638.77	.000
	1 = 26-40	16%		
	2 = 25 or younger	23%		
Education & Employment	Highest Grade			
	0 = High school degree or more	11%	467.44	.000
	1 = GED or less than HS degree	18%		
Education & Employment	Unemployed			
	0 = Currently employed	11%	318.08	.000
	1 = Currently unemployed	18%		
Education & Employment	Good Work History			
	0 = Stable work history	8%	352.17	.000
	1 = Unstable work history	15%		
Substance Abuse	Alcohol Problems			
	0 = No current problems	12%	264.62	.000
	1 = Current problems	28%		
Substance Abuse	Drug Problems			
	0 = No problems	12%	836.48	.000
	1 = Current problems	29%		
Social Networks	Family Problems			
	0 = No problems	12%	213.77	.000
	1 = Current problems	18%		
Social Networks	Married			
	0 = Married	10%	187.69	.000
	1 = Single	16%		
Social Networks	Social Support		361.23	
	0 = Social support present	9%		.000
	1 = No social support	15%		
Attitudes	Motivated to Change			
	0 = Offender motivated to change	8%	473.99	.000
	1 = Offender resistant to supervision	16%		

Note: Number of cases ranges from 31, 773 to 48,470 depending on risk factor.

Appendix 3. Cross Tabulations between Risk Factors and Re-arrest for Validation Sample

Domain	Variable	Arrest Rate	$\chi 2$	P
Criminal History	Prior Arrests			
	0 = No prior arrests	9%		
	1 = 1-2 prior arrests	11%	612.91	.000
	2 = 3-6 prior arrests	14%		
	3 = 7 or more prior arrests	20%		
Criminal History	Community Supervision Violations			
	0 = No prior CS violations	11%	369.56	.000
	1 = 1 or more CS violations	19%		
Criminal History	Varied Offending Pattern			
	0 = 1 type of offending	14%	196.50	.000
	1 = 2 or more types of offending	20%		
Criminal History	Institutional Adjustment			
	0 = No adjustment problems	12%	87.241	.000
	1 = Adjustment problems	21%		
Criminal History	Violent Offense			
	0 = No history or current violence	15%	59.047	.000
	1 = History or current violence	19%		
Criminal History	Age			
	0 = 41+	11%	499.76	.000
	1 = 26-40	16%		
	2 = 25 or younger	22%		
Education & Employment	Highest Grade			
	0 = High school degree or more	11%	502.72	.000
	1 = GED or less than HS degree	18%		
Education & Employment	Unemployed			
	0 = Currently employed	11%	379.277	.000
	1 = Currently unemployed	18%		
Education & Employment	Good Work History			
	0 = Stable work history	8%	371.27	.000
	1 = Unstable work history	15%		
Substance Abuse	Alcohol Problems			
	0 = No current problems	12%	283.03	.000
	1 = Current problems	29%		
Substance Abuse	Drug Problems			
	0 = No problems	12%	701.78	.000
	1 = Current problems	28%		
Social Networks	Family Problems			
	0 = No problems	12%	197.87	.000
	1 = Current problems	18%		
Social Networks	Married			
	0 = Married	11%	164.99	.000
	1 = Single	16%		
Social Networks	Social Support			
	0 = Social support present	9%	398.44	.000
	1 = No social support	15%		
Attitudes	Motivated to Change			
	0 = Offender motivated to change	8%	507.97	.000
	1 = Offender resistant to supervision	16%		

Note: Number of cases ranges from 31, 607 to 48,434 depending on risk factor.

Appendix 4. Cross Tabulations between Risk Factors and Re-arrest for Subsequent Case Plan Periods

Domain	Variable	Arrest Rate	$\chi 2$	P
Criminal History	Prior Arrests			
	0 = No prior arrests	6%		
	1 = 1-2 prior arrests	8%	3567.58	.000
	2 = 3-6 prior arrests	11%		
	3 = 7 or more prior arrests	17%		
Criminal History	Community Supervision Violations			
	0 = No prior CS violations	10%	2946.37	.000
	1 = 1 or more CS violations	19%		
Criminal History	Varied Offending Pattern			
	0 = 1 type of offending	11%	1679.04	.000
	1 = 2 or more types of offending	18%		
Criminal History	Institutional Adjustment			
	0 = No adjustment problems	11%	631.19	.000
	1 = Adjustment problems	21%		
Criminal History	Violent Offense			
	0 = No history or current violence	11%	304.23	.000
	1 = History or current violence	16%		
Criminal History	Age			
	0 = 41+	8%	3183.72	.000
	1 = 26-40	13%		
	2 = 25 or younger	19%		
Education & Employment	Highest Grade			
	0 = High school degree or more	8%	2509.84	.000
	1 = GED or less than HS degree	15%		
Education & Employment	Unemployed			
	0 = currently employed	9%	1235.60	.000
	1 = currently unemployed	15%		
Education & Employment	Good Work History			
	0 = Stable work history	6%	2083.60	.000
	1 = Unstable work history	12%		
Substance Abuse	Alcohol Problems			
	0 = No current problems	11%	1344.46	.000
	1 = Current problems	24%		
Substance Abuse	Drug Problems			
	0 = No problems	9%	5720.49	.000
	1 = Current problems	27%		
Social Networks	Family Problems			
	0 = No problems	9%	1254.19	.000
	1 = Current problems	15%		
Social Networks	Married			
	0 = Married	8%	1096.37	.000
	1 = Single	13%		
Social Networks	Social Support			
	0 = Social support present	9%	744.26	.000
	1 = No social support	12%		
Attitudes	Motivated to Change			
	0 = Offender motivated to change	7%	2039.84	.000
	1 = Offender resistant to supervision	13%		

Note: Number of cases ranges from 152,241 to 236,866 depending on risk factor.

Appendix 5. Scored PCRA Data Items

VARIABLE NAME	VARIABLE DESCRIPTION	SCORED ITEM
Date of Birth	Record offender's data of birth in MM/DD/YY format.	Captured in 1.7
# Adult Conv	Record the total number of adult convictions.	Captured in 1.2
# Other Arrests	Record the total number of other arrests.	Captured in 1.2
# Violent Arrests	Record the total number of prior arrests for a violent crime.	Captured in 1.3
# DV	Record the number of arrests for domestic violence.	Captured in 1.3
HXSONC	History of sex offending offenses without contact.	Captured in 1.3
HXSOC	History of sex offending with contact. Code Y for yes, N for no, and U for unknown.	Captured in 1.3
HXSOSR	History of sex offending statutory rape. Code Y for yes, N for no, and U for unknown.	Captured in 1.3
HXSOO	History of other sex offending. Code Y for yes, N for no, and U for unknown.	Captured in 1.3
Varied	How many different types of offenses has the offender engaged in (property, drug, sex, violent, order, other)	Captured in 1.4
Inst Adj1	Record the number of times an offender was written up during prior terms of incarceration	Captured in 1.6
Inst Adj2	Record the number of times the offender was officially punished for institutional infractions.	Captured in 1.6
CS Vio	During how many previous periods of supervision did the offender a) commit a new crime or b) have violations that were reported to the court or paroling authority.	Captured in 1.5
High Grade	Record the highest grade the offender completed. If received a GED code the highest grade completed in school. GED does not equal 12.	Captured in 2.1
Employed PSR	Was the offender employed at the time of the pre-sentence report? Code Y for yes, N for no, and U for unknown.	Captured in 2.2
Employed Arrest	Was the offender employed at the time of the arrest? Code Y for yes, N for no, and U for unknown.	Captured in 2.2
Alc Current	Does the offender have a current alcohol problem? Code Y for yes, N for no, and U for unknown.	Captured in 3.5
Drug Current	Does the offender have a current drug problem? Code Y for yes, N for no, and U for unknown.	Captured in 3.6

REFERENCES

Alexander, Melissa, and Scott VanBenschoten (2008). "The Evolution of Supervision in the Federal Probation System." *Federal Probation*, 72(2), 15-21.

