This is to certify that the
dissertation entitled

Risk Assessment: Toward Valid and Accurate Delinquency Predictions with African-Americans and Girls

presented by

Eyitayo Onifade

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Psychology

Major Professor's Signature

8/21/08

Date

RISK ASSESSMENT: TOWARD VALID AND ACCURATE DELINQUENCY PREDICTIONS WITH AFRICAN-AMERICANS AND GIRLS

By

Eyitayo Onifade

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Psychology

2008

ABSTRACT

RISK ASSESSMENT: TOWARD VALID AND ACCURATE DELINQUENCY PREDICTIONS WITH AFRICAN-AMERICANS AND GIRLS

By

Eyitayo Onifade

Recidivists typically comprise less than ten percent of the offender population, yet can commit the majority of offenses in a given year. Re-offenders share characteristics in common (criminogenic risk factors), which can be used to identify young offenders that are most likely to recidivate. These criminogenic risk factors are modeled and combined into multi-domain measures that juvenile justice systems use in a practice known as risk assessment. While the YLS/CMI is generally considered a good risk measure, little work has been done validating its predictive validity with African Americans and girls.

This study explored the risk-recidivism relationship between YLS/CMI score and re-offense rates over a 24 month period for African Americans, Whites, boys, and girls. The YLS/CMI demonstrated its ability to differentiate non-recidivists and re-offenders moderately well for each of the subgroups. Recidivists on average scored four points higher on the YLS/CMI than non-recidivists. In turn, the YLS/CMI had an AUC of .63, however, subgroup differences in

distribution of recidivism by risk level were found. African American low and moderate risk boys had substantially higher re-offense rates than their counterparts, while White low and moderate risk girls had substantially lower re-offense rates. This led to over and under predictions of re-offense for the two respective subgroups.

In analyzing the risk-recidivism relationship in this manner, the study was narrow in scope in that it did not account for risk factors excluded from the YLS/CMI on the community level like police surveillance levels that may moderate these relationships. Moreover, this study did not seek to ascertain the causal pathways between delinquency and the various risk factors. Further, socioeconomic status was not controlled for in parceling out possible moderating effects on the risk-recidivism relationship. Instead given the current stage in research on the dimensional identity of risk assessments like the YLS/CMI, the first priority was determining whether offense rates differed within risk levels by gender and racial group. Commentary could then be made about whether the instrument over-predicted or under-predicted future delinquency for the various groups. With that information future research can now focus on improving the dimensional identity of the YLS/CMI.

To Eniola for just being

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge and thank my beautiful wife, Tiffiani. We have been together on this journey so long, it's difficult, sometimes, to separate who I am from what she's encouraged me to become. We play the what if game, and in all honesty, her beautiful spirit inspires me to be better than I am. So without a doubt this dissertation, let alone my Ph.D would not have been possible without her to support and encourage me along the way. I thank Joy for letting me be a better scholar than husband at times. I thank Joy for listening to the same speech over and over again about the just us system and what we need to do about it. I thank Joy for being a great wife and even better Joy. You have my love, heart and soul plus.

Secondly, I would like to acknowledge and thank my beautiful family, Eniola, Sade, Kike, Kemi, Ciru, and Mommy. Believe it or not, our experiences together are the backbone of the dream, I dream for our people.

Thirdly, I would like to acknowledge and thank my brilliant father, who is my hero and the only human being I have ever strived to be like. It's my nature to toss directions to the side and dive into putting the pieces together based off a picture on the box. Fortunate for me

the blueprint has never been too far from hand and customer support is always available 24-7. Now that I am older and slightly wiser, I find myself reaching for the directions more and more and the product is finally manifest. Daddy, thanks for being my teacher.

Finally, I would like to acknowledge and thank my advisor, Dr. William Davidson. Four years ago, I walked into his office and I immediately realized this was the place for me, in spite of the weather and distance from home. Bill gave me room to grow and showed me real power and influence. He pushed me when I needed to be pushed and kept me grounded, when I was dreaming. I learned so much about research and research that makes a difference. But more importantly he acknowledged the enormity of life and allowed me to do the same. This made all the difference for me as a father, husband, brother and son; affording me the rare privilege of having a life beyond grad school. Thanks, Bill for everything.

TABLE OF CONTENTS

LIST OF TABLES....iv

INTRODUCTION....1

CHAPTER 1
The Criminal as Different: Crime Theory and the Impact of Poor Application....10

CHAPTER 2
Study Rationale....37

CHAPTER 3
Risk Assessment....44

CHAPTER 4
Validity, Accuracy and Dimensional Identity of Risk Assessments....57

CHAPTER 5
Rationale for Validating the YLS/CMI....79

CHAPTER 6
Methodology....83

CHAPTER 7
Results....91

CHAPTER 8
Discussion....105

CHAPTER 9
Conclusions....115

LITERATURE CITED....119

LIST OF TABLES

Table 1. Description of validation studies with risk assessment instruments.....60

Table 2. Results of validation studies.....61

Table 3. Dimensional identity of risk assessment instruments.....63

Table 4. Corrected Item-Total Correlations for the Intake and Probation Samples.....86

Table 5. Item Frequencies (% Yes).....91

Table 6. Subscale scores and ROC by race and gender.....96

Table 7. Descriptive Statistics.....97

Table 8. Hierarchical Logistic Regression for YLS/CMI score and New Offenses.....100

INTRODUCTION

There exists a rarely spoken of nexus between recidivism risk assessment and the reduction of minority and marginal group contact with the juvenile justice system. In 2007, the Office of Juvenile Justice and Delinquency Prevention [OJJDP] began to include standardized risk assessments in its database of best practices for systems plagued by disparate incarceration and arrest rates for girls and minorities (OJJDP, 2007). That said, the extent to which different racial and gender groups officially offend at disparate rates and severity is reflected in annually reported justice statistics and is hardly a matter of debate. For instance, African Americans comprise 16% of the youth population in the United States, yet comprise 29% of the delinquency caseload (Snyder & Sickmund, 2006). Moreover, despite accounting for no more than a third of the juvenile population, African Americans and Hispanic youth comprise nearly two-thirds of the juvenile detention population (Snyder & Sickmund, 2006). With respect to gender, girls are entering the justice system at an alarming rate for offenses traditionally associated with their male counterparts (Chesney-Lind &

Sheldon, 2004). For instance between 1970 and 2004, there was a 322% increase in violent crime for girls compared to a 107% increase in violent crime for boys in that same period (Chesney-Lind & Sheldon, 2004). While girls have historically comprised the majority of status offenders (e.g. truancy, running away, consumption of alcohol, incorrigibility); between 1994 and 2004 girls experienced an increase in arrest rates for Index Crimes like homicide, robbery, aggravated assault, arson, weapon possession and drug offending (Chesney-Lind & Sheldon, 2004). So while girls commit fewer officially reported offenses than boys, there is some concern with the extent to which the rate of increase is outpacing that of their male counterparts.

The causes of these disparities in officially recorded offense rates have been the subject of many empirical studies and qualified conjecture by professionals in fields as dissimilar as criminology and public health. However, one perspective in particular, the conflict theory of law, has managed to sustain itself over the past decades (Jackson, 1989). This theory suggests power is unequally distributed and crime control is often used to mitigate threats to the power group's interests (Jackson, 1989). Moreover, disparate arrest rates are a product of certain racial and gender groups receiving inordinate attention

from the criminal justice system that occurs on multiple levels of the justice process from differential street level policing to differential sentencing (Liska, 1992). This perspective may be reflected in the conflicting delinquency rates provided by national self report surveys versus official arrest report data. For example, two national self report studies [Monitoring the Future Survey and National Youth Survey] showed no significant increases in juvenile violence between 1980 and 2000, despite the aforementioned substantial fluctuations in arrest rates across race and gender groups (Howell, 2003).

In a pivotal meta-analysis, Pope and Feyerham (1992) found that most of the disparity in marginal group contact with law enforcement was found at the arrest and intake level of the justice process. Other studies attribute this differential treatment to unfavorable attitudes and perceptions that too many practitioners and professionals have of members of certain social groups (Guevera, Herz, & Spohn, 2006; Leiber & Fox, 2005; Schrantz & McElroy, 2000). In essence, the system is subject to a Pygmalion effect, a self-fulfilling prophecy where certain groups are expected to commit crimes and therefore are targeted by law enforcement while other groups are not subject to such treatment. This extra attention results in a larger

proportion of undesirable social behavior of certain groups being discovered and receiving a formal system response.

While the validity of this perspective is widely and hotly debated, the presumed role human bias plays in the disparity in minority and marginal group contact with the justice system has resulted in far-reaching policy implications. The Juvenile Justice and Delinquency Prevention Act of 2002 now requires that states receiving formula grants:

> "address juvenile delinquency prevention efforts and system improvement efforts designed to reduce, without establishing or requiring numerical standards or quotas, the disproportionate number of juvenile members of minority groups who come into contact with the juvenile justice system" (see section 223(a)(22))."

In the midst of this civil debate and system change, criminal justice experienced a concurrent but separate evolution of the practice known as recidivism risk assessment. A plethora of research has demonstrated that as few as 6% of offenders can commit the majority of offenses (Cottle, Lee, & Heilbrun, 2001; Wolfgang, Figlio, & Sellin, 1972; see also Schumacher & Kurz-Gwen, 2000). Consequently, a cornerstone strategy of crime reduction is the

identification of individuals comprising that small group of offenders (Bonta, 1996). In the past, this process of risk classification was managed by a practitioner or clinician who predicted whether a child was likely to re-offend based on the practitioner's professional opinion and experience. Since then psychometricians and researchers have demonstrated that practitioners and clinicians often ignore base-rates in addition to making spurious correlations and incorrect causal attributions (Grove and Meehl, 1996; Bonta, 1996). Formal actuarial assessments are consequently considered more accurate and less susceptible to bias than informal assessments (Grove and Meehl, 1996; Bonta, 1996). As such, a myriad of risk prediction instruments have become popular in the prediction of delinquency and recidivism.

Formal risk assessment instruments are typically comprised of risk factors (e.g. age of first offense) that have been demonstrated to be associated with delinquency and recidivism (Bonta, 1996). The cumulative effect of multiple risk factors present in a youth's life is then used to predict the likelihood of that child experiencing a particular undesirable outcome like further delinquency or recidivism. Moreover, modern instruments like the Youth Level of Service/Case Management Inventory utilize both

static and dynamic risk factors. That is to say these instruments assess risk factors that are either malleable or unresponsive to change efforts. For example, age of first offense is a static risk factor while education is dynamic because it can change. This practice of risk assessment is evolving in the direction of using combinations of risk factors to predict recidivism with recent research demonstrating risk patterns are just as important as total risk (Onifade et al., 2006). The early verdict of most researchers has largely been in favor of risk assessments, showing that the most recent generation of risk assessments are both valid and reliable predictors of delinquency and recidivism (Schwalbe, 2007).

Therein lays the nexus between risk assessment and the reduction of disproportionate minority and marginal group contact with the juvenile justice system. Theoretically, in so much as risk assessments are valid predictors of recidivism with diverse populations, risk assessments by their very nature could reduce the system's contribution to disparities in adjudication rates through squeezing out the biases found at the early stages of the justice process. However, little study has been conducted on the validity of risk assessments with girls, and African Americans. Therefore this posited curative relationship between risk

assessment and the reduction of disparities is presumptive at best. Furthermore, these instruments are administered by practitioners, begging the question how the presumed bias of the practitioners will not be reflected in something as basic as the scoring of the risk assessment. Moreover, it is likely that the risk-recidivism statistical relationship is naturally different for members of marginal groups. If conflict theory is correct, members of these groups would be at greater risk of having their criminal activity exposed due to disparate levels of policing, surveillance, and attention from the justice system in their communities. As even multi-domain risk assessments do not include policing levels, community surveillance, and justice system attention as predictive risk factors, the instruments should under-predict offending for these groups. This is especially disconcerting when one considers there is evidence that the very factors used in some risk assessments have been shown to have little to no relationship with delinquency for girls and other minorities. One should also keep in mind "arrest rates" only reflect the rate of crime officially recorded by police. Moreover, the differential validity of the instrument may be reflective of the inherent bias in the recorded and reported arrest rates, if juveniles with the

same risks have substantially and significantly different arrests rates.

The danger in presuming a curative relationship between risk assessment and justice system bias is demonstrated in the case of the North Carolina Assessment of Risk Instrument (NCAR). In a recent validation study of the measure, researchers compared the validity coefficients of African American (n=233) and White young offenders (n=231) over a 12 month period (Schwalbe, Fraser, Day, & Arnold, 2004). A Cox linear regression revealed that the NCAR showed significant differences between the high, moderate, and low risk offenders in their time to new offense. Furthermore, high risk offenders were nearly twice as likely to re-offend as moderate risk offenders. However, when the hazard analysis was extended partitioning out African Americans, Whites, boys and girls, the researchers found significant differences in the risk/offending patterns of the subpopulations. Males reached the base-rate of re-offense with lower scores on the NCAR than girls. In turn, African Americans also reached the base-rate of offense with lower scores on the NCAR than Whites. There were also differences in the pattern of offense with regards to the time it took for an offender of a particular group with an NCAR score to re-offend.

The NCAR study offered an ideal model of the validation process for risk assessment measures used in juvenile justice systems. The focus, then, of this study was to similarly examine the predictive validity of a risk assessment used in the local courts, the Youth Level of Service Case Management Inventory (YLS/CMI), across gender groups and with Whites and African-Americans in a young offender population. Clearly there is much work to be done in the field of risk assessment, validating this measure with a cross section of juvenile offenders. Thus, this study answers this question of predictive accuracy with regards to recidivism based on cumulative risk across racial and gender groups.

In addition to providing a systematic synthesis of juvenile risk assessment literature, the following literature review offered a review of the historical context of the criminal justice response to undesirable social behavior, provided a review of evidence and theories explaining disparate offending and incarceration rates, and reviewed ecological theory's role in improving the accuracy of risk assessments and improving upon their dimensional identity.

CHAPTER 1

The Criminal as Different: Crime Theory and the Impact of Poor Application

If one thinks of crime as undesirable social behavior so egregious that informal sanction, mores, and traditions are incapable of providing a sufficient response, then the criminal justice system is a formalized system of social control by society (Cullen & Agnew, 2006). Society has mechanisms in place that compel individuals to conform to its expectations and norms through punishment, deterrence, or rehabilitation and criminalization of those respective behaviors (Kraska, 2006). Actors in systems of justice response are faced with a tough question whose simple wording belies our collective difficulty in responding to it. Why do some people commit crimes, while others do not?

There are a number of theories offered in response to this basic question. Biochemistry, Strain, Labeling, Rational Choice, Social Control, Social Disorganization, and Social Learning are all popular theoretical perspectives explaining undesirable social behavior so egregious that informal sanctions and traditions are incapable of providing a sufficient response (Cullen & Agnew, 2006).

Biogenetic theories offer perhaps the oldest explanation of criminal behavior as exemplified by Cesare Lombroso, who in as early as 1876 contended criminality was an inheritable trait with identifiable physical characteristics (Burke, 2001). Contemporary biogenetic theories suggest heredity, temperament, traits and physical maladies cause a predisposition to crime (Walsh, 2000).

Strain theorists following in the line of Durkheim, Merton, Cohen, Agnew and Messner essentially state social structures within society encourage or discourage citizens to commit crime (Akers, 2000). Depending on the adequacy of social structures in facilitating the goal achievement of individuals, those individual's perception of their means and opportunities affect their willingness to adhere to society rules in achieving those goals (Akers, 2000).

Labeling theorists hypothesize individuals are highly susceptible to the stigmas society assigns to them; consequently the individual has little choice but to conform to the behavioral expectations of them whether those behaviors are deviant or not (Braithwaite, 1989).

Rational Choice theorists offer that individuals are rational and engage in cost/benefit analysis when determining a best course of action; therefore, individuals freely choose to act in their best interest according to

their calculations (Liska, 1987). Crime then is a result of law's failure to act swiftly, with certainty, and severity to deviant behavior, consequently criminal behavior is individually calculated to be in the best interest of the individual offender (Liska, 1987).

Social Control theorists believe that people accumulate relationships, values, norms, and beliefs through a socialization process which encourage or discourage crime (Cullen & Agnew, 2006). The socialization process occurs through direct control (punishments and rewards), indirect control (identification with non-criminals) and internal control (an active conscience). Crime is therefore a product of degradation of social controls (Cullen & Agnew, 2006).

Social Disorganization theorists attribute crime to the absence or breakdown of community institutions and communal relationships that typically facilitate cooperation for mutual benefit among members of a community (Jensen, 2003).

Social Learning theorists explain crime as a learned behavior which occurs through the modeling of others' behavior, the results of which are reinforced by certain outcomes (Cullen & Agnew, 2006).

The one common denominator in most of these perspectives is that those that commit crimes are different in their person or context; in turn, the justice response has always hinged on the identification of individuals with those differences and addressing the criminogenic (crime causing) factors that comprise those differences between offenders and law-abiding citizens. In identifying criminogenic differences, society has tried everything from phrenology to profiling to standardized risk assessments. Each of these presupposes criminals have identifiable differences that can be relied upon to separate out individuals that society can target with criminal justice resources. The differences between these methods depend on the theory from which they originate. With the wide cross-section of theories explaining crime, the resulting system is an amalgamation of crime reduction models with often counterintuitive and conflicting means that either widen or narrow the adjudication net depending on any number of factors including but not limited to age, race, class, gender, community, and policing philosophy (Pope & Feyerham, 1992). Specifically, these disparate methods of crime control occur in three ways:

- through emphasizing law enforcement in particular communities

- the criminalization of undesirable behavior associated with certain groups
- the decision making of criminal justice practitioners at various junctions of the adjudication process (Schrantz & McElroy, 2000).

Disparate Methods of Crime Control

First, with regards to over-emphasis of law enforcement in certain communities, communities vary in the degree to which they use the criminal justice system to respond to undesirable and maladaptive behavior (Stucky, 2005). It is also widely observed that socioeconomic conditions like adequate education and employment opportunities, access to health care, and treatment programs can reduce problem solving to the overuse of the criminal justice system in certain communities (Loftin & McDowall, 1982; Maguire, 2001; Stucky, 2005).

Indeed, studies of police strength and policing levels have demonstrated that economic class and race are significant factors in policing levels (Stucky, 2005). Both economic and racial conflict theories presume resources are finite and, consequently, power groups act in their best interests, which is often detrimental to other groups. As this relates to economic class or race, policing levels are less related to crime levels than they are the perceived

threat from the marginalized group (Stucky, 2005). Thus, local government systems that are more susceptible to political pressure from the power group are more likely to increase policing resources in marginal communities (Stucky, 2005; Wilson, 1968).

In support of this contention, several studies have demonstrated a positive association between inequality and policing levels (Sever, 2003; Jackson & Carroll, 1981; Loftin & McDowall, 1982; Maguire, 2001). In a three-year longitudinal study of 943 cities, Stucky (2005) demonstrated that there was a significant positive association between African-American presence and police employment. Moreover, cities that had partisan elections were more likely than their counterparts to have more police employees per 1000 residents. Kent and Jacobs (2005) found similar results, when examining policing levels in large US cities between 1980 and 2000. Cities with appointed managers as opposed to elected officials were less likely to hire additional police support in spite of minority presence.

The difference in communities' willingness to use the criminal justice system is also dependent upon the availability of alternative systems that can appropriately address the respective undesirable and maladaptive

behaviors of individuals in that community. For instance, it is commonly observed that nearly 7% of inmates admitted to jails annually have a mental illness (Freudenberg, 2002). Moreover, 75% of this number has a co-occurring substance abuse disorder (Freudenberg, 2002). Consequently, mental health professionals and other advocates often recommend diversion to community-based services that treat illnesses; however, in 1994 as few as 34% of jails responding to a survey stated that they offered diversion services for the mentally ill (Freudenberg, 2002).

Due to economic restraints, huge gaps in allocations of mental health resources remain intractable for low income communities, making criminal justice a primary source for treatment of the homeless and indigent. This may explain why some communities treat maladaptive behavior as a treatable illness; while other communities are more likely to use punitive sanctions available through law enforcement due primarily to disparate funding levels between the criminal justice and mental health systems (Weisman, Lamberti, & Price, 2004).

This is further demonstrated by the extent to which juvenile offenders often have dual status as victims in the child welfare system. Some communities use contact with juvenile justice as a trigger mechanism for diversion into

the child welfare system that is better suited to deal with family issues (Phillips & Bloom, 1998). However, this practice varies by community and thus contributes to disparities in officially reported offense rates for various groups (Phillips & Bloom, 1998). Some communities allow their schools to adopt zero- tolerance practices which ultimately decrease the ratio of reported and unreported delinquency (Mauer, 1999). Consequently, these communities will have a higher rate of "crime" than their counterparts whether the prevalence of the underlying undesirable behavior is similar or not. Differences in the degree to which communities rely on the juvenile justice system to respond to delinquency affect the reported crime rate for that community.

Second, communities can selectively enforce laws based on group membership (Taxman, Byrne, & Pattavinia, 2005). For instance, African Americans are more likely than Whites to be arrested for drug offenses, yet Whites have higher drug usage rates according to public health sources (Freudenberg, 2002). Over the past two decades, sentencing guidelines for drug involvement have dramatically changed. The majority of changes are in regards to the use of prisons to incarcerate and thus incapacitate offenders for long periods of time. Possession of a controlled substance

typically resulted in a sentence of 18 months in 1990, yet had increased to 30 months over the next ten years (BJS, 2004).

Given the differences in arrest rate for drug offenses in African American communities, the further criminalization of drug involvement disproportionately affected this group (Taxman, Byrne, & Pattavinia, 2005). A specific example commonly offered is that of laws targeting groups associated with the trafficking of crack cocaine. Crack cocaine use and sale is largely associated with African Americans, while powder cocaine is associated with other groups; however sentencing mandates require significantly longer sentences for possessing smaller amounts of crack than powder cocaine, in spite of their comparable deleterious effects on the user (Mauer, 1999). While the justification for this difference in sentencing is that the trafficking of crack has far greater deleterious effects on communities, the selective criminalization in itself contributes to disparities in system contact. This effect is further seen in the treatment of female status offenders. In a review of self-report studies conducted at the peak of arrest disparities between girls and boys for status offenses; Steffensmeier and Steffensmeier (1980) found that boys reported running

away just as much as girls, yet girls had a substantially higher likelihood of being arrested for this behavior.

One also notes selective criminalization of marginal groups can occur through targeting certain neighborhoods and areas with justice resources and policies (Eck & Weisburd, 2004). Curfews are often passed that make it a violation for youth of a certain age to be out after a certain time at night, yet those ordinances are differently enforced at a neighborhood level.

In an effort to fight the proliferation of gangs in certain communities, ordinances have been passed that prohibit individuals that "look like" gang-members from loitering or aggregating in a particular area for extended periods of time. The Supreme Court struck down a city ordinance that used just such language in prohibiting said behavior, but recommended the law be re-drafted to specify such loitering cannot be done with the purpose of facilitating "criminal conduct." With greater surveillance of these communities, the contributions of criminals to the larger cultural trends of the community they belong to can also become a source of selective criminalization.

For instance, the City of Atlanta has passed an ordinance prohibiting "sagging", a fashion trend that originated in prisons but matriculated into the larger

youth culture over time and is now commonly associated with African American males. Opponents of the code contend this law unfairly targets African American youth while allowing other popular fashion trends such as low ride cut jeans and tank tops that are popular amongst other groups.

Finally, disparities in contact can occur through the basic decision-making of justice officials. Prosecutors, for instance, often have wide latitude as to plea negotiations. These decisions can be swayed by prevailing public opinions at the time of those decisions. Despite sentencing guidelines and mandates, prosecutorial discretion allows for significant departures from these guidelines based on the decisions of the prosecution which are not reviewable under the law (Kerstetter, 1990). Prosecutors thus have discretion in three ways: the filing of charges, the level of charges and the dismissal of charges (Albonetti, 1987).

The US Sentencing Commission revealed that in 2002, the average departure rate was 17%; however, districts varied as much as 35% across the United States (Hartley, Maddan, & Spohn, 2007). Moreover, this rate of departure from sentencing guidelines varies by offense type with drug offenses and property offenses differing in departure rates by as much as 15% (Hartley, Maddan, & Spohn, 2007).

Numerous studies have indicated that while legal factors such as strength of evidence (Albonetti, 1987) and seriousness of the crime (Albonetti, 1987; Schmidt & Steury, 1989) affect a prosecutor's decision to pursue a case, characteristics that are legally irrelevant also have impact, such as racial composition of the suspect and victim (Keil & Vito, 1989; Spohn & Spears, 1996), the defendant's gender (Nagel & Hagan, 1983; Spohn & Spears, 1996) or employment status (Schmidt & Steury, 1989). Moreover, studies have shown that aforementioned crack-powder cocaine sentencing disparity is exacerbated by prosecutorial discretion where prosecutors are less likely to depart from sentencing guidelines for African Americans than Whites (Hartley, Maddan, & Spohn, 2007).

Furthermore, police have some degree of choice in whether to warn and let go, or arrest and detain, with that decision affected by many factors including their appraisal of the offender. Moreover, a defendant's access to resources also plays a part in whether alternative means of treatment or problem-solving are used to deal with their behavior.

Disparate System Contact Rates

One then quickly arrives at the sensitive subject spoken to earlier, where individuals that share

characteristics in common, criminogenic or not, have disparate amounts of contact with the justice system as a result of the aforementioned crime control methods. For instance, in the past decade, the number of girls referred to the court for property and person offenses increased at a greater rate than for boys. Indeed, young girls with maltreatment and neglect histories were twice as likely as their peers to be arrested in adulthood (Maxfield & Widom, 1996) with much of that disparity attributed to a failure of service providers to adequately identify and respond to the needs of this vulnerable population when the youth were most responsive to care. While 9% of female status offenders were placed in detention facilities in the mid 80s, only 1.5% of their male counterparts received such treatment (Schwartz, Steketee, & Schneider, 1989).

Moreover, young black offenders were 30 times more likely than white offenders to be placed in detention facilities for drug offenses despite sharing similar drug usage rates (Pope and Feyerherm, 1992). In a widely cited criminal justice study, Alfred Blumstein attributed 24% of racial disparities in imprisonment to criminal histories and racial bias. While African Americans constituted 15% of drug users in 1998, African Americans comprised 37% of those arrested for drug offenses. The numbers of

disparities are innumerable and as such warrant special attention to any crime reduction methods that depend on identifying differences.

Impact of Disparate Contact

The impact of disparate contact with the justice system is far-reaching; affecting individuals, the community and law-enforcement at large. For our justice system to be viable, the public must have confidence in the justice process. While it is convenient and even effective to identify criminogenic differences amongst individuals, this element of law enforcement should not devolve into discrimination. Quite simply, disparate contact naturally concentrates iatrogenic effects of the justice process in communities that experience the disparity. Moreover, we see iatrogenic effects across a broad spectrum of social living, from educational achievement to community health to political empowerment.

For instance, a growing body of research has demonstrated the role that disparate contact plays in destabilizing at-risk communities via the breakdown of families (Courtwright, 1996; Freudenberg, 2002). Latina and African American women are 3 and 7 times, respectively, as likely as White women to be incarcerated in their lifetime. Women that are incarcerated are often separated from their

children, and their children often end up in non-familial foster care where the children are at higher risk for psychological, educational and social problems (Phillips & Bloom, 1998). Moreover, nearly 48% of women behind bars report being sexually assaulted while incarcerated (Harlow, 1998). Thus, these women are disproportionately at risk for sexually transmitted infections in addition to having mental health issues due to violent sexual trauma (Freudenberg, 2001). These personal and social problems are compounded by the fact that nearly 40% of incarcerated women find themselves homeless upon release (Richie, 2001).