Andrews, Donald A., James Bonta, and Robert D. Hoge (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19-52.

Andrews, Donald A., and James Bonta (1994). The Psychology of Criminal Conduct. Cincinnati, OH: Anderson Press.

Andrews, Donald A., and James Bonta (1995). The Level of Service Inventory - Revised. Toronto: Multi-Health Systems.

Andrews Donald and James Bonta (1998). The Psychology of Criminal Conduct. Cincinnati, OH: Anderson Press.

Andrews, Donald A., and James Bonta (2006). The Psychology of Criminal Conduct, LexisNexis.

Baber, Laura (2010). "Results-based Framework for Post-Conviction Supervision Recidivism Analysis." *Federal Probation*, 74(3), 5-10.

Bonta, James (1996). Risk-Needs Assessment and Treatment. In A. T. Harland (Ed.), Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply (pp. 18-32). Thousand Oaks, CA: Sage.

Bonta, James, and Donald A. Andrews (2007). "Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation," (User Report 2007-06). Ottawa: Public Safety Canada.

Bonta, James, Guy Bourgon, Tanya Rugge, Terri-Lynne Scott, Annie K. Yessine, Leticia Gutierrez, and Jobina Li (2010). "The Strategic Training Initiative in Community Supervision: Risk-Need-Responsivity in the Real World," (User Report 2010-11). Ottawa: Public Safety Canada.

Bonta, James, and Stephen Wormith (2007). Risk and Need Assessment. In Gill McIvor & Peter Raynor (Eds.), Developments in Social Work with Offenders. London, England: Jessica Kingsley Publishers.

Burgess, Ernest W. (1928). Factors Determining Success or Failure on Parole. In A. Bruce, E. Burgess, & A. Harno (Eds.), Prediction and Classification: Criminal Justice Decision Making. Chicago, IL: University of Chicago Press.

Campbell , Mary Ann, Sheila French, and Paul Gendreau (2007). "Assessing the Utility of Risk Assessment Tools and Personality Measures in the Prediction of Violent Recidivism for Adult Offenders." (User Report 2007-04). Ottawa, Ontario: Public Safety Canada.

Connolly, Michele M. (2003). "A Critical Examination of Actuarial Offender-Based Prediction Assessments: Guidance for the Next Generation of Assessments." U.S. Department of Justice, National Institute of Justice (NCJ 202982).

Dawes, Robyn M. (1979). "The Robust Beauty of Improper Linear Models in Decision Making." *American Psychologist*, 34(7), 571–582.

Gendreau, Paul, Tracy Little, and Claire Goggin (1996). "A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works!" *Criminology*, 34(4), 48-69.

Gottfredson, Don M., and Howard N. Snyder (2005). "The Mathematics of Risk Classification: Changing Data Into Valid Instruments for the Juvenile Courts." National Center for Juvenile Justice, Office of Juvenile Justice and Delinquency Prevention (NCJ 209158).

Gottfredson, Stephen. D., and Don M. Gottfredson (1979). "Screening for Risk: A Comparison of Methods." Criminal Justice and Behavior, 7(3), 315-330.

Gottredson, Don M. (1987). Prediction and Classification in Criminal Justice Decision Making. In Don M. Gottredson and Michael Tonry (Eds.), Prediction and Classification: Criminal Justice Decision-Making. Chicago: University of Chicago Press.

Harcourt, Bernard (2007). Against Prediction. Chicago, IL: University of Chicago Press.

Hubbard, Dana J., and Travis C. Pratt. (2002). "A Meta-Analysis of the Predictors of Delinquency Among Girls." *Journal of Offender Rehabilitation*, 34:1-13.

IBM (2004). Strategic Assessment: Federal probation and pretrial services system. Washington, DC.

Latessa, Edward J., Richard Lemke, Matthew Makarios, Paula Smith, and Christopher T. Lowenkamp (2010). "The Creation and Validation of the Ohio Risk Assessment System (ORAS)." *Federal Probation*, 74(1), 16-22.

Lowenkamp, Christopher and Edward Latessa (2004). Increasing the Effectiveness of Correctional Programming through the Risk Principle: Identifying Offenders for Residential Placement. *Criminology and Public Policy*, 4(1), 501-528.

McEwan, Troy E., Paul E. Mullen, and Rachel MacKenzie (2009). "A Study of the Predictors of Persistence in Stalking Situations." *Law & Human Behavior*, 33: 149-158.

Meehl, Paul (1954). Clinical vs. Statistical Prediction. Minneapolis: University of Minnesota Press.

Menard, Scott (2002). Applied Logistic Regression Analysis (2nd Edition). Thousand Oaks: Sage Publications.

Monahan, John (1981). Predicting Violent Behavior: An Assessment of Clinical Techniques. Beverly Hills: Sage.

O'Rourke, Margaret (2008). "Risk Assessment: An Overview." Available online at http://www.stateclaims.ie/ClinicalIndemnityScheme/publications/RiskAssessmentOverview.pdf.

Rice, Marnie E., and Grant T. Harris (2005). "Comparing Effect Sizes in Follow-up Studies: ROC Area, Cohen's d, and r." *Law and Human Behavior*, 29(5): 615-620.

Sawyer, Jack (1966). Measurement and Prediction, Clinical and Statistical. *Psychological Bulletin*, 66(3), 178-200.

Silver, Eric, William R. Smith, & Steven Banks (2000). "Constructing Actuarial Devices for Predicting Recidivism: A Comparison of Methods." *Criminal Justice and Behavior*, 27, 732-763.

Simourd, Linda, and Don A. Andrews (1994). "Correlates of Delinquency: A Look at Gender Differences." *Forum on Corrections Research*, Vol. 6: 26-31

VanBenschoten, Scott (2008). "Risk/Needs Assessment: Is This the Best We Can Do?" *Federal Probation*, 72(2), 38-42.

Van Voorhis, Patricia and Kelly Brown (1996). "Risk Classification in the 1990s." U.S. Department of Justice, National Institute of Corrections.

Walklate, Sandra. (1999) "Is It Possible to Assess Risk?" *Risk Management*, Vol. 1(4), 45-53.

Actuarial and Clinical Assessment of Criminogenic Needs: Identifying Supervision Priorities among Federal Probation Officers

Draft: Submitted to the Journal of Crime and Justice for Review

J.C. Oleson
University of Auckland

Scott VanBenschoten
Charles Robinson
Christopher T. Lowenkamp
Administrative Office of the United States Courts

For many years, "rehabilitation" was a dirty word in corrections, something not to be uttered aloud in polite circles. When Robert Martinson famously concluded that "[w]ith few and isolated exceptions, the rehabilitation efforts that have been reported so far have had no appreciable effect on recidivism" (1974: 25), it sounded a death knell for the rehabilitative movement in the United States (Halleck and Witte, 1977). "Nothing works" reigned for years. Connotative of Lombrosoian crime-as-disease, the rehabilitative ideal was besieged from all sides: the left attacked it for being a patronizing euphemism for vengeful punishments; the law-and-order right attacked it for being soft on crime and for ignoring real harms and real wrongdoing (Allen, 1981). In fact, in the late 1970s, the prognosis for rehabilitation was so bleak that, years later, Francis Cullen (2005) was able to distinguish the twelve individuals who singlehandedly "saved" rehabilitation.

Cullen's pantheon included three Canadian researchers: Paul Gendreau, Don Andrews, and James Bonta. Paul Gendreau was among the first criminologists to challenge Martinson's findings (Petersilia, 2004) and, with his colleagues, later inverted Martinson's skeptical question, "What Works?" into a promising answer: "What Works!" (Gendreau, Little, and Goggin, 1996). Don Andrews and James Bonta, authors of a popular textbook on the psychology of criminal behavior (Andrews and Bonta,

2010), are known for applying the science of behavioral change to the problem of recidivism. To this end, their classification tool, the Level of Service Inventory, Revised (LSI-R) (Andrews and Bonta, 1995), has been widely adopted around the world.

Collectively, Gendreau, Andrews, and Bonta did a great deal to reinvigorate rehabilitation. Perhaps most importantly, they identified three key principles of effective interventions that have radically transformed the practice of community corrections: risk, need, and responsivity (Andrews, Bonta, and Hoge, 1990; Andrews and Dowden, 2007; Gendreau, 1996). The risk principle essentially states that the level of service (treatment and programming) should be matched to the offender's level of recidivism risk: high-risk offenders should get high treatment dosages; low-risk offenders – who actually can become *more likely* to recidivate when subjected to intensive interventions (Andrews and Dowden, 1999; Lowenkamp and Latessa, 2005; Lowenkamp, Latessa, and Holsinger, 2006) – should receive only modest dosages (Lowenkamp *et al.*, 2006). The need principle states that interventions should be focused on the dynamic risk factors ("criminogenic needs") that are significantly correlated with criminal behavior. Some researchers describe "the big four" risk/need factors: antisocial attitudes, antisocial peers, antisocial cognition, and a history of antisocial behavior. Other researchers add four more factors – family and marital problems, school or work difficulties, lack of involvement in non-criminal leisure activities and recreation, and substance abuse – and describe "the central eight" (Andrews and Dowden, 2007). While community corrections officers may also wish to address non-criminogenic needs (e.g., self-esteem, mental health, learning disabilities, or lack of parenting skills), the need principle suggests that resources should be first focused on changing factors that will reduce the likelihood of reoffending (Lowenkamp *et al.*, 2006). The responsivity principle states that styles and modes of interventions should be tailored to offenders' temperaments, motivations, learning styles, genders, and ethnicities. While anxiety, mental health, or IQ score may not be strongly correlated to recidivism (i.e., are not criminogenic needs), these responsivity factors *may* interfere with the ability to

make effective use of treatment interventions. Matching offenders to appropriate services is essential to ensure that offenders can profit from the treatment they receive (Andrews, Bonta, and Hoge, 1990). Responsivity to treatment also depends upon qualities of the staff and the treatment setting (Birgden, 2004).

The principles of risk, need, and responsivity lie at the heart of a body of work known as evidence-based practices (EBP). The notion of evidence-based practices first originated in medicine (e.g., Sackett *et al.*, 2000) but has been adapted to other fields, including community corrections. Promising the ability to effect real reductions in recidivism in a time of economic scarcity, EBP exerts a growing influence that is difficult to overstate. The National Institute of Corrections recently issued a "box set" of EBP documents targeting eight different stakeholders in the criminal justice system: (1) community corrections agencies (Crime and Justice Institute, 2009), (2) practitioners and treatment providers (Scott and Crime and Justice Institute, 2008), (3) pretrial services (VanNostrand and Crime and Justice Institute, 2007), (4) members of state judiciaries (Warren and Crime and Justice Institute, 2008), (5) prosecutors (Fahey and Crime and Justice Institute, 2008), (6) defense counsel (Weibrecht and Criminal Justice Institute, 2008), (7) jails and detention facilities (Christensen and Criminal Justice Institute, 2008), and (8) prisons (Serin and Criminal Justice Institute, 2005).

Evidence-based practice has been especially influential in Commonwealth countries (e.g., Canada, Australia, and the United Kingdom) (Petersilia, 2004). Indeed, some authors suggest that the evidence-based practices movement "has constituted a revolution in the way criminal conduct is managed in Canada, Britain, Europe, Australia, and New Zealand" (Ward, Melser, and Yates, 2007: 3). That said, EBP are central to the guidance issued in the United States by the National Institute of Justice (e.g., Jalbert *et al.*, 2011) and the National Institute of Corrections (e.g., Taxman, Shepardson, and Byrne, 2004). And, increasingly, state correctional systems are relying upon EBP, too (see, e.g., Aos, Miller, and Drake, 2006).

Background and Research Question

Evidence-based practices are reshaping the federal probation and pretrial services system, as well. John Hughes, Assistant Director for the Office of Probation and Pretrial Services at the Administrative Office of the U.S. Courts (AOUSC), traces the first watershed moment to 2000, when two things occurred: (1) almost all chief probation officers from the nation's 94 districts gathered for a futures-planning conference and (2) the Administrative Office contracted with independent consultants to obtain a strategic assessment of the federal probation and pretrial services system (Hughes, 2008). Alexander and VanBenschoten (2008) suggest slightly more recent starting points. First, a second futures-planning conference in 2002 created a "charter for excellence" (reproduced in Hughes, 2008: 6), a document that pledges: "We are outcome driven and strive to make our communities safer and to make a positive difference in the lives of those we serve." Second, the 2004 report on the strategic assessment made the central recommendation that federal probation "become a results-driven organization with a comprehensive outcome measurement system" (Alexander and VanBenschoten, 2008: 321; Hughes, 2008: 6).

The authors are in agreement about many of the changes that have transformed federal probation and pretrial services throughout the last decade. They agree that the drafting of the charter forf excellence and contracting for the strategic assessment were bold, groundbreaking steps. They agree that probation and pretrial services has shifted away from counting outputs and has moved toward measuring outcomes. They agree that the timely and accurate collection of data is an essential aspect of becoming a results-driven organization. And they agree that evidence-based practices are part and parcel of contemporary probation and pretrial services work, informing the federal supervision policy and guiding the efforts of the AOUSC.

Recently, probation staff members at the AOUSC have developed a new, third-generation risk-needs assessment instrument: the Federal Post Conviction Risk Assessment (PCRA). Staff provided numerous PCRA training sessions at four regional meetings convened during 2010 and 2011. Federal probation officers from districts in the greater Washington, DC, metropolitan region gathered in Washington, DC; officers from districts in the eastern United States gathered in Charlotte, NC; officers from districts in the middle of the country gathered in Detroit, MI; and officers from districts in the western United States, including Pacific islands, gathered in Salt Lake City, UT. Approximately 150-350 officers attended each of the training meetings, and data was collected from the sessions.

During these sessions, it became clear that probation officers made more consistent and more accurate assessments of offender risk when they used the PCRA than when they relied upon unstructured clinical judgment (see Oleson *et al.*, 2011). Whether this difference was due to the officers' variance in assigning uniform meaning to terminology related to risk categories or to their fundamentally seeing the case differently, these results are consistent with a robust body of work (e.g., Ægisdóttir *et al.*, 2006; Grove and Meehl, 1996; Grove *et al.*, 2000; Meehl, 1954; Monahan *et al.*, 2001), collected over many decades, indicating that actuarial prediction outperforms unstructured clinical judgment in predictions of risk.