One observes further iatrogenic effects in the cases of education and family ecology. Congress passed legislation in 1994 prohibiting prison inmates from receiving Pell grants in pursuing higher education, thus limiting prisoners' access to professional skill sets that ease re-entry to society (Fellner & Mauer, 1998). Thus, for groups such as that of the African American community, where males have a 29% chance of spending time in prison at one point in their lives, disparate contact with the justice system affects this group's very ability to gain upward social mobility through education.

The African American community also has a considerable imbalance in the adult male-female ratio (Courtwright,

1996). While boys outnumber girls at birth, by the age of 40 this cohort has 86 males for every 100 females. With the rate at which the remaining men are incarcerated or under criminal justice supervision, the number of eligible males for marriage is chronically low. Furthermore, felons are largely disenfranchised; consequently a group like African Americans has nearly 13% of its adult male population ineligible to vote (Fellner & Mauer, 1998), weakening the very ability of these communities to influence the policies that set the parameters of their lives.

Communities experiencing disparate contact with the criminal legal system have fewer informal mechanisms of crime control. Joblessness, poverty, and a highly transitory population are believed to be associated with poor neighborhood cohesion. As a result, the ability of authority figures to transmit values in tune with the larger society expectations to children is highly threatened. These communities develop a somewhat antagonistic relationship with law enforcement, where police are perceived as equally as threatening as criminals. This leads to lower rates of participation in crime reduction programs or even basic reporting of crimes to the police.

The very basis of democracy is the presumption of equality; therefore it is generally agreed that it is wrong in principle to treat equals differently unless there is some reasonable circumstance warranting such treatment (May & Sharratt, 1994). In considering most crime theory, we see that crime reduction methods based on identifying differences saddles the line between violating our principles and addressing a social problem in the most practical way, despite iatrogenic effects. So in resorting to reducing crime through the identification of differences, we must carefully consider whether there is some reasonable circumstance warranting such treatment.

Juvenile Justice as Differential Treatment

The very basis of the argument for the necessity of juvenile justice is the presumption that youth require differential treatment from their adult counterparts due to differences in their etiological pathways to crime (Shepard, 1998). Adolescence is second only to infancy with regards to rate of change and growth, where identity formation and social integration primarily occur (Erikson, 1963). Both of these developmental stages are highly susceptible to the influence of peer, educational, familial and social factors.

Adolescents typically are moving toward autonomy from parental figures, whilst undergoing rapid neurobiological changes that affect emotional and abstract reasoning. Developmental neurobiologists have gone so far as to use magnetic resonance imaging of the brain to show how youth and adults differ in the way they think, with adolescents using their limbic system whereas adults use their frontal cortex in decision-making of an emotional nature (Siegel, 1999). The limbic system is responsible for instinctive reactions whereas the frontal cortex is usually relied upon for complex reasoning and thought.

Developmental psychologists observe that youth involved with the justice system typically come from homes and communities that are marked by harmful environmental factors and thus they lack the consistent and reliable care-giving necessary for their moral development (Bandura & McDonald, 1963). Youth presumably are most given to undesirable social behavior that violates codified laws because of some lack of supervision or failure on the part of the adult and communal actors in the child's life; therefore warranting rehabilitation of the youth as opposed to retributive justice which is traditionally considered the priority of adult systems.

One notes that without this perspective and in treating juveniles like adults, society used methods that would be considered senseless and draconian today. Prior to the adoption of this perspective, at least ten children under 14 years of age were executed for their crimes (Streib, 1987). Juvenile courts were thus born out of the "parens partriae" doctrine, which implied the courts have an obligation to take charge of and rehabilitate delinquent juveniles (Anderson, 1988; Scott & Grisso, 1997).

The Source of Disparity

Between system contributions to disparities and intrinsic differences between groups like that of juveniles and adults, there is considerable room for debate as to why these differences exist. Moreover, in acknowledging identifiable differences and acting to reduce crime through special treatment of individuals possessing those characteristics, society has impetus to consider the advantages and disadvantages of just such a practice.

Higher education, for instance has experienced a chronic under-representation of minorities and has long since incorporated standardized tests into their admissions, in an effort to "objectively" identify students most likely to succeed in college. The median correlation of the SAT I with GPA for Blacks reported in a 2001

comprehensive review was .33 for African Americans and was .44 for Whites (Young, 2001). Moreover studies consistently demonstrated that the SAT I under-predicts college GPA performance for women as compared to men. As standardized tests are so important in American society because of their role in college admissions; essentially deciding who and who does not get access to our educational resources the bedrock of social mobility in this country, the persistence of differential predictive validity in standardized testing is problematic to say the least. As a result, the education field is constantly in a state of debate and legal challenge as to the merits of its classification and prediction system (Zwick, 2004). The parallel example of the SAT I that is taken by 40% of graduating US high school students (Young, 2001) is quite salient in that it reflects how measures can become so entrenched in practice, the field is almost wedded to them. This is despite empirical evidence pointing to fundamental flaws in their predictive validity with girls and some minority groups (Young, 2001).

Attribution Theory (Heider, 1958) could explain the affinity for most of the paradigmatic traditions cited in these debates. Self-serving causal attributions or fundamental attribution errors are at play casting the individual or the environment wholly responsible for the

disparate outcome. Criminal justice theorists have long since developed any number of theories to explain away specific differences in differential offending/system contact rates that focus either on the individual or the environment. For instance, Strain Theory suggests disadvantaged groups feel some conflict between aspirations of success and their expectations of achievement given societal and structural systems (Agnew, 1992). When one turns her or his attention to biological deficit theories, one finds they typically imply criminal behavior is by its nature pathological. As a result, disparities in incarceration rates reflect endemic maladaptive disorders in those groups (Loeber & LeBlanc, 1990). Simply put, some groups merely have a disproportionate number of individuals that are genetically or otherwise biologically predisposed to crime. On the external attribution side, social disorganization theories attribute delinquency to spatial and ecological factors that encourage crime in some contexts and reduce it in others (Sampson, Raudenbush & Earls, 1997).

While these debates are common and are certainly conflict-laden, clearly some degree of support and evidence exists for each one of these perspectives. Thus, the greater axiomatic conclusion that we can draw from multiple

competing theories concurrently explaining a single social condition is that the causal pathways are in fact quite complex and interactive. The framework then that is best suited to account for differential offending and, in turn, our differential responses to undesirable social behavior of juveniles must account for behavior in context and thus must be comprehensive.

If we review the logic statement to this point we are essentially saying the following: Crime warrants a formal response from society. Criminals are different than law-abiding citizens; ergo the justice system is heavily dependent upon identifiable differences to prosecute criminals for their deviant behavior. Moreover, the accuracy of that system depends on the comprehensiveness of the theoretical framework the system response is derived from. Hence, as much of the theoretical basis for crime reduction is based either on external or internal causal attributions of crime, unjust disparities that disadvantage certain social groups arise. If the juvenile justice system incorporates risk assessment as a mechanism for objectively and correctly identifying youth most likely to offend that model's success not only hinges on the presumption of differences amongst individuals. The reliability or fallacy of that presumption also rests on how comprehensive the

theoretical framework one works from in parceling out relationships between crime's cause and effect.

A Comprehensive Theoretical Framework

An alternative and more comprehensive model to use in validating and improving upon risk measures is Bronfenbrenner's Bioecological Systems theory (1979). Bronfenbrenner (1979) offered that the development of a child occurs within a larger relationship context between its biology and the various dimensions of its environment. Thus, as the child matures the interaction between the child's own biological make-up, his immediate family/communal environment and societal background collectively impact the child's development.

When the child experiences a negative outcome or engages in undesirable social behavior like crime, the causal source lays not only in the realm of their environment or the individual (cognitive, genetic and physical attributes), but rather the cumulative interaction between the two, which are both multi-leveled.

Within the immediate environment of the child (separate of their cognitive, genetic and physical characteristics) rests the microsystem (Bronfenbrenner, 1979). This system includes the family, school, neighborhood and childcare structures. As these factors are

closest to the child, they have the greatest impact on the child's well-being. It should be noted that the effect is bidirectional; so for example, just as the child is affected by their school institution, the school institution is affected by the child. This creates multiple feedback loops that reinforce and initiate behaviors throughout the maturation process. Between each of these structures comprising the microsystem lays a connecting tissue referred to as the mesosystem (Bronfenbrenner, 1979). Just as the family is connected to child, the family is also connected to the childcare and school and neighborhood.

Adjacent to the microsystem and mesosystem is the exosystem which is comprised of the larger social systems in which the child is not directly involved (Bronfenbrenner, 1979). For instance, a child may not have any interaction with a school board but their teacher does, and whatever effect the school board has on their teacher is augmented and filtered down to the child through the child and teacher's interaction in the classroom.

The outermost layer is referred to as the macrosystem (Bronfenbrenner, 1979). The macrosystem is comprised of values, customs and laws. These principles give shape to the interactions at the other levels. For instance, a

society that uses work hours and the demand of the skill to determine the amount of compensation for labor will inevitably require that workers with skills in less demand work long hours to support their families. This will affect children by limiting the amount of time the parent spends with the child.

This ecological framework has implications for society's response to undesirable social behavior. At the macrosystem level our values and principles affect the basic determination of which undesirable behaviors are so egregious they warrant formal sanctions. They affect our attitudes, which in turn affect our implementation and consistency of application of these laws in the microsystems of youths' lives. These effects ripple into the exosystem of principle actors of the juvenile's life in many ways. For instance, attitudes in a community that reflect zero-tolerance for drug use exert considerable pressure on law enforcement officers to satisfy their expectation of having a drug-free community. This in turn affects the officer's relationship with the youth directly. It may even affect the youth indirectly through the parent or school via the mesosystem in a positive or negative manner depending on the internal environment of the child and their respective affinity for drug involvement.

When one traces these causal pathways from criminogenic factors to delinquent acts to justice response, the inordinate number of possible mediators and moderators becomes obvious. Moreover, one realizes that an interactive system such as this is primed for any number of differential outcomes and differential treatments of individuals. The system is then faced with a question of how to accommodate the complexity of the system yet manage the group with the single relevant outcome of offending. In criminal justice, when powerful individuals and groups have sole discretion in deciding what those relationships are and in turn what those relationships mean with regards to differences between groups and individuals in behavior, the system becomes highly reflective of the beliefs and prejudices most prevalent in that population. Therein lie the cost of change models that are informal, of limited causal attribution scope and unchecked for inconsistencies in implementation.

From this premise one arrives at the notion of creating additive ecological models that account for each of these factors at the various levels and then determines critical points. When certain thresholds are met the likelihood of offending becomes so great, specific action from the justice system addressing those needs is

warranted. This is the theoretical framework behind risk assessment. The problem then arises when we consider that typical risk assessment goes no further than assessing the micro and mesosystem. Rarely does this expand to include the macrosystem. Nor does the common practice of simply adding the number of risks that exist in a youth's life account for the interaction between those risks and how particular combinations of criminogenic factors either mitigate or exacerbate delinquency. So in considering risk assessment as a solution to disparate contact, we must also recognize that this solvency may be limited by the comprehensiveness of risk measures and the lack of validation with diverse populations.

In the following sections, the rationale for using risk assessments is reviewed. In addition, previous research validating contemporary risk assessment with regards to predictive accuracy and dimensional identity is reviewed. The remainder of this review will demonstrate the necessity of validating the YLS/CMI with a diverse population, using both additive models and risk combinations to predict the likelihood of offense.

Chapter 2

Study Rationale

Given that over 2 million young offenders are arrested annually, costing $14.4 billion for law enforcement, courts, detention, residential placement and block-grants (National Center on Addiction and Substance Abuse, 2004), there is considerable impetus to identify repeat offenders early in the justice process. Recidivists are of particular concern because as few as 6% of offenders can commit the majority of offenses through recidivism (Snyder & Sickmund, 2006). Moreover, treating each juvenile as though they have an equal likelihood of returning to the system without intervention leads to wasted resources and inadequately serviced youth in need (Howell, 2003). There is great need for matching interventions with the need set of young offenders, consequently, dependable tools for identifying re-offenders and assessing areas of criminogenic risk are highly desired by systems (Howell, 2003).

The process for sorting offenders into groups based on criminogenic risk is vulnerable to enabling wide disparities in system contact based on group membership status for the previously mentioned reasons. Risk assessments and decisions based on group membership status

can lead to over-predictions and under-predictions of future criminal activity. For example, since the 1960s at least two-thirds of studies have found some degree of disproportionate racial minority contact with law enforcement that was attributable to over-reliance on subjective human judgments at the intake and arrest level, according to a Pope and Feyerham's meta-analysis (1992).

Furthermore, girls are increasingly becoming vulnerable to a "net-widening" effect due to a difference in their criminogenic developmental pathways (Chesney-Lind & Sheldon, 2004). Girls are unique in that their first contact with the system is often due to relatively minor offenses, while their male counterparts exhibit a far greater range of severity in offending (Chesney-Lind & Sheldon, 2004). Programming often mixes girls regardless of stage in criminogenic development. Net-widening for girls is exacerbated by their status offenses being formally prosecuted at a greater rate than their male counterparts due to prevailing opinions about their risk for truancy, incorrigibility and running-away. As such, this initial contact invites greater attention from the legal system, and this attention facilitates contact with higher risk offenders in the system and essentially exacerbates their criminogenic risk through peer group association.

There is also some evidence that commonly recognized risk factors for boys do not have the same degree of association with offending for girls, so girls that exhibit similar risks as boys are not necessarily as likely to re-offend. This can lead to over-programming of girls in the justice system when they exhibit risks similar to that of their male counterparts. Risk assessments can as result over-predict delinquency for girls if not properly validated with this group.

These factors coupled with the observation that minor and low risk offenders are made worse by extended and intense contact with the justice system have driven the resurgence in popularity of actuarial risk assessment measures (Andrews, Kiessling, Robinson, & Mickus, 1986). Indeed, in the 2003 Comprehensive Framework for Juvenile Justice, the US Office of Juvenile Justice and Delinquency Prevention offered actuarial prediction instruments as tools in creating a system of equitable and accurate disposition decision-making (Howell, 2003) by standardizing assessment practices and using prediction methods that are similar across jurisdictions and populations (Grove & Meehl, 1996).