We were interested in whether using the PCRA might also improve the ability of federal probation officers to assess criminogenic needs. On the one hand, research suggests – and our data bore out – that actuarial prediction of risk outperforms clinical judgment; on the other hand, federal probation officers are considered to be the "crème de la crème" of community corrections officers (Buddress, 1997: 6). They are well educated, well trained, and often come to the federal system with substantial practical experience from state, county, and other local systems. Perhaps the dynamic nature of criminogenic needs lends itself to clinical judgment in a way that general risk of recidivism does not.

Methods

Data was collected at the PCRA regional training meetings convened during 2010 and 2011. In advance of the training, each officer was required to participate in 8 hours of online training that reviewed the fundamentals of risk, need, and responsivity (Andrews, Bonta, and Hoge, 1990). The trainers explained to the participating officers that they would be asked to assess an offender's risk of recidivism and criminogenic needs (explained as targets for supervision designed to reduce risk by changing behavior), based on a videotaped mock intake interview and supplementary written documentation. Specifically, the trainers informed the officers that they would be asked to place the offender in the case vignette in one of four risk categories (low, low/moderate, moderate, or high) and to identify the offender's three most important targets to reduce offender risk through supervision and interventions (criminogenic needs), in rank order. Although the probation officers were in a large group setting, officers did not discuss their rankings of risk or identification of criminogenic needs until they had submitted their data collection forms.

The case vignette consisted of a 24-minute mock intake interview (based upon an actual case, with identifiers and key case details modified in order to protect the offender's anonymity). The probation officer in the vignette conducted a typical supervision intake interview, asking the offender – a 47-year-old man with a long history of methamphetamine addiction and firearms charges – a series of questions about the offender's criminal behavior, employment, social networks, cognitions, substance abuse, time in custody, and current accommodations. Supplemental written materials included a mock presentence report and mock release paperwork from the Federal Bureau of Prisons. The offender in the vignette was steadily employed and, although he was divorced, lived in a stable residence with his parents. He participated in treatment, remained free of drug use, and could articulate a relapse

prevention plan. He did not associate with antisocial peers and was in the planning process of developing a prosocial network.

After the video concluded, officers were given as much time as needed to identify the risk level and three top criminogenic needs. Officers typically took between five and ten minutes to review the supplementary material and submit a complete data collection form.

Results

The correct risk score, according to the PCRA, was a 6, representing the bottom end of the "low/moderate" risk scale. PCRA scores ranging between 0 and 5 are considered "low risk," 6-9 are "low/moderate risk," 10-12 are "moderate risk," and 13-18 are "high risk."

The most pressing criminogenic needs in the training case, according to the PCRA, were related to criminal history, substance abuse, antisocial peers, and leisure/recreation. These represent half of the eight major categories of criminogenic needs: (1) antisocial attitudes; (2) antisocial peers; (3) antisocial cognition; (4) a history of antisocial behavior; (5) family and marital problems; (6) school or work difficulties; (7) lack of involvement in non-criminal leisure activities and recreation; and (8) substance abuse.

Top Three Criminogenic Needs as Identified by Probation Officers

Asked to identify the top three criminogenic needs, 1,040 probation officers generated a very substantial list that we have coded into 27 different categories. These are listed in Table 1, below, along with the number of probation officers who listed each given need as first, second, or third.

TABLE 1: Top Three Criminogenic Needs

Identified Need	Listed First	Listed Second	Listed Third
Substance Abuse (Drugs /Alcohol)	634 (61%)	207 (20%)	106 (10%)

Criminal Peers	143 (14%)	165 (16%)	156 (15%)
Mental Health	91 (9%)	254 (24%)	153 (15%)
Employment	41 (4%)	128 (12%)	153 (15%)
Criminal History/Community Supervision Violations	23 (2%)	30 (3%)	35 (3%)
Firearms	22 (2%)	39 (4%)	36 (3%)
Criminal Attitudes	19 (2%)	16 (2%)	17 (2%)
Family and Marital Problems	12 (1%)	45 (4%)	74 (7%)
Leisure/Recreation	8 (1%)	9 (1%)	22 (2%)
Income	6 (1%)	17 (2%)	44 (4%)
None Listed	5 (0%)	26 (3%)	94 (9%)
Housing	5 (0%)	13 (1%)	42 (4%)
Medical	5 (0%)	9 (1%)	22 (2%)
Criminal Recidivism/ Violations This Time	5 (0%)	9 (1%)	8 (1%)
Social Support	5 (0%)	23 (2%)	13 (1%)
Relationships	4 (0%)	15 (1%)	19 (2%)
Isolation	3 (0%)	2 (0%)	3 (0%)
Education	2 (0%)	4 (0%)	6 (1%)
Structure	2 (0%)	5 (0%)	9 (1%)
Monitoring/Surveillance/Drug Testing	1 (0%)	4 (0%)	2 (0%)
Safety	1 (0%)	0 (0%)	0 (0%)
Self-Control	1 (0%)	0 (0%)	2 (0%)
Violence	1 (0%)	3 (0%)	3 (0%)
Stress/Grief/Guilt	1 (0%)	5 (0%)	10 (1%)
	0 (0%)	6 (1%)	4 (0%)

Child Support			
Spiritual Matters/Cultural Matters/Native American Matters	0 (0%)	4 (0%)	5 (0%)
Transportation/Drivers License	0 (0%)	2 (0%)	2 (0%)

Table 1: Top three named criminogenic needs as identified by 1,040 probation officers, by number and percentage.

The response rate was very good. Only 5 of the 1,040 participating probation officers (less than one percent) did not list a first-ranked criminogenic need. Three percent did not list a second-place need and 9% did not list a third-ranked need. Most officers, however, identified three needs deemed relevant to the case. The ten most frequently identified needs will be described briefly below.

Table 1 indicates that 91 percent of the probation officers correctly identified substance abuse as one of the top three relevant criminogenic needs in the case. More than half of the officers (61 percent) identified it as *the* top criminogenic need. (The offender had a history of heavy methamphetamine use, and while he was free of drug use at the time of the interview, his period of sobriety had been short.)

Almost half of the officers (45 percent) also correctly identified criminal peers as one of the top three relevant criminogenic needs in the case. Fourteen percent named it as *the* top criminogenic need. (The offender had divorced because of his drug use, and while he was no longer associating with antisocial peers, he had not yet formed new, prosocial friendships. He described his hopes for creating a prosocial network through involvement with his son's race car team.) The identification of criminal peers as a relevant criminogenic need was impressive. Research suggests that antisocial peers are a particularly important influence in recidivism (Gendreau, Little, and Goggin, 1996).

Many officers identified mental health as one of the top three criminogenic needs in the case. In fact, more officers (48 percent) counted mental health as one of the top three needs than counted criminal peers as such (45 percent). While the offender's interview revealed potential mental health issues, mental health is not strongly statistically correlated with recidivism. It is not counted among top-

tier risk/needs factors; it is not even included among the "central eight" (Andrews and Dowden, 2007). Mental health is not a criminogenic need. Rather, it is a responsivity factor – a barrier – something to be addressed inasmuch as it interferes with the successful delivery of treatment interventions.

Only 8 percent of the officers identified criminal history or community supervision violations as a top target of supervision and only 2 percent identified it as the top need in the case. Criminal history, a static measure, is not a dynamic risk factor like criminal attitudes or criminal peers (Flores et al., 2005); it is, however, strongly predictive of future behavior, and it was the offender's criminal history that raised his risk score out of the low-risk range into the low/moderate risk range. Such a lengthy criminal history may indicate a need to treat criminal history as a criminogenic need and "[b]uild up low-risk non-criminal alternative behaviour in risky situations" (Andrews and Dowden, 2007: 446 tbl. 2).