As the use of risk assessments has become widespread, the success of juvenile justice has become dependent upon

the validity, accuracy and dimensional identity of the prediction instruments used in this cornerstone of the justice process. Risk assessment as it is referred to here is the prediction of offending and classification of juveniles based on their likelihood to offend (Bonta, 1996; see also Gottfredson & Moriarty, 2006).

The predictive validity of risk assessment measures thus refers to the extent to which these instruments predict delinquency, which is reflected in the significance and strength of relationship between risk scores and delinquency outcomes (Gottfredson & Moriarty, 2006). In turn, accuracy is the rate at which risk assessments correctly identify non-offenders and offenders. Dimensional identity is said to exist when the same relationship between delinquency and risk assessments is true for all subpopulations upon which the measure has been administered (Von Eye & Bergman, 2003).

In systems that use risk assessments in their decision-making, the success of the system at reducing delinquent offending is directly related to the performance of their measures in the aforementioned regards to validity, accuracy and dimensional identity. Wiebush and researchers (1999) found that 39 states reported using risk assessments and classification instruments with young

offenders. The use of actuarial risk assessments is thus quite common across many jurisdictions. However, there has been little attempt to systematically review the success of these measures over the past decade in predicting delinquency.

In so much as risk assessments are valid predictors of recidivism with diverse populations, risk assessments by their very nature could reduce the systems contribution to disparities in adjudication rates through squeezing out the biases found at the early stages of the justice process. Theoretically these instruments should do a better job than informal assessments at identifying youth likely to offend repeatedly, regardless of race or gender, thus members of marginal groups that are low risk would conceivably be eligible for the same programming as their counterparts, which in many cases includes being refracted from the system. For these groups, this in turn results in "narrowing of the net", the reduction of disproportionate marginal group contact with law enforcement.

However, little study has been conducted on the validity of risk assessments with girls and African Americans, so this posited curative relationship between risk assessment and the reduction of disparities is presumptive at best. Moreover, it is likely that the risk-

recidivism statistical relationship is naturally different for members of marginal groups.

If conflict theory is correct, members of these groups would be at greater risk of having their criminal activity exposed due to disparate levels of policing, surveillance, and attention from the justice system in their communities. As even multi-domain risk assessments do not include policing levels, community surveillance, and justice system attention as predictive risk factors, the instruments should under-predict offending for African Americans. This is especially disconcerting when one considers there is evidence that the very factors used in some risk assessments have been shown to have little to no relationship with delinquency for girls leading to over-predictions of delinquency for this group and further entrenching girls in the justice system when case-management decisions are based on those assessments.

Consequently, there is need for validation of instruments used by systems to identify youth most likely to recidivate. This study was an attempt to validate the YLS/CMI, which is used in the local jurisdiction with a wide cross-section of young offenders including African-Americans and girls. In choosing a predictive instrument, a comparative review of the performances of various popular

risk assessment instruments was necessary. In the following section, a synthesis of this literature is presented. The purpose then of the following review is to summarize and critique the extent to which delinquency risk measures have been validated with attention to the usefulness of these relationships with regards to their accuracy and dimensional identity. Moreover, attention is given to the specific performance of the YLS/CMI compared to other instruments.

Chapter 3

Risk Assessment

Risk assessment rests on the premise that a small number of offenders commit the majority of offenses through repeat offending, which is well evidenced in the literature (Wolfgang, Figlio, & Sellin, 1972; Moffit, 1993; see also Snyder & Sickmund, 2006). For instance, Schumacher and Kurz-Gwen (2000) demonstrated in a longitudinal study with 70,000 youth that as few as 8 percent of offenders commit more than half of offenses in a given year. In order for risk assessment to work, this small group of offenders must share characteristics in common that differentiate them from their fellow offenders.

Risk Factors

The literature is thus replete with research indicating that characteristics of an individual's life on multiple levels mark those offenders most likely to commit the bulk of offenses (Cottle, Lee, & Heilbrun, 2001). Moffit (1993) articulated these differences as risk factors, which set offenders on either of two delinquency paths: adolescent limited and life course persistent offending. In the case of the former, the majority of offenders limited their delinquency to adolescence, while

for the latter; a minority of young offenders (6%) continued to offend well into adulthood. Moffit noted that life-course persistent offenders were characterized by committing offenses at young ages and poor prenatal care.

Since then a number of risk factors like family structure (Myner, et al., 1998), learning disabilities (Duncan, et al., 1995), gender (Dembo, et al., 1998), offense type (Archwamety & Katsiyannis, 1998) etc, have all been associated with delinquency. In turn, a Cottles, Lee and Heilbrun's (2001) meta-analysis of 23 studies, representing 15,265 youth revealed that age at first offense, family problems, peer groups and family history of criminality were the strongest single predictors of recidivism.

It should be noted, however, that the use of single risk factors to predict the outcomes of offenders is replete with high false positive and false negative classification rates thus making single dimensioned assessments inadequate in the case management process (Risler, Sutphen, & Shields, 2000). For example, we can know in general that individuals of certain racial groups or of certain gender or of a certain age raised in a distressed community are more likely to commit offenses than their peers without those risk factors. We also can

know individuals that do share those risks in common can have very different outcomes, some desirable while others are undesirable. Thus, in classifying youth solely on the basis of these risks, the usefulness of our predictions will be hampered by high false positive and false negative rates.

Risk Factors for Girls

Moreover, using single risk factors is also problematic because there is some indication that certain risk factors are more salient with some groups than others. Specifically with regards to gender, criminological theory has largely focused on males and been generalized to females (Belknap, 1996). Simound and Andrews's (1994) meta-analysis went as far to suggest that predictors for males and females are largely the same, with criminal history, antisocial behaviors, and delinquent peers each serving as strong predictors of delinquency for both genders. Hubbard and Pratt (2002), who also conducted a meta-analysis, criticized Simound and Andrews's conclusions for its lack of focus upon the maladjustment of young female offenders. Hubbard and Pratt (2002) asserted that there should have been special emphasis on sexual abuse and maltreatment of girls in childhood because of the disparate rate of such abuse against females.

Other studies have suggested that mood disorders (Mcreynolds, 2005), neighborhood disadvantage (Mcreynolds, 2005), and truancy (Hishinuma et. al., 2005) as risk factors are moderated by gender. Hubbard and Pratt's (2002) meta-analysis included 97 effect size estimates representing 5,981 cases, and revealed that while commonly mentioned risk factors like peer deviancy, anti-social behavior and prior history are strong predictors; childhood maltreatment and school relationships are both strong predictors of delinquency for females. Therefore in using single factors that are highly correlated with delinquency but not necessarily accurate in individually predicting offense given race, gender and other moderating factors, systems waste resources and perpetuate unjust biases against entire classes of people.

Risk Factors for African Americans

While studies have indicated that African Americans share many of the same risks for delinquency as White youth (i.e. family, substance abuse, education, personality, offense history, etc.), researchers have noted the relationship between those risks and delinquency are not the same for the two groups (Myers, Taylor, Alvy, Arrington, & Richardson, 1992). Certain criminogenic risks are present to a greater degree in African American

communities than others and several factors moderate the relationship between common risk factors and delinquency for African Americans. For example, African Americans are more likely to experience poverty, poor schooling, single-parent households, and are more likely to witness violence in their communities. Often the prevalence of these risk factors can result in high levels of adaptability and resiliency, confounding typical relationships between those factors and offending (Kliewer & Kung, 1998). Studies of coping strategies and delinquency have shown that emotion-focused coping strategies are associated with negative behavioral outcomes, whereas active-coping is associated with positive behavioral outcomes (Compas et al., 1998). These studies have also shown the difference in prevalence of these strategies in the two groups is due to differences in available resources.

Furthermore, some research has indicated that African American youth are particularly vulnerable to certain risk factors due to stress-induced factors related to their status as African Americans (i.e., class oppression, low self-efficacy, and barriers to goal achievement) (McGee, 2003). Moreover, some studies have indicated the single factor of threat perception in response to everyday situations has a substantial impact on delinquency risk

(Kliewer, Fearnow, & Walton, 1998). African American young offenders in vulnerable communities often exhibit higher levels of "threat awareness" resulting in a greater likelihood of using coping strategies that may or may not be socially desirable (Kliewer, Fearnow, & Walton, 1998). Failure to account for factors related to resiliency can lead to over-predicting offense rates in groups of African Americans that have highly adaptive members. Moreover, failure to account for structural factors and community conditions that may exacerbate delinquency risk can lead to under-predictions of risk.

Ecological Perspective and Risk Assessment

The ecological perspective may be more desirable as a basis for developing risk assessment measures. As articulated by Rappaport, the "ecological viewpoint should be regarded as an orientation emphasizing relationships among persons and their social and physical environment. Conceptually the term implies that there are neither inadequate persons nor inadequate environments, but rather that the fit between persons and environments may be in relative accord or discord (p. 2, Rappaport, 1977)".

The failure of oversimplified risk models that do not account for person-environment fit reflects the theoretical work of Patterson (1989) and Elliot (1985) on the

developmental pathways of juvenile criminality, which suggests that the undesirable behavior of youth is both complicated and multifaceted. In some cases delinquency is a product of psychological pathology, in other cases a result of detrimental environmental factors or in still other cases some combination of both, intrinsically implying that juvenile offenders are a rather heterogeneous population (Cullen & Agnew, 2006).Thus while a myriad of risk factors have been individually correlated with delinquency, it is the combination and cumulative effect of risk factors that best predict delinquency (Schwalbe, Fraser, Day & Cooley, 2006).

Psychometricians have consequently attempted to develop instruments that account for cumulative risk across several domains in additive models where higher scores reflect higher likelihoods of delinquency, classifying youth as high, moderate or low risk with corresponding rates of re-offense associated with each risk category.

While the widespread use of risk assessments reflects the resurgence in their popularity, the actual first use of risk assessment dates back to as early as 1928 with the Burgess instrument which was used by the Illinois parole board in parole decisions (Burgess, 1928). The 21 item Burgess Instrument outperformed prison psychiatrists in

predicting parole failures for 3,000 Illinois offenders (Burgess, 1928). Its early use demonstrated the intended purpose of structured risk assessments in juvenile justice; that being to predict offending and in doing so risk assessment aides in case management and policy decisions through objectively and consistently assessing juveniles given the complexity of juvenile criminality.

Re-Emergence of Dynamic Risk Assessment Measures

In 1996, Bonta conducted a meta-analysis of correctional classification literature and described the evolution of risk assessment instruments, concluding that there have been at least three generations of risk measures. The *first generation* was comprised of clinicians and practitioners offering their professional opinion about the likelihood of particular delinquency outcomes (Bonta, 1996). While the clinical method is still in use today, the accuracy and validity of predictions made in this manner has been shown to be inadequate (Lowenkamp, Holsinger, & Latessa, 2001; Van Voorhiss & Brown, 1997). This has been demonstrated in the work of Grove and Meehl (1996) who reviewed 136 studies since the 1920s across disciplines in which clinical and actuarial practices were compared. The researchers found that 64 of those studies favored actuarial practices versus eight siding in favor of

clinical/human decision making (Grove & Meehl, 1996). In assessing risk, human judgments result in high false positive and false negative rates for several reasons. Clinicians and practitioners often erroneously ignore base rates of offense, make spurious correlations and are swayed by incorrect causal attributions (Gottfredson & Moriarty, 2006).

The *second generation* of risk assessment was comprised of risk measures using static risk factors like race or age of first crime to predict outcomes (Bonta, 1996). The risk factors used in such measures were theoretically derived and had statistically significant relationships with delinquency. For instance, with regards to the latter aforementioned risk factor, researchers like Moffit offered the supposition that there are life course persistent offenders and adolescence limited offenders, explaining away differences in offense rates between juveniles as being heavily influenced by age at which first offense is committed (Moffit, 1993). This supposition was reflected in offense rates which showed that age of first offense was a strong indicator of future delinquency. In turn other risk factors like race and gender were also strong indicators of future delinquency and were thus included in many risk assessment measures.

The risk measures offered an advantage over the clinical method with regards to accuracy and validity (Bonta, 1996). Thus risk assessment was initially steeped in the individual differences tradition; however the failure to consider the context of these factors across multiple levels in individuals' lives reduced the effectiveness of these instruments in correctly predicting crime outcomes and in guiding case management (Bonta, 1996; Wiebush et al. 1995). As static factors like race and age of first offense are not malleable to change, second generation risk assessment provided little useful information that could be used in intervention efforts.

There was also considerable evidence that certain risk factors included in second generation prediction instruments were more salient with certain groups than others. For instance, the occurrence of hyperactivity is more common in boys than girls (Lahey et al., 1999) and in turn neighborhood risks pertain to a larger portion of Black youth than Whites (Loeber & Farrington, 1999) therefore when using these factors to assess risk, some groups automatically exhibit elevated risk levels on these risk assessments that did not bear out in delinquency outcomes.

Differences in level of association between certain risk factors and delinquency for different social groups pose a threat to the dimensional identity of measures used with diverse populations. Again, dimensional identity refers to the generalizability of validity and accuracy of prediction models across various subgroups of population like race or gender. Specifically in the case of juvenile justice it has become fundamentally clear that ignoring poor dimensional identity and ignoring the person-environment fit in risk assessment has tremendous costs in terms of justice and distribution of resources (Bonta, 2002). This occurs through biases prejudicing the system against members of certain groups or poorly distributing resources to wrongly classified individuals.

The *third generation* of actuarial assessments represented a departure from the individual differences tradition and an attempt to adopt an ecological model with dynamic risk factors. Bronfenbrenner (1979) offered that the undesirable social behavior of youth was a product of poor fit between various systems on multiple levels in which an individual fits. The micro-level was characterized as layer of interaction between the youth, family and immediate environment, while the mesosystem was characterized as connecting layer between the family,

social services and school. Finally, the macro-system referred to cultural expectations, laws and views which produce a cascade effect through the various layers of the youth's external and internal environment that ultimately impact the child. As this pertains to risk assessment, the failure or success of an instrument at predicting the undesirable social behavior of juveniles is a direct reflection on its ability to capture information on the person in larger contexts of his or her environment. As such, instruments of prior generations that overly focused on individual and fixed characteristics of the individual were inadequate and thus were doomed to failure. Third generation actuarial assessments reflected these series of interactions by using the cumulative risks of several risk factors on multiple levels in predicting offense. Consequently the power of risk assessments has become dependent upon the risk factors or risk subscales that they tap into. In a widely cited meta-analysis, Cottle, Lee and Heilbrun's (2001) synthesis of risk factor literature demonstrated that offense history, psychological factors, school behavior, attitudes toward authority, gender factors and peer factors are each strongly associated with re-arrest, probation violation and commitment. In turn many

risk instruments tap into these domains in predicting delinquency.