Approximately 9 percent of the officers identified monitoring potential possession of firearms as a target of supervision; 2 percent identified it as the top target in the case. Firearms are not a criminogenic need.

Of the items identified, the following were considered on the list even though they were not present in the case: employment (31 percent), criminal attitudes (6 percent), family dysfunction (12 percent) and a sustainable income (7 percent). All of these items were stable, yet some officers interpreted them as problematic and in need of intervention.

Only 4 percent of the officers identified recreation as a top-three criminogenic need; 1 percent identified it as *the* top criminogenic need in the case. Problems associated with recreation and leisure are a criminogenic need that appears among the "central eight" (Andrews and Dowden, 2007). (As noted above, the offender was no longer associating with antisocial peers, which reduces the dynamic risk associated with that need, but the offender was not involved in any prosocial leisure activities. He indicated that he thought his son's racing team would serve as a prosocial activity, but had not yet participated at the time of the interview.)

Table 1 indicates that officers named 16 other targets for supervision, all with modest numbers of officers listing them in their top three: housing; medical; criminal recidivism/violations this time; social support; relationships; isolation; education; structure; monitoring/surveillance/drug testing; safety; self-control; violence; stress/grief/guilt; child support; spiritual matters/cultural matters/Native American matters; and transportation/driver's license. In broad strokes, some of these categories resemble previously-described categories. For example, "social support," "isolation," and "relationships" all have aspects in common with both "criminal peers" and "family." "Stress/grief/guilt" may be closely linked to "mental health." However, instead of assuming what the probation officers intended and collapsing the categories, we retained the 27 discrete categories.

Risk as Identified by Probation Officers

Probation officers exhibited tremendous variation in their clinical assessment of the top three criminogenic needs. They also exhibited great variation in using their clinical judgment to assign offenders to a risk category. Table 2, below, depicts the officers' assessments of the offender's risk category.

Table 2: Estimated Risk Category

Risk Category	N	%
Low	28	3
Low/Moderate	323	31
Moderate	531	51
High	158	15

Table 2: Estimated risk category level as identified by 1,040 probation officers, by number and percentage.

Scores on the PCRA range from a low of 0 to a high of 18. According to the training staff, the offender's actual risk level was 6, representing the bottom end of the "low/moderate risk" category. Using their unstructured clinical judgment, nearly one-third of the officers (31 percent) correctly placed the offender in the low/moderate risk category, but two-thirds over-classified him: 51 percent of the

officers classified him as “moderate risk" and 15 percent of the officers classified him as “high risk.” This may be an example of the “precautionary principle” (Ansbro, 2010; Kemshall, 1998), a tendency to over-classify when potential costs of missing a true positive (i.e., not identifying a high-risk offender as such) can be enormous for a probation officer but costs of false positives (i.e., wrongly identifying a low-risk offender as high-risk), while serious, are not borne by the officer.

Assessment of offender risk also has implications for identification of criminogenic needs. As demonstrated below, a greater percentage of officers who correctly assessed the offender’s level of recidivism risk also accurately assessed his criminogenic needs than did officers who overestimated his risk.

Identification of Criminogenic Needs According to Risk Estimates

The data suggest that officers’ estimates of risk shaped their evaluations of criminogenic needs. For analytical purposes, we grouped the 27 different types of named criminogenic needs into four broad categories: top-tier criminogenic needs in the case (consisting of criminal attitudes, criminal peers, self-control, social support, and structure); lower-tier criminogenic needs in the case (education, employment, family and marital problems, housing, income, leisure/recreation, relationships, and substance abuse [drugs/alcohol]); responsivity factors or barriers to treatment in the case (child support, isolation, medical, mental health, monitoring/surveillance/drug testing, safety, spiritual matters/cultural matters/Native American matters, stress/grief/guilt, and transportation/drivers license); and controlling strategies in the case (criminal history/community supervision violations, firearms, criminal recidivism violations this time, and violence). Table 3, below shows the items and the specific categories to which they are assigned.

Table 3: Category Assignment of Supervision Targets

Top Tier	Lower Tier	Responsivity Factors	Controlling Strategies
Criminal Attitudes	Education	Child Support	Criminal History Concerns
Criminal Peers	Employment	Isolation	Firearms Monitoring
Self Control	Family Problems	Medical Issues	Monitoring Potential Violence
Social Support and Structure	Housing	Mental Health	Monitoring Supervision Violations
	Income	Monitoring	
	Leisure	Drug Testing	
	Relationships	Safety Issues	
	Substance Abuse	Spiritual Matters	
		Culture	
		Stress	
		Guilt	
		Transportation	

Officers who viewed the offender as representing a low or low/moderate risk were more likely to identify a top-tier criminogenic need (criminal attitudes, criminal peers, self-control, social support, or structure) than were officers who viewed the offender as constituting a moderate or high risk. Understandably, officers who viewed the offender as moderate or high risk were more likely to identify controlling strategies (e.g., searches related to the offender's criminal history, history with firearms, or potential for violence). Table 3, below, identifies the number of officers who, as part of their top three criminogenic needs, identified a top-tier criminogenic need, a lower-tier criminogenic need, a responsivity factor, and a controlling strategy, as grouped by estimated risk level.

Table 4: Identified Needs, by Estimated Risk Level

Risk	Top-Tier Identified[A]		Lower-Tier Identified[B]		Responsivity Factor Identified[C]		Controlling Strategy Identified[D]	
	Yes	No	Yes	No	Yes	No	Yes	No
Low	64%	36%	89%	11%	54%	46%	0%	100%
Low/Moderate	64%	26%	97%	3%	46%	54%	14%	86%
Moderate	44%	56%	97%	3%	53%	47%	22%	78%
High	42%	58%	98%	2%	61%	39%	26%	74%

Table 3: Identification of top-tier criminogenic need, lower-tier criminogenic need, responsivity factor, and controlling strategy, as grouped by 1,040 officers' estimated risk level.
A: X^2 (3) = 38.964; p < .001
B: X^2 (3) = 6.325; p < .100
C: X^2 (3) = 9.778; p < .05
D: X^2 (3) = 19.967; p < .001

Table 3 reveals systematic differences in the assessment of criminogenic needs as a function of estimated risk level. Probation officers who evaluated the offender as representing a low or low/moderate level of recidivism risk were more likely (64 percent) to identify a top-tier criminogenic need than were officers who evaluated the offender as representing a moderate risk (44 percent) or a high level of risk (42 percent). This relationship was strongly statistically significant (X^2 (3) = 38.964; p < .001).

The relationship between estimated risk level and identification of lower-tier criminogenic needs (e.g., family and marital problems, leisure/recreation, or substance abuse [drugs/alcohol]) was not as defined as the relationship between risk and top-tier needs (X^2 (3) = 6.325; p < .100). Only 89 percent of officers who estimated the offender's risk level as low identified a lower-tier need in their top three needs, while 97 percent of those who estimated his risk as low/moderate or moderate and 98 percent of those who estimated his risk as high did so.

Only 46 percent of officers who correctly assessed the offender's risk level as low/moderate identified a responsivity factor (e.g., medical issues, mental health, or stress) when naming their top three criminogenic needs. On the other hand, 54 percent of those who identified the offender's risk level as low and 53 percent of those who identified it as moderate did so. Among those who assessed the offender as high risk, 61 percent identified a responsivity factor when identifying the top three criminogenic needs. This relationship was weakly statistically significant (X^2 (3) = 9.778; p < .05).