The third generation of actuarial assessments is comprised of measures using dynamic and static risk factors to determine the likelihood of the youth coming in further contact with the system. Dynamic risk factors are risk factors that are malleable to change efforts of clinicians and practitioners, whereas static factors are fixed and rigid characteristics of a youth or their environment that cannot be changed. This cumulative multi-domain approach is reflective of the axiom spoken to earlier: the more risks a juvenile experiences in multiple areas of their lives the more likely that juvenile will engage in some form of delinquency in the future. For these instruments to be effective, Bonta strongly advocated periodically validating these measures on populations of interest to system workers. Some researchers have recommended that measures be validated every two years. With at least 39 states using risk assessments, adherence to that recommendation would mean that a number of studies have been conducted in the past decade on the validity of risk measures. Unfortunately, since 1996 there have been no systematic reviews of studies validating juvenile risk assessments qualifying as third generation instruments.

Chapter 4

Validity, Accuracy and Dimensional Identity of Risk Assessments

Predictive validity in the case of actuarial risk assessments refers to the extent to which the scale predicts a delinquency outcome (Walters, 2006). In that way an instrument can be validated in a number of ways and in fact have been. For instance, a study could show that there is a significant correlation between risk score and recidivism. Another study could show that recidivists and non-recidivists have significantly different scores on the instrument. Yet still another study could show that there are significant differences between juveniles that have been classified as "low or high risk" in recidivism rates. Researchers then have wide latitude in determining the validity of their risk assessment measures, making comparisons extremely difficult.

However, several researchers have made recommendations as to how predictive validity studies be conducted and what validity coefficients should be reported. Nunnally (1978) noted that the accuracy of predictive measures are always higher with samples with which they are created than with independent samples; therefore validity studies should

utilize independent samples. DeComo, Wagner, and Baird (1994) recommended that instruments differentiate between high risk groups and low risk groups by at least 4 to 1 or 30% on predicted delinquency outcomes. This means that the recidivism rate for juveniles classified as high risk should be four times that of juveniles classified as low risk using a prediction instrument.

Much of the risk prediction literature has focused on the actual statistics and validity coefficients used in describing the relationship between risk assessments and delinquency outcomes. In a widely reported criticism of validation practices solely using correlations or association statistics to validate measures, Loeber and Dishion (1983) argued that the validity of predictive instruments depends on their ability to minimize false positive and false negative error rates with regards to predicting re-offense. The base-rate of re-offense represents the odds of correctly guessing that any given offender will re-offend. It is possible for a measure's risk scores to be significantly correlated with a delinquency outcome yet incorrectly identify the vast majority of its offenders as recidivists or non-recidivists. Therefore, if any instrument does not correctly identify re-offenders at a greater rate than

chance, such instruments offer little practical advantage with regards to case management in spite of possible significant correlations between risk and offense outcomes.

Since 1983 a number of validity coefficients which account for base-rates have been offered as superior alternatives to correlations such as Area Under the Curve (AUCs), and Relative Improvement Over Chance (RIOC) statistics. Given the myriad of validation protocols, this review reported how common each of these methods/techniques was in reported validity studies.

Results of Systematic Review

As reported in Table 1, between 1996 and 2006, 23 studies were published with the majority (61%) presented in the past 5 years. Sample sizes ranged from 35 to 9,692 participants with a mean sample size of 1,498 (SD=2,748). Of the studies reporting age (20 of 23), the average age of participants was 14.5. Of the studies reporting gender (18 of 23), sample compositions for males ranged from 54% to 100%, with a mean of 70% (SD=11%). Of the studies reporting racial demographics (20 of 23), 12% to 100% of samples were White youth; in the typical study 45% (SD=25%) were Whites.

Table 1. Description of validation studies with risk assessment instruments.

Author	N	Age	Boys	White	FU	Instrument	a	b	c	d	e	f	g	h	i	Design
NCCD, 2001	954	15	72%	18%	12	ACPRA	X	X		X				X	X	M
Hodges, 2000	1460	12	64%	N	6	CAFAS	X	X	X	X		X	X			C
Quist, 2001	35	15	54%	23%	12	CAFAS	X	X	X	X		X	X			C
Frank, 1999	804	13	56%	76%	6	FISCA	X	X		X		X	X	X		M
Risler, 2000	181	14	65%	20%	48	FORAI	X	X		X				X	X	C
Funk, 1999	500	14	77%	13%	36	Funk Model	X	X		X				X	X	C
Huff, 1999	1173	15	78%	85%	8	IRA	X	X		X				X	X	M
Grinberg, 2005	404	13	59%	40%	N	LCQRAI	X	X	X				X		X	M
Nowicka, 2005	154	17	100%	100%	46	LSI	X	X		X	X	X	X	X	X	C
Schwalbe, 2004	464	14	75%	50%	12	NCAR	X	X		X				X	X	M
Schwalbe, 2006	9534	14	77%	43%	24	NCAR	X	X		X				X	X	C
Krysik, 2002	4754	13	64%	55%	12	NCCDRA		X		X				X		M
Sharkey, 2003	159	13	67%	30%	18	OCPDLRA	X	X		X				X	X	C
Smith, 1999	298	16	N	35%	N	PRI	X	X		X		X	X	X		C
Jimmerson, 2004	423	15	66%	40%	6	SBARA	X	X	X	X			X	X	X	M
Turner, 2005	1165	16	74%	19%	12	SDRRC	X	X		X		X	X	X	X	M
WSIPP, 2004	9692	N	N	N	18	WSJCA	X	X		X		X	X	X	X	C
Jung, 1999	250	14	66%	49%	6	WSJCA	X	X		X		X	X	X	X	C

(Table 1. continued)

Schmidt, 2005	107	15	63%	71%	60	YLS/CMI	X	X		X	X	X	X	X	X	C
Marczyk, 2003	1995	16	N/A	12%	12	YLS/CMI	X	X		X	X	X	X	X	X	C
McKinnon, 2004	102	N	N	N/A	42	YLS/CMI	X	X		X	X	X	X	X	X	C
Catchpole, 2003	74	N	N	55%	12	YLS/CMI	X	X		X	X	X	X	X	X	C
Flores, 2004	167	15	79%	70%	12	YLS/CMI	X	X		X	X	X	X	X	X	M
Average	1498	14.5	70%	45%	16	YLS Most Common	94%	100%	19%	94%	13%	44%	56%	88%	81%	
Range	35 to 9692	12 to 17	56 to 100%	12 to 100%	6 to 60											

Note: N = Not available

Column Labels: F.U. = Follow-up Length in Months

a = Education; b = Family; c = Community; d = Substance; e = Leisure; f = Attitude; g = Personality;

h = Offense History; i = Peer

*Design Column Entries = C = Correlational Design; M = Means Difference Testing

Table 2. Results of validation studies.

Author	Scale	Criterion	Dependent Variable Type	Validity Coefficient	Design
NCCD, 2001	Risk Level	Recidivism	Arrest	χ^2=90, p<.01	M
Hodges, 2000	Risk Score	Delinquency	S.R. Contact	β=.1, p<.01	C
Quist, 2001	Risk Level	Recidivism	Arrest	r=.41, p<.01	C
Frank, 1999	Risk	Recidivism	Detention	χ^2=17, p<.01	M

(Table 2. Continued)

Risler, 2000	Risk Score	Recidivism	Arrest	RIOC=.42	C
Funk, 1999	Risk Score	Recidivism	Arrest	R^2=.21	C
Huff, 1999	Risk Level	Recidivism	Arrest	χ^2=16, p<.01	M
Grinberg, 2005	Detention Status	Risk Score	Risk Score	F=49.1, p<.01	M
Nowicka, 2005	Risk Score	Recidivism	Conviction	AUC=.77	C
Schwalbe, 2004	Risk Score	Recidivism	Arrest lag	χ^2=15.6, p<.01	M
Schwalbe, 2006	Risk Score	Recidivism	Arrest lag	HR=1.08, p<.01	C
Krysik, 2002	Risk Level	Recidivism	Arrest	PPM=3.2 (<4 cut-off)	M
Sharkey, 2003	Risk Score	Recidivism	Arrest	AUC=.60	C
Smith, 1999	Risk Score	Recidivism	Arrest	AUC=.78	C
Jimmerson, 2004	Risk Score	Recidivism	Arrest	X^2 = 82.9 p<.01	M
Turner, 2005	Risk Level	Recidivism	Arrest	X^2 =88.3, p<.01	M
WSIPP, 2004	Risk Score	Recidivism	Arrest	AUC=.64	C
Jung, 1999	Risk Score	Recidivism	Arrest	AUC=.64	C
Schmidt, 2005	Risk Score	Recidivism	Arrest	AUC=.61	C
Marczyk, 2003	Risk Score	Recidivism	Arrest	R^2=.01, non-sig(AUC=53%)	C
McKinnon,2004	Risk Score	Recidivism	Arrest	AUC=.57	C

(Table 2 Continued)

Catchpole, 2003	Risk Score	Recidivism	Arrest	AUC=.78	C
Flores, 2004	Risk Level	Recidivism	Arrest	X^2=74.6, p<.01	M

Table 3. Dimensional identity of risk assessment instruments.

Author	Method	Sample Size		Mean Score		Recidivism Rate: Low		Recidivism Rate: High		Validity Coefficient	
		Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
NCCD, 2001	M	270	684	9	9	N/A	N/A	N/A	N/A	N/A	N/A
Hodges, 2000	C	518	942	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Funk, 1999	C	112	388	N/A	N/A	N/A	N/A	N/A	N/A	R^2=.20	R^2=.17
Schmidt, 2005	C	40	67	19.1	15.7	0	0.18	0.4	0.57	r=.14, ns	r=.25, p<.05
Jung, 1999	C	90	173	11.9	11.1	N/A	N/A	N/A	N/A	F= .11, ns.	
Sharkey, 2003	C	53	106	15.5	14.6	0.41	0.4	0.58	0.54	ROC=.52	ROC=.65
Huff, 1996	M	259	914	N/A	N/A	0.1	0.14	0.38	0.45	RIOC=.26	RIOC=.35
Schwalbe, 2004	M	116	348	8.1	8.6	N/A	N/A	N/A	N/A	ns.	
Turner, 2005	M	268	768	22.1	18.9	0.05	0.1	0.22	0.38	r=.21, p<.05	r=.24, p<.05
Jimmerson, 2004	C	149	274	N/A	N/A	N/A	N/A	N/A	N/A	R^2=1.00	R^2=.42
WSIPP, 2004	M	N/A	N/A	N/A	N/A	0.06	0.13	0.18	0.36	N/A	
Flores, 2004	C	324	1180	20.9	19.6	0.12	0.1	0.44	0.58	r=.32, p<.05	r=.31, p<.05

Results of Systematic Review

As reported in Table 1, between 1996 and 2006, 23 studies were published with the majority (61%) presented in the past 5 years. Sample sizes ranged from 35 to 9,692 participants with a mean sample size of 1,498 (SD=2,748). Of the studies reporting age (20 of 23), the average age of participants was 14.5. Of the studies reporting gender (18 of 23), sample compositions for males ranged from 54% to 100%, with a mean of 70% (SD=11%). Of the studies reporting racial demographics (20 of 23), 12% to 100% of samples were White youth; in the typical study 45% (SD=25%) were Whites.

As demonstrated in Table 1, 16 separate risk assessment measures were used in these studies. The Youth Level of Service/Case Management Inventory (YLS/CMI) was the most widely used with 26% of studies conducted with this instrument. The Child and Adolescent Functional Assessment Scale (CAFAS) and North Carolina Assessment of Risk Instrument (NCARI) were each used twice. Twenty-nine percent of the instruments had fewer than 10 items, with a single measure comprised of 60 items. Each of the studies used multi-domain sampling. It is important to note that no studies accounted for surveillance levels or macro-system variables outlined by Bronfenbrenner (1979). Table 1 details the overlap of domains used in the study's

instruments. Education, family, substance abuse and offense history were assessed in nearly all instruments. The community and leisure time domains were only found on a small percentage of risk instruments. The Life Challenges Questionnaire Risk Assessment Instrument and Alamedo County Placement Risk Assessments were the only instruments that used the self-report of the juvenile rather than multiple sources.

Validity and Accuracy

As evidenced in Table 1, researchers used two primary designs to validate the instruments. Either the study used a correlation design (n=14) or a means-difference design. Means-difference designs were between risk level groups with membership in comparison groups determined by risk assessment scores of study participants (n=9). In the correlational studies, researchers examined the level of association between risk score and delinquency outcome. In the means differences designs, researchers checked for significant differences between groups of youth of different risk levels.

The correlation studies varied in sample size from 35 to 9,672. The follow-up periods ranged in length from 6 months to 60 months with the typical study following the youth for 19 months subsequent of the initial assessment.

For example, Quist and Matshazi (2000) found a significant correlation ($r=.41$, $p<.05$) between risk score on the CAFAS and recidivist arrest for 35 juveniles from a youth rehabilitation group home that were observed over a 12 month period post-assessment. It should be noted that studies merely reporting correlations between scores and delinquency outcomes have been criticized as being insensitive to base rates and provide little information as to the false positive and false negative predictive accuracy of the instrument (Loeber & Dishion, 1983). While Quist and Matshazi simply reported the correlation between recidivist arrest and risk score in validating the CAFAS, this relationship was reported in a number of ways in other studies that failed to provide an idea of the correct classification rate of the validated instrument. For instance, Hodges and Kim (2000) were unique in that they constructed a logistic regression model for the relationship between CAFAS risk score and contact with law in a study with 1,460 youth that were observed over a six month period. Controlling for background variables, they found that the total CAFAS score increased the log-odds of contact with the law by 1% for every point a youth scored on the CAFAS. While the CAFAS was concluded to be a valid

predictor of delinquency, there was no report of how accurate the instrument was in making this prediction.