There was a strong relationship (X^2 (3) = 19.967; p < .001) between the estimated level of recidivism risk and the identification of a controlling strategy (e.g., interventions related to the offender's history with firearm violations or potential for violence) as one of the top three criminogenic

needs. While 0 percent of the officers who estimated the offender's risk as low identified a controlling strategy as a need, 14 percent of those who estimated it as low/moderate, 22 percent of those who estimated it as moderate, and 26 percent of those who estimated it as high did so.

As a general matter, officers who correctly identified the offender's risk level as low/moderate were more likely than most to identify a top-tier criminogenic need, were equally likely to identify lower-tier needs, were less likely than most to name a responsivity factor, and were less likely than most to identify a controlling strategy. Officers who assessed the offender's risk level as low also did well, but officers who over-classified the offender's risk as moderate or high did less well.

We thought it might be useful to "drill down" into the data and to examine, by clinical judgment risk category assignment, the percentage of officers who identified a top-tier need, a lower-tier need, a responsivity factor, or a controlling strategy as their number-one listed criminogenic need. This information is depicted in Table 4, below.

Table 5: First-Listed Need, by Estimated Risk Level

Risk	Top-Tier Identified	Lower-Tier Identified	Responsivity Factor Identified	Controlling Strategy Identified
Low	36%	46%	18%	0%
Low/Moderate	24%	66%	8%	2%
Moderate	13%	71%	10%	6%
High	7%	71%	13%	9%

Table 4: Identification of top-tier criminogenic need, lower-tier criminogenic need, responsivity factor, and controlling strategy as first-listed need, as grouped by 1,040 officers' estimated risk level.

There is a strong statistical relationship between the estimated level of recidivism risk and the identification of top-tier, lower-tier, responsivity factors, or controlling strategies as first-listed need (X^2 (9) = 51.694; $p < .001$). Officers who deemed the offender a low risk were more likely (36 percent) to identify a top-tier need as the first-listed criminogenic need, followed by those who deemed him a low/moderate risk (24 percent), a moderate risk (13 percent), or a high risk (7 percent). On the other hand, officers who viewed the offender as low/moderate risk were less likely (46 percent) to identify a

lower-tier need as the first-listed criminogenic need than were those who viewed him as low risk (66 percent) or those who viewed him as moderate or high risk (both 71 percent). No clear pattern emerged in the identification of responsivity factors as a first-listed need, although those who correctly identified the offender as low/moderate risk were less likely to name a responsivity factor than other officers. A linear trend emerged in the identification of controlling strategies: while 0 percent of officers who deemed the offender a low risk named a controlling factor as their first-listed need, 2 percent of those who deemed him low/moderate risk did, 6 percent of those who deemed him moderate risk did, and 9 percent of those who deemed him high risk did. Similar patterns, although less distinct, also were evident in the distribution of second-listed needs. These are depicted in Table 5, below.

Table 6: Second-Listed Need, by Estimated Risk Level

Risk	Top-Tier Identified	Lower-Tier Identified	Responsivity Factor Identified	Controlling Strategy Identified
Low	22%	52%	26%	0%
Low/Moderate	29%	43%	22%	6%
Moderate	17%	44%	28%	11%
High	15%	40%	37%	9%

Table 4: Identification of top-tier criminogenic need, lower-tier criminogenic need, responsivity factor, and controlling strategy as second-listed need, as grouped by 1,040 officers' estimated risk level.

There is a strong statistical relationship between the estimated level of recidivism risk and the identification of top-tier, lower-tier, responsivity factors, or controlling strategies as first-listed need (X^2 (9) = 30.577; $p < .001$). Officers who correctly deemed the offender a low/moderate risk were more likely (29 percent) to identify a top-tier need as their second-listed need than were other officers. In general, officers who estimated risk as low or low/moderate were somewhat more likely to identify a top-tier need than those who estimated risk as moderate or high. Generally, as estimated risk increased, the percentages of officers identifying a lower-tier need as a second-named need decreased. Conversely, as estimated risk increased, the percentage of officers identifying a responsivity factor as a second-named need generally increased. The appearance of controlling strategies generally increased as

estimated risk level increased, although more officers who deemed the offender a moderate risk (11 percent) listed controlling strategies as their second-named need than officers who deemed him a high risk (9 percent). Similar patterns emerged in the distribution of third-listed needs, although they are also indistinct. Table 6, below, displays the data.

Table 7: Third-Listed Need, by Estimated Risk Level

Risk	Top-Tier Identified	Lower-Tier Identified	Responsivity Factor Identified	Controlling Strategy Identified
Low	28%	48%	24%	0%
Low/Moderate	21%	50%	21%	8%
Moderate	19%	51%	21%	9%
High	35%	43%	19%	12%

Table 4: Identification of top-tier criminogenic need, lower-tier criminogenic need, responsivity factor, and controlling strategy as third-listed need, as grouped by 1,040 officers' estimated risk level.

The distribution of third-listed needs was not statistically significant (X^2 (9) = 8.303; p < .504). Although it was not the case among first- and second-listed needs, for third-listed criminogenic needs, a greater percentage of officers who viewed the offender as high risk identified a top-tier need (35 percent) than did other officers. There do not appear to be clear trends, however, for the identification of top-tier or lower-tier needs as third-named criminogenic needs. There are trends in the responsivity and controlling strategy data, though. As estimated risk level increased, the percentage of officers naming a responsivity factor as second-named need decreased, from 24 percent (minimal risk) to 19 percent (high risk). On the other hand, as estimated risk level increased, the percentage of officers identifying a controlling strategy as second-named need increased, from 0 percent (minimal risk) to 12 percent (high risk).

Discussion

The federal probation and pretrial services system is in a dramatic state of punctuated evolution (Hughes, 2008: Alexander and VanBenschoten, 2008). Building upon the system's charter for excellence

and a comprehensive strategic assessment (Hughes, 2008), the federal probation and pretrial services system of today is results-driven and is committed to evidence-based practices (Judicial Conference, 2006). Essential to the successful adoption of evidence-based practices (Cullen and Gendreau, 2000) is an understanding of the principles of risk, needs, and responsivity (Andrews and Dowden, 2007).

The regional PCRA training sessions convened during 2010 and 2011 were a valuable step in this direction. They also afforded the authors a valuable opportunity to determine whether federal probation officers could better assess risk of recidivism and identify criminogenic needs when using the PCRA than when exercising their unstructured professional judgment. A substantial body of research suggested that they would (e.g., Grove and Meehl, 1996; Grove *et al.*, 2000; Meehl, 1954; Monahan *et al.*, 2001), but despite more than 50 years of research (Ægisdóttir *et al.*, 2006), many community corrections officers from jurisdictions around the world resent –and resist – actuarial assessments (Fitzgibbon, Hamilton, and Richardson, 2010; Horsefield, 2003; Lynch, 1998; Snyder, Ervin, and Snyder-Joy, 1996; VanBenschoten, 2008). Such adherence to custom and common sense while eschewing empirical, scientific knowledge about what works in corrections has been characterized as "correctional quackery" by some researchers (Flores et al., 2005; Latessa, Cullen, and Gendreau, 2002). The fact is that sound professional judgment informed by a valid actuarial instrument is superior to professional judgment alone. Similarly, the actuarial tool must be in the hands of a trained professional who can apply the tool correctly to the facts in a case, appropriately interpret the results, and override results in the limited instances where such an override is appropriate.