Each of the studies with correlation designs used risk score as a predictor of a delinquency outcome, thus reporting the magnitude and direction of relationships between relevant variables. However, the specific delinquency outcome varied from study to study. As seen in Table 2, eleven of the fourteen correlational studies used recidivism arrest as the predicted delinquency variable associated with risk score. Contact with law, time to first arrest and conviction were each respectively used once as predicted variables in the remaining correlation studies.

There was some variability in the manner in which these associations were used in concluding that the measure was valid, which is reflected in the reported validity coefficients. For example, a single study by Quist and Matshazi (2000) simply reported the correlation between risk score and arrest and its significance level in validating the CAFAS ($r=.41$, $p<.01$). However, two studies (Funk, 1999; Marczyk et al., 2003) reported the amount of variation in recidivism explained by risk score. The Marczyk study was the single study that reported that its respective risk assessment instrument was not valid given

the fact that 1% of the variation was accounted for by risk score.

The remaining ten correlation studies used regressions models to predict outcomes. The correct classification rate was reported in determining the accuracy and validity of the measure. Seven studies reported Area Under the Curve/Receiver Operator Coefficients. The AUCs ranged in value from 57% to 92%. In general, the higher the AUC the better the classification rate. As such, the difference between the AUC and the base rate of recidivism is the improvement over chance in classifying juveniles correctly with that particular instrument. Thus, the FORAI instrument which was used by Risler and others (2000) is reported as having the greatest improvement over chance in predicting recidivist arrest (AUC = .92; improvement over chance = .42), while Mckinnon (2004) reported that the YLS/CMI had the lowest predictive accuracy (AUC = .57; improvement over chance = .07). It should be noted several studies were conducted in this manner with the YLS/CMI. Schmidt and others (2005) in a longitudinal study over 60 months with 107 youth reported an AUC of .61. In turn, Catchpole and others (2003) reported an AUC of .78 in their study over 12 months with 74 youth. Studies reporting AUCs are further reported in Table 2 and the respective instruments were

each at least 10% better than chance at correctly classifying the juveniles. Of those reporting an AUC, the average was .68.

In the means-difference testing studies (n=9), juveniles were grouped by risk level. Recidivism was then collected during a follow-up period and compared across the risk level groups. For example, Flores, Travis and Latessa (2004) found significant differences in recidivist arrest rates over 12 month period between juveniles classified as low, moderate and high. In this sample of 1,679 youth, high risk youth were 44% more likely to be re-arrested than low risk youth. The remaining studies were similar in nature in that they looked for differences in recidivism rates between youth of different risk levels. In these studies, length of follow-up ranged from 6 months to 12 months. The delinquency outcomes used were detention (n=1), arrest (n=5), time to first arrest (n=1), and conviction (n=1).The sample sizes ranged from 250 to 4,574. As reported in Table 2, each of the instruments validated in this manner had significant differences in delinquency outcomes between low and high risk young offenders. There were two studies which used unique ways to determine the validity of their instruments. Grinberg and others (2005) took a different approach in their study with 404 youth validating the Life

Challenges Questionaire Risk Assessment Instrument. A sample of non-detained youth (n=246) and detained youth (n=87) were compared on the instrument, with significant differences in scores being found as reported in Table 2. The researchers consequently concluded that the instrument was a valid predictor of delinquency. In the Krysik and Lecroy study (2002) with the National Council of Crime Delinquency Risk Assessment assessed on 4,754 youth, they used the DeComo, Wagner and Baird (1994) [PPM] validation method in which high risk and low risk groups were expected to differ in recidivism rates by 30% or by a 4 to 1 ratio. With this instrument the PPM was less than 4, with 76% of high risk youth offending and 24% of low risk youth offending. With the exception of that study, each of the instruments validated were concluded to be valid due to significant differences in delinquency outcomes between low and high risk young offenders (Table 2).

Dimensional Identity

As stated earlier, dimensional identity refers to the consistency in prediction of delinquency outcomes for an instrument across subpopulations in a sample (e.g. race and gender). While theoreticians have repeatedly recommended that validation studies examine the validity of prediction instruments with girls and minorities, few studies have

actually done so. As shown in Table 3, in only 12 of the 23 studies was the differential validity of risk measures by gender explored. Only five of the 23 studies included an exploration of differential validity with ethnic and racial groups. Only a single study (Schwalbe, 2004) reported the correct classification rates separately by gender and ethnicity. Further, few even tested for significant differences in recidivism rates by risk level between groups. Tables 3 & 4 show the respective sample sizes and validity coefficients of these groups where reported. Studies checked for dimensional identity in a number of ways. Some studies checked for differences in recidivism rates between groups that shared risk levels but differed in gender (n=4) or ethnicity (n=2). Others looked for interactions between gender (n=2)/ethnicity (n=1), recidivism and scoring in delinquency outcomes. Still others (gender, n=2; ethnicity, n=1) reported respective correlations for each group between score and delinquency outcomes.

Given that only a few studies examined dimensional identity, it was consistently reported that the higher the risk (by group) of the youth, the higher the recidivism rate regardless of race or gender. When this pattern was further examined, notable differences emerged that merit

note and provide a direction for further study. Three of the nine studies that examined recidivism rates for low risk juveniles separately by race and gender reported differences of at least 9% (Schmidt et al., 2005; Turner et al., 2005 and Flores et al., 2004). However, no significance testing was conducted to determine if these differences were reliable. In each case where there was an observed difference greater than 9%, girls and non-minorities recidivated less than boys and minorities despite being classified in the same risk category. Moreover, when examining trends with high risk youth, 6 of the 9 studies showed differences greater than 9% for girls/minorities and their male/non-minority counterparts. Again, girls and non-minorities that shared risk classifications recidivated less than the respective male minority group.

In the studies reporting correlations, there were absolute differences in the validity coefficients as a function of gender and racial groups. Again, the statistical difference between the coefficients was not examined. In the two studies where race and gender were added to regression models, both found race and gender to be irrelevant. In perhaps the strongest study of dimensional identity, Schwalbe and others (2004) assessed

464 youth with the NCAR and observed their recidivist arrest over 12 months. While a quarter of the sample was comprised of girls, more than half were comprised of minorities. The NCAR showed a significant relationship between score and recidivist arrest; however the researchers took a further step in examining this relationship with regards to race and gender. Performing a Cox regression they found significant differences in the time to first arrest and pattern of arrest over time. Boys reached the base rate of offense at lower scores than girls. Researchers therefore concluded that the NCAR had poor dimensional identity.

Based on the literature reviewed here, the practical appeal of risk assessment instruments appears to far exceed their scientific foundation. While it is the case that the research to date provides encouraging results, much more work needs to be done before the current excitement over these methods is scientifically justified. A number of key points are relevant.

First, in the last decade there have been relatively few scientific investigations of the validity of risk assessment. Further, the studies summarized used a wide variety of methodologies and relatively short follow-up intervals. Of the 23 which this review identifies, it

appeared that reasonably representative samples were employed. Of the 23, 16 different instruments were employed and only on instrument appeared in more than two investigations. Further, the average follow-up time period was only 16 months and only eight studies included follow-ups. It should also be noted that there was wide variation in what delinquency outcomes were actually predicted. While future arrest was quite common, several studies attempted to predict when such arrests would occur in the future. However, little rationale was provided for the different outcomes employed. This is important because there is little research supporting the validity of risk assessment measures in predicting arrest, let alone time to first arrest. In examining the study conducted by Schwalbe and others (2004) it is clear that the likelihood of arrest can change over time and that trajectory is not necessarily linear. As such, the appropriateness of simply adding this variable to logistic regression models should be considered when examining this variable. Moreover, no attention was given to the accuracy of these predictions over time. No study checked for changes in risk scores over time. The studies offered no commentary as to how long the predictions should be considered accurate for the assessed

youth; therefore we cannot know how often the assessment should be given over the course of time.

Second, the multiple instruments demonstrated only modest agreement about what was important. The domains of education, family, substance abuse, offense history, and peers were consistently covered, but community, leisure, attitude, and personality were present less than half the time. Essentially, the concern expressed by Bonta (1996) continues. The risk assessment may well benefit from conceptual clarity and consensus as the field matures. For example, the ecological perspective may provide a more desirable as a basis for developing risk assessment measures. The failure of oversimplified risk models that do not account for person-environment fit reflects the theoretical work of Patterson (1989) and Elliot (1985) on the developmental pathways of juvenile criminality which suggests that the undesirable behavior of youth is both complicated and multifaceted. In some cases delinquency is a product of psychological pathology, in other cases a result of detrimental environmental factors or in still other cases some combination of both, intrinsically implying that juvenile offenders are a rather heterogeneous population (Cullen & Agnew, 2006).Thus, while a myriad of risk factors have been individually correlated with

delinquency, it is the combination and cumulative effect of risk factors that best predict delinquency (Schwalbe, Fraser, Day & Cooley, 2006).

Third, researchers seemed satisfied to classify juveniles as low, moderate or high risk and observe a significant difference in recidivism rates between the groups. While statistical significance is certainly an important criterion, strength of association and accuracy of prediction are more relevant in this case. For example, only one study used the recommended practice of looking for a 4 to 1 ratio between recidivism rates of the high and low groups. In examining the reported recidivism rates by risk level, it is clear that jurisdictions differ in recidivism rates that are considered low and those that are considered high. As seen in Table 3 (Sharkey, 2003), low risk girls assessed with the OCPDLRA recidivated twice as much as high risk girls assessed with the SDRRC. What then does it mean for a juvenile to be low risk if a quarter of this number will re-offend but half of this group will re-offend according to another instrument? This highlights the necessity of validating these measures repeatedly on different populations in different settings.

Fourth, the dimensional identity of these instruments is not well established. Research in this area is very much

in its infancy. There were 23 studies examining the validity of risk measures, yet only 12 were conducted on girls and 5 with non-majority ethnic groups. The success of these instruments varied from study to study and in fact the determination of the consistency of the measure across gender and ethnicity is unclear. Most studies simply reported correlations and recidivism rates for the respective groups and then used "inter-ocular t-tests" to decide if the differences were important. However, given the small sample sizes and limited replication of results it was impossible to attribute with any certainty differences in predictive validity to the dimensional identity of the instruments. Moreover, there were only two instruments, the YLS/CMI and NCAR that were validated more than once with diverse populations. In the case of the latter, the creator was the single individual validating the measure. In turn, the instrument with the greatest overall accuracy, the FORAI, was also only studied once and did not report its accuracy and validity with girls and diverse populations. Rather than determining that these instruments were valid and accurate with diverse populations, these studies indicate the necessity of further research with better samples and consistent methodology.

This review was able to provide us with a glimpse of what an ideal validation study should look like. Studies should have large enough samples that classification based on risk level can be made not only on the general populace but with both gender groups and relevant racial groups in mind. While checking for differences in delinquency rates between risk levels is useful, studies should also check that there is a 4 to 1 ratio between high and low risk groups. Most importantly, researchers should use their risk models on a case by case basis to predict relevant outcomes for their samples reporting the correct classification rates for the different racial and gender groups. More attention should also be given to the specific delinquency outcome of time to first offense. Also, when validating each of these instruments other researchers should be encouraged to follow-up these studies periodically with checks in different populations.

Chapter 5

Rationale for Validating the YLS/CMI

There is some evidence in the literature supporting perspectives that assert risk assessments will under-predict recidivism for African American youth due to exclusion of relevant factors at various levels of a youth's life like exposure to policing, resiliency, socio-economic status. This evidence also supports perspectives that assert girls are relatively minor offenders and offense rates are lower for female offenders in spite of sharing risks in common with their male counterparts. Consequently, instruments are highly likely to over-predict offending for this group.

The focus then of this study was the predictive validity of a widely used and popular risk measure, the Youth Level of Service Case Management Inventory, across gender groups and with Whites and Blacks in a particular juvenile justice system. The YLS/CMI is a 42 item risk assessment instrument with eight subscales [Education, Family, Drug Involvement or Substance Abuse, Personality, Attitude, Peer Association, Free-Time use, and Offense History]. Seven of the eight subscales are dynamic risk factors that have been shown in previous research to be

associated with offending. Offenders scoring between 0 and 8 are considered low risk, 9 and 22 are moderate risk, and 23 and 42 are high risk for re-offense.

In two previous studies, validating the YLS/CMI as a predictor of recidivism for minority groups and girls, the instruments received mixed but generally positive ratings. In the one study Jung and Rawana (1999) tracked 250 youth that were administered the YLS/CMI over six months, ultimately correlating their risk scores with their recidivism rates. They compared the risk-recidivism relationship of boys and girls; in addition to that of Native Canadians and non-Native Canadians. They found that the instrument differentiated between recidivists and non-recidivists for all subgroups. When we turn our attention to girls and Blacks, there have been relatively few studies that examined risk-recidivism relationships across groups. For instance, our previous validation work with the YLS/CMI with a one year follow-up suggested the possibility that the YLS/CMI under-predicted recidivism for African-American boys; however, given limited sample sizes for the various subgroups conclusions about the instrument's dimensional identity were tentative at best, necessitating larger sample sizes and a longer follow-up period.

Since then Flores, Travis, and Latessa (2004) further explored these differences in predictive validity between gender groups and racial groups on the YLS/CMI in one of the largest studies to date. Their sample was comprised 1,671 youth in total, with 79% of the sample being male and 21% being female. In turn 67% were white and the remaining portion reported simply as "nonwhite." A significant correlation between YLS/CMI score and re-arrest was found for each group with girls having the highest validity coefficient, $r=.318$, $p<.05$; and nonwhites having the lowest validity coefficient, $r=.208$, $p<.05$. However, when the actual recidivism rate by risk level analysis was performed there was some indication that differential predictive validity existed for some groups despite shared directionality of the relationship between risk score and recidivism for each group. A chi-square analysis revealed that the recidivism rate by risk level distribution was substantially different for the respective groups. In other words the rate of re-arrest for "low" risk white juveniles (8%) was substantially lower than that of "low" risk juveniles (40%) in the nonwhite sample. Moreover, the recidivism rate of "moderate" risk boys was greater than that of "very high" risk girls. Flores and others concluded that the YLS/CMI adequately predicts recidivism for the

various subgroups, but requires re-norming of the risk levels for those various subgroups.