Studying probation officers' assessments of risk and criminogenic needs at the PCRA training meetings confirms the importance of adhering to evidence-based practices and underscores the essential role of officer training. Asked to identify the three most important targets of supervision to reduce risk by changing behavior in a training case, 1,040 probation officers (who had viewed the same video and examined the same materials) generated literally dozens of different answers. Even after

coding their answers into groups, officers had identified 27 different categories. Some of their assessments were on-point: more than 90 percent of the officers identified substance abuse as one of the top three targets and nearly half identified criminal peers as one of the three. But many of the targets named were not relevant in this case (e.g., employment difficulties or problematic family circumstances) and many of the "needs" the officers named in the case were not criminogenic needs at all (e.g., mental health, stress, or child support). This, however, is in no way unusual. In Edward Latessa's 2004 August Vollmer Award address, he observed that he often asks "correctional staff who work with offender's [*sic*] day in and day out what they think are the major risk factors associated with criminal conduct. They are often all over the map, and needless to say, I am often amazed with the list they come up with" (2004: 551). When Flores and his colleagues (2005) asked 171 juvenile justice officers to identify the most important criminogenic needs of juvenile delinquents, only 42.1 percent could name one of the "big four" and only 6.4 percent of the officers could name two. The authors wrote, "[I]t should come as no surprise that many rehabilitative efforts fail to produce positive treatment effects when those responsible for delivering the interventions are largely unaware of the most relevant criminogenic needs to target with those services" (2005: 12).

The reasons that officers struggled to identify the relevant criminogenic needs in the case is that they failed to have a uniform understanding of terms used to categorize risk and/or they disagreed with each other on the actual risk presented. While 31 percent correctly evaluated the offender's recidivism risk as "low/moderate," two-thirds of the officers estimated it as being higher. More than half (51 percent) assessed the risk level as being "moderate," and 15 percent assessed it as "high," even though, according to the PCRA, the offender's actual risk score was at the bottom of the "low/moderate" range (a PCRA score of 6). Such over-classification of risk may be evidence of the so-called precautionary principle (Ansbro, 2010; Kemshall, 1998) and should be worrying because evidence suggests that over-

supervising low-risk offenders can make them *worse*, affirmatively increasing their likelihood of recidivism (Lowenkamp and Latessa, 2004).

Presumably, regular use of the PCRA will improve the accuracy and consistency of risk assessments. After receiving training on the tool, probation officers asked to re-evaluate the offender's case showed greater consensus in their evaluations. During this second assessment, using the PCRA, no officers assessed the offender's risk level as "high," and more than 90 percent correctly assessed his risk level as "low" (Oleson et al., 2011).

Probation officers who correctly placed the offender into the "low/moderate risk" category also tended to identify top-tier criminogenic needs in their top-three needs and they tended not to identify responsivity factors as "needs." On the other hand, a greater percentage of the officers who deemed the offender to be a "high" risk named lower-tier criminogenic needs, responsivity factors, and controlling strategies in their top-three needs. The data suggest that officers who identified the offender as "low risk" might be more likely to address relevant criminogenic needs from the "big four" (i.e., antisocial attitudes, antisocial peers, antisocial cognition, and a history of antisocial behavior) and the "central eight" (i.e., family and marital problems, school or work difficulties, lack of involvement in non-criminal leisure activities and recreation, and substance abuse) than officers who identified the offender as "high risk" and might therefore be more likely to structure the offender's supervision around responsivity issues (e.g., physical health, mental health, or stress) and controlling strategies (e.g., drug testing, residence searches for firearms or drugs). What this means, of course, is that assessment of recidivism risk and criminogenic need will play a pivotal role in defining what kind of supervision experience a given offender will have. The need for accuracy and consistency in these foundational assessments, in our view, means that VanBenschoten was right when he stated that the use of an actuarial risk/needs tool is "[t]he cornerstone of effective supervision" (2008: 38). Given that actuarial risk assessment is one of the few useful checks against over-classification and the precautionary

principle (Bonta and Motiuk, 1990; Bonta, 2008), it means that Zinger was right when he wrote, "Failure to conduct actuarial risk assessments or consider its results is irrational, unscientific, and unprofessional" (2004: 607).

The evidence suggests that the PCRA serves an essential role in the ongoing evolution of the federal probation and pretrial services system. By operating as a check against the precautionary principle and over-classification, risk/needs assessment instruments of this kind can reduce recidivism among low-risk offenders by ensuring that they are not over-supervised and can reduce recidivism among high-risk offenders by ensuring that these individuals are provided with treatment interventions that correspond to their criminogenic needs. This focus on outcomes and evidence underlies the new rehabilitation of "what works" (Taxman, Shepardson, and Byrne, 2004). To be sure, this is a form of rehabilitation that takes public safety as its ultimate object – not the transformation of every individual offender (Robinson, 2002). But as we move away from "nothing works" (Martinson, 1974), through "what works" (Latessa and Lowenkamp, 2006; Petersilia, 2004), and toward "making what works work" (Andrews, 2006), it is entirely possible that "rehabilitation" will not only cease to be a dirty word but can become the watchword of a well-trained and professional federal probation and pretrial services system.

References

Ægisdóttir, S. et al., 2006. The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction. *Counseling Psychologist*, 34(3), 341-382.

Alexander, M. and VanBenschoten, S., 2008. The Evolution of Supervision within the Federal Probation and Pretrial System. *Journal of Offender Rehabilitation*, 47(3), 319-337.

Allen, F. A., 1981. *The Decline of the Rehabilitative Ideal: Penal Policy and Social Purpose*. New Haven, CT: Yale University Press.

Andrews, D. A., 2006. Enhancing Adherence to Risk-Need-Responsivity: Making Quality a Matter of Policy. *Criminology and Public Policy*, 5, 595-602.

Andrews, D.A. and Bonta, J., 1995. *The Level of Service Inventory-Revised*. Toronto: Multi-Health Systems.

Andrews, D.A. and Bonta, J., 2010. *The Psychology of Criminal Conduct* (5th ed.). New Providence, NJ: Anderson Publishing.

Andrews, D.A., Bonta, J., and Hoge, R.D., 1990. Classification for Effective Rehabilitation: Rediscovering Psychology. *Criminal Justice and Behavior*, 17, 19-52.

Andrews, D. A. and Dowden, C., 1999. A Meta-Analytic Investigation into Effective Correctional Intervention for Female Offenders. *Forum on Corrections Research*, 11, 18-21.

Andrews, D.A. and Dowden, C., 2007. The Risk-Need-Responsivity Model of Assessment and Human Service in Prevention and Corrections: Crime-Prevention Jurisprudence. *Canadian Journal of Criminology and Criminal Justice*, 49(4), 439-464.

Ansbro, M., 2010. The Nuts and Bolts of Risk Assessment: When the Clinical and Actuarial Conflict. *Howard Journal*, *49*(3), 252-268.

Birgden, A., 2004. Therapeutic Jurisprudence and Responsivity: Find the Will and the Way in Offender Rehabilitation. *Psychology, Crime & Law*, 10(3), 283-295.

Bonta, J., 2008. Offender Risk Assessment and Sentencing. *Canadian Journal of Criminology and Criminal Justice*, 49(4), 519-529.

Bonta, J. and Motiuk, L. L., 1990. Classification to Correctional Halfway Houses: A Quasi-Experimental Evaluation. *Criminology*, 28, 497-506.