Thus the YLS/CMI has received the greatest amount of attention from researchers and has proven to be relatively accurate in identifying offenders with a great likelihood of repeat contact with the juvenile justice system. However, despite the amount of attention it has received few studies have examined its validity specifically with girls and African Americans. In that context, this study answered two fundamental questions:

1. How well does the YLS/CMI predict recidivism for African Americans and girls?
2. Are there differences in the risk-recidivism relationship between these understudied groups and the traditional sample?

Chapter 6

Methodology

Setting

This study took place in a medium sized Midwestern county. The United States Census estimated the population of this county to be 277,000 in 2006. There are 60,940 youth under the age of 18 in the county. Female youth comprise 51% of the population, while African Americans comprise 11% of the population.

The Juvenile Court in this county serves youth in a number of programs including substance abuse, residential services, day treatments, sex offender services, diversion, anger management, family services, detention, probation, and truancy services.

Sampling and Data Collection

This study drew upon archival data from a pooled sample of two divisions in the court from March 2004 to March 2006. Juveniles from the Intake Division were assessed on the YLS/CMI upon entry into the system; while juveniles from the Probation Division comprised a cross-sectional sample in that they represented all juveniles receiving probation services in that time-period. The two samples were selected to represent youth entering the

juvenile justice system and those already under the jurisdiction of the court. There were no refusals.

Demographic Information. Ultimately this sample included 968 youth between the ages of 9 and 18; with a mean age of 14 years 6 months. There were 274 girls (28%) and 694 boys (72%). There were 509 White youth and 459 Black youth.

Data Sources. Demographic variable data was collected by the court through self-report of participants. Participants reported their age, address, school grade, gender, and race/ethnicity. Participants had the option of selecting African-American, White, Hispanic, or other.

Youth Level of Service/Case Management Inventory. The YLS/CMI has 42 items which comprise 8 subscales: Offense History, Family Circumstances, Education, Peer Group, Substance Abuse, Leisure Activity, Attitude, and Personality. Previous research (Schmidt, Hoge & Gomes, 2005) has demonstrated that the YLS/CMS has high internal consistency and reliability, which has been confirmed by this researcher and is reported elsewhere. Each item is scored dichotomously with a "yes" response receiving a single point and a "no" response receiving no points. These scores can be tallied for a risk score corresponding to each subscale and a total score. The creators of the

instrument have offered recommended cut scores for low, moderate and high classifications. Scores below 9 are low, below 24 are moderate and above 23 high.

Subscales. Items on each subscale of the YLS/CMI had dichotomously scored items [no=0, yes=1]. An examination of the subscale structure is presented below, but the eight original subscales (Hoge, Andrews, & Leschied, 2002) and their item configuration are listed here.

a) Prior and Current Offenses, Adjudications. Five items focusing on official criminal history (e.g. item I-a, "Three or More Prior Convictions").

b) Leisure/Recreation. Three items focusing on of free time (e.g. item II-b, "Could Make Better Use of Time").

c) Education/Employment. Seven items focusing on the school behavior and performance (e.g. item III-a, "Disruptive Classroom Behavior").

d) Peer Relations. Four items focusing on the characteristics of friends and acquaintances (e.g. item IV-c, "Lack of Positive Acquaintances").

e) Substance Abuse. Five items focusing on pattern of alcohol and illegal drug use and abuse (e.g. item V-d, "Substance Abuse Interferes with Life").

f) Family Circumstances/Parenting. Six items assessing experience of parental discipline and relationships (e.g. item VI-b, "Difficulty in Controlling Behavior").

g) Attitudes/Orientation. Five items reflecting criminogenic and prosocial tendencies (e.g. item VII-c, "Actively Rejecting Help")

h) Personality/Behavior. Seven items reflecting disruptive patterns of behavior (e.g. item VIII-g, "Verbally Aggressive, Impudent").

The item-total correlation for the sample is provided below.

Table 4.

Corrected Item-Total Correlations for the Intake and Probation Samples

Item	Corrected Item-Total Correlation
Offense History	α=.81
Three or More Prior Adjudications	0.31
Two or More Failures to Comply	0.40

(Table 4 Continued)

Prior Probation	0.51
Prior Detention	0.49
Three or More Current Adjudications	0.00
Family Circumstances;	α=.65
Inadequate Supervision	0.28
Difficulty in Controlling Behavior	0.65
Inappropriate Discipline	0.28
Inconsistent Parenting	0.42
Poor Relations/Father-Child	0.30
Poor Relations/Mother-Child	0.31
Education	α=.66
Disruptive Classroom Behavior	0.46
Disruptive Behavior on School Property	0.32
Low Achievement	0.56
Problems With Peers	0.33
Problems With Teachers	0.41
Truancy	0.47
Unemployed/Not Seeking Employment	0.27

(Table 4 Continued)

Peer Relationships	α=.70
Some Delinquent Acquaintances	0.32
Some Delinquent Friends	0.48
No or Few Positive Acquaintances	0.44
No or Few Positive Friends	0.40
Substance Use	α=.75
Occasional Drug Use	0.44
Chronic Drug Use	0.47
Chronic Alcohol Use	0.17
Substance Abuse Interferes With Life	0.47
Substance Abuse Linked to Offense(s)	0.30
Leisure	α=.79
Limited Organized Activities	0.46
Could Make Better Use of Time	0.50
No Personal Interests	0.45
Personality	α=.68
Inflated Self-Esteem	0.07
Physically Aggressive	0.44
Tantrums	0.42

(Table 4 Continued)

Short Attention Span	0.38
Poor Frustration Tolerance	0.53
Inadequate Guilt Feelings	0.42
Verbally Aggressive/Impudent	0.44
Attitude	α=.67
Pro-Criminal Attitudes	0.54
Not Seeking Help	0.51
Actively Rejecting Help	0.19
Defies Authority	0.60
Callous	0.48

Variables. Demographic variable data was collected through self-report of participants. Recidivism was defined as any new criminal petitions up to 24 months following the first YLS/CMI administration. Recidivism data was derived from the data management system of the court. Tickets and truancy violations were not counted. When referring to re-offense or recidivism, this actually a reference to new petitions a juvenile may have received. Two pieces of information are included on the recidivism collection form: recidivism (yes or no), and total number of charges. The form is completed by court staff using the court's data

management system and entered into a database to be used for analysis.

Procedure. Court personnel had four full days (32 hours total) of training in the YLS/CMI prior to administration. This training consisted of explaining definitions and scoring criteria for each items, practice cases and coding, and inter-rater reliability checks. During training 369 pairs of interviewers rated 36 different cases. The resulting percent agreement was 90%. Following training, one-twentieth of all coded cases were randomly checked by a second interviewer from an audio tape of the interview (N = 27). In all cases inter-rater reliability exceeded 90%.

Chapter 7

Results

Descriptive Statistics

The mean YLS/CMI for young offenders was 11.7 (SD=8.4), which was in the moderate risk range. Table 5 provides the item by item frequencies for each racial and gender group. On average, 28% of youth exhibited risk on any given item. The following items had the greatest likelihood of a youth exhibiting risk: poor relationship with father (51%), disruptive behavior in classroom (40%), disruptive behavior on school grounds (43%), low achievement (54%), some delinquent acquaintances (55%), some delinquent friends (42%), limited organized activities (57%), could make better use of time (61%), short attention span (41%), and poor frustration tolerance (49%).

Table 5. Item Frequencies (% Yes)

	African American		Whites	
Item	Boys	Girls	Boys	Girls
Offense History				
Prior Convictions	8%	5%	5%	5%
Failure to Comply*	15%	6%	10%	8%

(Table 5 Continued)

Prior Probation	17%	14%	17%	16%
Prior Custody	18%	20%	17%	14%
Current Convictions	3%	<1%	<1%	3%
Family Circumstances				
Inadequate Supervision*	32%	14%	27%	18%
Difficulty in Controlling behavior	40%	34%	40%	38%
Inappropriate Parental Discipline	15%	13%	23%	19%
Inconsistent Parenting	31%	24%	37%	32%
Poor Relationship with Father*	60%	47%	46%	51%
Poor Relationship with Mother*	26%	32%	28%	31%
Education				
Disruptive Behavior in Classroom*	50%	41%	41%	28%
Disruptive Behavior of School Grounds*	46%	49%	39%	38%
Low Achievement*	60%	45%	56%	53%
Problems with Peers*	30%	46%	31%	27%
Problems with Teachers*	38%	31%	36%	26%

(Table 5 Continued)

Truancy*	41%	34%	33%	43%
Unemployed	6%	3%	6%	6%
Peer Group				
Some Delinquent Acquaintances	60%	49%	55%	55%
Some Delinquent Friends*	47%	41%	35%	45%
Some Positive Acquaintances*	36%	19%	30%	23%
Some Positive Friends*	30%	20%	26%	16%
Drug Involvement/Substance Abuse				
Occasional Drug Use*	34%	19%	39%	40%
Chronic Drug Use*	21%	8%	20%	20%
Chronic Alcohol Use*	2%	<1%	9%	9%
Substance Use Interferes with Life*	18%	8%	22%	23%
Substance Abuse Linked to Offense*	9%	2%	17%	19%
Free time Use				
Limited Organized Activities	56%	62%	56%	53%
Better Use of Free Time	66%	62%	57%	60%

(Table 5 Continued)

No Personal Interests	25%	28%	21%	20%
Personality				
Inflated Self Esteem*	7%	2%	6%	2%
Physically Aggressive	34%	41%	34%	30%
Tantrums*	24%	23%	30%	38%
Short Attention Span*	45%	30%	55%	32%
Poor Frustration Tolerance	50%	44%	52%	51%
Inadequate Guilt	21%	14%	19%	13%
Verbally Aggressive Attitudes	27%	35%	26%	32%
Pro-criminal Attitudes*	26%	16%	18%	13%
Not Seeking Help*	23%	12%	17%	16%
Rejecting Help	11%	6%	8%	8%
Defies Authority*	24%	14%	18%	12%
Callous*	15%	6%	10%	5%

Group Differences on Frequency of Exhibited Risk

Mean scores for African American boys and girls, and White boys and girls are detailed in Table 6. Girls (M=10.6) scored significantly lower on the YLS/CMI than boys (M=12.1), $t(968)=-2.6$, $p<.01$, however there were no significant differences in scores for African Americans

(M=11.8) and Whites (M=11.5). Within the two racial groups there were no significant differences in mean score for boys or girls.

There were however, significant differences in distribution of exhibited risk on individual items by race and gender group as seen in Table 5. Boys exhibited risk at a significantly greater rate than girls on the following items: failure to comply, inadequate supervision, poor relationship with father, disruptive behavior in the classroom, low achievement, problems with teachers, positive friends, occasional drug use, chronic drug use, chronic alcohol use, substance abuse inferring with life, inflated self-esteem, short attention span, pro-criminal attitudes, not seeking help, and defying authority.

Racial differences in distribution of exhibited risk were limited to drug involvement items. Table 6 shows that with the exception of the drug involvement and attitude subscales, there was little variation in mean score across subgroups. Moreover African American female youth exhibited significantly lower risk for re-offense on the drug involvement subscale than the other subgroups.

Table 6. Subscale scores and ROC by race and gender

	African Americans				Whites			
	Boys	AUC	Girls	AUC	Boys	AUC	Girls	AUC
Offense History	0.6	n.s	0.5	n.s	0.5	0.57	0.5	n.s
Family Circumstances	2	n.s	2	0.66	2	n.s.	1.9	0.75
Education	2.7	0.59	2.5	0.61	2.4	0.59	2.2	0.67
Peer Association	1.7	0.6	1.3	n.s.	1.5	0.62	1.4	0.66
Drug Involvement	0.8	0.58	0.4	n.s.	1.1	0.59	1.1	n.s
Free time	1.5	n.s	1.5	n.s.	1.3	0.62	1.3	n.s
Personality	2.1	n.s	1.9	0.65	2.2	0.58	2	0.69
Attitude	1	0.55	0.5	n.s.	0.7	n.s.	0.5	0.61
Overall	12.4	0.6	10.6	0.675	11.7	0.63	10.9	0.72

As seen in Table 7, this trend is reflected in the distribution of subgroup members by risk level. Overall, 45% of the sample was low risk, 42% moderate risk and 14% high risk. African American female offenders and African American male offenders exhibited significant differences in risk score level with twice as large a proportion of boys (18%) than girls (6%) rated as high risk. Girls overall were under-represented in the high risk group compared to boys.

In turn, for each group nearly half of their members were rated as low risk for re-offense, while no more than one in five exhibited a high risk of recidivism. When

examining mean scores on each of the subscales by gender and racial group few differences emerged.

Table 7. Descriptive Statistics

		N	Mean YLS/CMI Score (SD)	Risk Level and Recidivism Rate						
				Low	RR	Moderate	RR	High	RR	
White		509	11.5 (8.4)	46%	23%	42%	42%	13%	59%	X^2=34.8, p<.05
	Girls	146	10.9 (8.4)	48%	11%	42%	34%	10%	60%	
	Boys	363	11.7 (8.4)	45%	28%	42%	44%	14%	59%	
Black		459	11.8 (8.4)	44%	36%	41%	56%	15%	61%	X^2=21.6, p<.05
	Girls	128	10.2 (7.3)	56%	23%	38%	41%	6%	63%	
	Boys	331	12.5 (8.5)	40%	43%	43%	61%	18%	61%	
Total		968	11.7 (8.4)	45%	29%	42%	48%	14%	60%	X^2=54.9, p<.05
	Boys	694	12.1 (8.5)	42%	35%	42%	52%	16%	60%	
	Girls	274	10.6 (8.0)	52%	17%	40%	37%	8%	61%	

Significant Differences between Subgroups in Risk Level-Recidivism Relationship

(Table 7 Continued)
a. White Low Risk Boys versus Girls
X^2=7.8, p<.05
b. Black Low Risk Boys versus Girls
X^2=8.2, p<.05
c. Black Moderate Risk Boys versus Girls
X^2=6.0, p<.05
d. Low Risk Girls versus Boys
X^2=14.5, p<.05
e. Moderate Risk Girls versus Boys
X^2=7.3, p<.05
f. Black Low Risk Boys versus White Low Risk Boys
X^2=6.8, p<.05
g. Black Moderate Risk Boys versus White Moderate Risk Boys
X^2=8.1, p<.05

Question 1. How well does the YLS/CMI predict recidivism for African Americans and girls?