Buddress, L. A. N., 1997. Federal Probation and Pretrial Services – A Cost-Effective and Successful Community Corrections System. *Federal Probation*, *61*(1), 5-12.

Christensen, G.E. and Crime and Justice Institute, 2008. *Our System of Corrections: Do Jails Play a Role in Improving Offender Outcomes?* Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Crime and Justice Institute, 2009. *Implementing Evidence-Based Policy and Practice in Community Corrections* (2nd ed). Washington, DC: U.S. Department of Justice, National Institute of Corrections

Cullen, F. T., 2005. The Twelve People Who Saved Rehabilitation: How the Science of Criminology Made a Difference. *Criminology*, 43(1), 1-42.

Cullen, F. and Gendreau, P., 2000. Assessing Correctional Rehabilitation: Policy, Practice, and Prospects. In J. Horney (Ed.), *Criminal Justice 2000* (pp. 109-175). Washington, DC: U.S. Department of Justice.

Fahey, J. and Crime and Justice Institute, 2008. *Using Research to Promote Public Safety: A Prosecutor's Primer on Evidence-Based Practice*. Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Fitzgibbon, W., Hamilton, C., & Richardson, M., 2010. A Risky Business: An Examination of Irish Probation Officers' Attitudes towards Risk Assessment. *Probation Journal*, *57*(2), 163-174.

Flores, A. W. et al.., 2005. Evidence of Professionalism or Quackery: Measuring Practitioner Awareness of Risk/Need Factors and Treatment Strategies. *Federal Probation*, 69(2), 9-14.

Gendreau, P., 1996. The Principles of Effective Interventions with Offenders. In A. T. Harland (Ed.), *Choosing Correctional Options that Work: Defining Demand and Evaluating the Supply* (pp. 117-130). Thousand Oaks, CA: SAGE.

Gendreau, P., Little, T., and Goggin, C., 1996. A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works! *Criminology*, 34(4), 575-607.

Grove, W. M., & Meehl, P. E., 1996. Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy. *Psychology, Public Policy, and Law*, *2*, 293-323.

Grove, W. M. et al., 2000. Clinical Versus Mechanical Prediction: A Meta-Analysis. *Psychological Assessment*, *12*(1), 19-30.

Halleck, S. L. and Witte, A. D., 1977. Is Rehabilitation Dead? *Crime & Delinquency*, 23, 372-383.

Horsefield, A., 2003. Risk Assessment: Who Needs It? *Probation Journal*, *50*(4), 374-379.

Hughes, J., 2008. Results-Based Management in Federal Probation and Pretrial Services. *Federal Probation*, 72(2), 4-14.

Jalbert, S. K., et al.., 2011. *A Multi-Site Evaluation of Reduced Probation Caseload Size in an Evidence-Based Practice Setting* (2006-IJ-CX-0011). Washington, DC: U.S. Department of Justice, National Institute of Justice.

Judicial Conference, 2006. *September Proceedings of the Judicial Conference of the United States*. Washington, DC: Administrative Office of the United States Courts.

Kemshall, H., 1998. Defensible Decisions for Risk: Or ''It's the Doers Wot Get the Blame.'' *Probation Journal*, *45*(2), 67–72.

Latessa, E. J., 2004. The Challenge of Change: Correctional Programs and Evidence-Based Practices. *Criminology & Public Policy*, 3(4), 547-560.

Latessa, E. J., Cullen, F.T., and Gendreau, P., 2002. Beyond Correctional Quackery—Professionalism and the Possibility of Effective Treatment. *Federal Probation*, 66(2), 43-49.

Latessa, E. J. and Lowenkamp, C. T., 2006. What Works in Reducing Recidivism? *University of St. Thomas Law Review*, 3(3), 521-535.

Lowenkamp, C. T., & Latessa, E. J., 2004. Understanding the Risk Principle: How and Why Correctional Interventions Can Harm Low-Risk Offenders, in *Topics in community corrections* (pp. 3-8). Washington, DC: National Institute of Corrections.

Lowenkamp, C. T. and Latessa, E. J., 2005. Increasing the Effectiveness of Correctional Programming through the Risk Principle: Identifying Offenders for Residential Placement. *Criminology and Public Policy*, 4, 501-528.

Lowenkamp, C. T., Latessa, E. J., & Holsinger, A. M., 2006. The Risk Principle in Action: What Have We Learned from 13,676 Offenders and 97 Correctional Programs? *Crime & Delinquency*, *52*, 77-93.

Lowenkamp, C. T., et al.., 2006. Adhering to the Risk and Need Principles: Does it Matter for Supervision-Based Programs? *Federal Probation*, 70(3), 3-8.

Lynch, M., 1998. Waste Managers? The New Penology, Crime Fighting, and Parole Agent Identity. *Law and Society Review,* 32(4), 839–869.

Martinson, R., 1974. What Works? Questions and Answers about Prison Reform. *Public Interest*, 35 (Spring), 22-35.

Meehl, P. E., 1954. *Clinical Versus Mechanical Prediction*. Minneapolis, MN: University of Minnesota Press.

Monahan, J. et al., 2001. *Rethinking Risk Assessment*. New York: Oxford University Press.

Oleson, J.C. et al., 2011. Training to See Risk: Measuring the Accuracy of Clinical and Actuarial Risk Assessments among Federal Probation Officers. *Federal Probation*, in press.

Petersilia, J., 2004. What Works in Prisoner Reentry? Reviewing and Questioning the Evidence. *Federal Probation*, 68(2), 4-8.

Robinson, G., 2002. Exploring Risk Management in Probation Practice. *Punishment and Society*, *4*(1), 5-25.

Sackett, D. L., et al., 2000. *Evidence-Based Medicine: How to Practice and Teach EBM* (2nd ed.). New York: Churchill Livingstone.

Schneider, A. L., Ervin, L., and Snyder-Joy, Z., 1996. Further Exploration of the Flight from Discretion: The Role of Risk/Need Instruments in Probation Supervision Decisions. *Journal of Criminal Justice*, 24(2), 109-121.

Scott, W. and Crime and Justice Institute, 2008. *Effective Clinical Practices in Treating Clients in the Criminal Justice System*. Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Serin, R. and Crime and Justice Institute, 2005. *Evidence-Based Practice: Principles for Enhancing Correctional Results in Prisons*. Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Taxman, F. S., Shepardson, E. S., and Byrne, J. M., 2004. *Tools of the Trade: A Guide to Incorporating Science into Practice*. Washington, DC: National Institute of Corrections.

VanBenschoten, S. (2008). Risk/Needs Assessment: Is This the Best We Can Do? *Federal Probation*, *72*(2), 38-42.

VanNostrand, M. and Crime and Justice Institute, 2007. *Legal and Evidence-Based Practices: Applications of Legal Principles, Laws, and Research to the Field of Pretrial Services*. Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Ward, T., Melser, J., and Yates, P., 2007. Reconstructing the Risk-Need-Responsivity Model: A Theoretical Elaboration and Evaluation. *Aggression and Violent Behavior*, 12, 208-228.

Warren, R., and Crime and Justice Institute, 2007. *Evidence-Based Practice to Reduce Recidivism: Implications for State Judiciaries*. Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Weibrecht, K. and Crime and Justice Institute, 2008. *Evidence-Based Practices and Criminal Defense: Opportunities, Challenges, and Practical Considerations*. Washington, DC: U.S. Department of Justice, National Institute of Corrections.

Zinger, I., 2004. Actuarial Risk Assessment and Human Rights: A Commentary. *Canadian Journal of Criminology and Criminal Justice*, 46(5), 607-620.