Over a two year follow-up period, 41% of assessed youth re-offended (n=968). There were substantial differences in re-offense rates by groups (see table 7). For instance, boys (46%) re-offended at a significantly higher rate than girls (29%). Furthermore, over half of African American male youth in the study recidivated (54%). This rate was followed in size by White male youth in which 39% of this group re-offended. African American female and White female youth had lower recidivism rates with 32% and 26% respectively re-offending over the twenty-four month period following assessment.

Risk-Recidivism Relationship

Overall, the YLS/CMI demonstrated its ability to differentiate between non-recidivists and re-offenders over a two year period moderately well (see table 8). In each

group, recidivists scored significantly higher than non-recidivists on the YLS/CMI ($p<.01$). The difference in mean scores between recidivists and non-recidivists was as follows: African American males (3.5), African American females (5.1), White males (4), and White females (6).

Logistic regression allows us to examine the relationship between the continuous variable, YLS score, and the dichotomous variable, re-offense. Moreover, the statistical technique can provide information on incremental percentage increase in likelihood of offense for each point scored on the YLS/CMI. For example in each of the groups, there was approximately a 1.06 percent increase in the likelihood of re-offense for each point scored on the YLS/CMI. From 3% to 13% of the variation in re-offense rate was explained by YLS/CMI, thus the effect size for each group was small but significant. Overall, 7% [Nagelkerke R^2] of the variation in offense rates was explained by total risk score.

Table 8. Hierarchical Logistic Regression for YLS/CMI score and New Offenses

Group	Exp(B)	95% C.I. for Exp(B)	Nagelkirke R^2	Correct Classification Rate
African American				
Boys	1.04	1.01 to 1.07	.03	60%
Girls	1.09	1.03 to 1.1	.10	69%
Whites				
Boys	1.05	1.02 to 1.08	.06	62%
Girls	1.09	1.03 to 1.1	.13	75%

In practical terms this resulted in 30% of re-offenders being correctly identified with logistic regression and 81% of non-recidivists being correctly identified (see table 8). The overall correct classification rate was 64%. African American boys had the lowest correct classification rate with 60% correctly identified. Girls had the highest correct classification rate with 75% of offenders correctly identified. So while the relationship between YLS/CMI score and recidivism had a relatively small effect size, predictions based on these relationships produced a wide range of correct classification rates based on group membership. African

American boys were incorrectly classified at the greatest rate with 47% of boys mis-identified based on YLS/CMI score. Both girl groups were correctly classified over 65% of the time.

Given the recommendation by methodologist that Receiver Operating Characteristics and its corresponding AUC coefficients better reflect correct classification rates in a binary classification system by controlling for base rate of re-offense, ROC curve analysis was also conducted. The AUC refers to the sensitivity of the instrument versus the specificity. Sensitivity is the proportion of true positives; that it to say the number of offenders predicted to re-offend that did in fact re-offend. Specificity is the number of true negatives; that is to say the number of offenders predicted to not re-offend that did not in fact re-offend.

The YLS/CMI had an AUC of .641, $p<.05$ overall, indicating it performed 14% better than chance in correctly classifying offenders by risk score. The YLS/CMI produced similar AUCs for each of the subgroups (Girls, AUC=.68; Boys, AUC=.61; White, AUC=.66; Black, AUC=.63) (see table 7).

Furthermore, there was a significant difference in the distribution of re-offenders by risk level ($X^2=54.9$,

p<.05) with high risk offenders being more than twice as likely to re-offend as low risk offenders. When examining distribution of re-offenders by risk level for each of the gender and racial groups, there was positive relationship between offense level and offense rate. This can be seen in Table 7 with recidivism rates becoming increasingly higher as risk level increases for each of the gender and racial subgroups.

Question 2. Are there differences in the risk-recidivism relationship between these understudied groups and the traditional sample?

Due to the small overall effect and sample size constraints, differences between groups were not able to be detected using logistic regression. Furthermore, an examination of the relationship between cumulative risk score and re-offense did not reveal any significant differences in AUC for the gender groups (Girls, AUC=.68; Boys, AUC=.61) or racial groups (White, AUC=.66; Black, AUC=.63). Despite non-significant differences in over-all classification rates, there were substantial practical differences in the distribution of correctly identified recidivists versus non-recidivists. There was a significant difference in true negatives between White girls, White boys and African American boys. Using a binomial proportion

difference test, we examined whether these proportions were significantly different. There was a significant difference between true positives for African American boys and each of the other groups as reflected in the overall classification rates found in table 8.

There were several significant differences within risk levels by subgroup as shown in Table 7 [Chi-Square significant difference test coefficients are reported at the bottom of Table 7]. For instance, White low risk boys and African American low risk girls were twice as likely to re-offend as White low risk females. Moreover, African American low risk boys were four times as likely to re-offend as White low risk girls. In turn, low risk boys exhibited greater re-offense rates than girls overall. When examining subgroup difference in recidivism within the moderate risk level, similar differences are apparent. African American moderate risk male youth had substantially higher rates of re-offense than the other subgroups with nearly 61% re-offending, while White moderate risk female youth had the lowest recidivism rate of that risk level (34%), which was even lower than that of Black low risk male youth. It should be noted however that there were no significant differences in recidivism rates between Black female youth and White female youth. When comparisons were

made of high risk subgroups, there are no identifiable differences in recidivism rates by race or gender. Generally, two-thirds of high risk offenders recidivated over the two year period.

Recidivism-Subscale Relationship

Given the differences in re-offense rates within risk levels by racial and gender subgroups, further analysis was performed examining the correct classification rates for each of the subscales by gender and racial groups. Table 6 shows that White female youth had five subscales [family, education, peer, personality, attitudes] with AUCs that were significantly better than chance in correctly classifying non-recidivists and recidivists. White male youth in turn had six subscales where this was the case [offense history, education, peer group, drug involvement, free time, and personality]. Black female youth had the fewest subscales that predicted re-offense better than chance with family, education, and personality subscales meeting this criteria. Black male youth in turn had four subscales [education, peer, substance abuse/use, and attitude].

Chapter 8

Discussion

Juvenile justice systems have a vested interest in correctly predicting which young offenders are most likely to come in repeated contact with the legal system. There is considerable evidence that repeat offenders commit the majority of offenses, despite comprising a numerical minority of all offenders. Thus treating all offenders as though they have a similar likelihood of recidivating can quickly exhaust system resources rehabilitating and incapacitating the majority of young offenders who are not likely to re-offend regardless of system response to their first offense. This issue compounded by differential offending rates amongst racial and gender groups that some paradigms attribute to differential prosecution and policing. Thus the current practice of formal risk assessment is highly recommended because in theory, 1) risk assessments correctly predict recidivist based on risk level, and 2) risk assessments can help reduce disproportionate minority contact through assessing risk based on criminogenic dynamic risk factors.

The problem with this approach lies in the fact that few risk assessments have been validated across racial and

gender groups and these instruments very well may suffer from poor dimensional, resulting in over and under-prediction of re-offense depending on the subgroup. This is possible due to the inclusion and exclusion of risk factors that are relevant for some groups and irrelevant for other groups. For instance, no risk prediction instruments check for risk due to surveillance level and zero-tolerance policing policies that exist in some neighborhoods as opposed to others. As race and class stratification continue to permeate U.S. neighborhoods, it can be postulated that the re-offense rates of Black males for instance should be under-predicted by risk assessments due to the exclusion of macro-level policing factors on risk assessments. Moreover, while risk assessments are supposed to solicit objective and verifiable information on youth, there is still some degree of judgment and latitude that assessors have in scoring youth on the instruments. Thus, if some theories are correct biases can still conceivably trickle into the scoring on these measures. Thus it is highly recommended by proponents of the risk assessment model of delinquency prevention that researchers validate these instruments on a wide cross-section of juveniles, especially with regards to African Americans and girls.

In that context this study offered some examination of the dimensional identity of the YLS/CMI, a relatively popular risk assessment measure with juveniles. In keeping with other findings reported in the literature on this instrument, this study found the YLS/CMI does moderately well in identifying young re-offenders over a two year period. Moreover, as risk level increases, so do actual re-offense rates. High risk offenders, re-offended at twice the rate of low risk offenders. This result was also encouraging because there were three times as many low risk juveniles as high risk juveniles, allowing for the possibility to divert or refract nearly 45% of the offender population based on risk level. This would reduce caseload sizes and free up programming resources for moderate and high risk offenders. The risk prediction was shown to be dependable over-all by the ROC analysis, which indicated that the YLS/CMI was 14% better than chance at classifying youth. While the amount of variation in offense rates explained by the YLS/CMI was relatively small, the over-all correct classification rate was over 60%. It can be concluded that in examining the YLS/CMI's performance with the over-all offender population, it is a powerful instrument deserving of its popularity with justice officials. However, this over-all conclusion says nothing

for its performance when parceling out subgroups that may have had different risk-recidivism relationships. There are then two questions that immediately stem from this original analysis. Does the YLS/CMI predict re-offense with subgroups like African Americans and girls? Are their differences in these risk prediction relationships? While the questions appear similar, we see that it is possible for the instrument predict re-offense for these groups, but have differences with regards to the degree for which those predictions are correct. Moreover, the risk factors that contribute to those classifications may differ by subgroup with regard to predictive validity.

Question 1. How well does the YLS/CMI predict recidivism for African Americans and girls?

When examining these predictions by race and gender, we see that each group had a significant relationship between risk and recidivism. Recidivists for both girls and boys of both racial groups exhibited significantly higher YLS/CMI scores than non-recidivists. The amount of variation in offending explained by YLS/CMI score ranged from 3% for African American boys to 13% for White girls. While this effect size was clearly small for each subgroup, the correct classification rate of offenders was still greater than 50% for each subgroup. Despite the small

effect size for African American boys, 60% of these offenders were correctly classified as non-re-offenders or re-offenders. The ROC analysis for each subgroup also indicated the YLS/CMI was a good predictor of re-offense for each group, with AUCs ranging from .61 with boys and .68 for girls. Finally, for each group, re-offense rates significantly and substantially increased by risk level.

Again, these findings support the YLS/CMI as a predictor of recidivism with young offenders. However, it is also clear from considering the results that were possible differences in the validity coefficients for each subgroup. Thus, if case management decisions are made by risk score or risk level differences in rates of delinquency can still exist, which brings us to the second question of this study.

Question 2. Are there differences in the risk-recidivism relationship between African Americans, girls and the traditional sample?

In examining differences between the various groups, the first round of differences emerges from scoring on the instrument itself. Girls on averaged scored two points lower on the YLS/CMI than boys. There were no significant differences in scoring based on race. The item by item frequency analysis revealed that were a substantial number

of YLS/CMI items that boys were more likely exhibit risk on than girls. With the exception of drug involvement and attitudes, the variation in scores by subgroup did not aggregate in overall subscales.

These small differences in item by item exhibited risk did result in larger proportions of some groups being rated as high risk. More than twice as many African American boys were classified as high risk than African American girls. However, there were no significant differences in the distribution of low risk offenders by subgroups. Moreover, at least four in ten young offenders were considered low risk. Thus, each group has nearly half of its population which could be eligible to be refracted or diverted from the system.

There were further differences in the recidivism rates by subgroup. Nearly twenty percent more boys than girls re-offended. African American males had the highest re-offense rate, while White females had the lowest re-offense rates. The literature suggests that the YLS/CMI typically has a correlation coefficient of .20 to .30. At the .05 alpha level, to have a 60% chance of detecting a difference for correlations between group, this would require sample sizes for each group in excess of 1,300 assessed youth. Therefore, while a regression was conducted to answer the

first question, the effect size produced by this regression was not used in determining whether differences between groups exist. Thus due to sample size constraints, significant differences could not be detected using regression, differences in the effect sizes which ranged from 3% to 13% were not found to be significant. However, when examining the correct classification rates by subgroup using proportion difference test, significant differences were detected between African American boys, White boys and the two girl groups. The YLS/CMI performed much better with girls than with boys. This may be due to the low recidivism rate of girls and corresponding large group of offenders classified as low risk. While both boy groups had a larger re-offense rate but similarly sized low risk groups, leading to a larger false negative rate for boys than girls. This meant that the YLS/CMI under-predicted re-offense for African American boys in particular. Furthermore, compared to the other groups the YLS/CMI over-predicted re-offending for White girls.

The YLS/CMI under-predicted recidivism for Black low risk male youth, while over-predicting recidivism for White low risk female youth. This trend of over and under prediction for these groups was also true of moderate risk groups. Only high risk youth had similar rates of re-

offense regardless of racial or gender group. It should be noted however that within each subgroup, re-offense rates incrementally increased by risk level suggesting that weighting scores by race and gender grouping could improve the dimensional identity of the YLS/CMI.

Given these findings, further analysis of the predictive validity of each subscale by subgroups revealed that while the subgroups scored similarly on the respective subscales, there were differences between groups in correct classification rates by subscale. The majority of YLS/CMI subscales could individually predict re-offense for White male youth whereas as few as three subscales on the instrument individually predicted re-offense for Black female youth. In turn, education was the single criminogenic risk factor that could predict re-offense for each of the four subgroups, while offense history only had predictive validity with White male youth. Both subgroups of girls demonstrated a unique predictive risk relationship for family circumstances and personality.

These different predictive relationships between criminogenic factors and recidivism by race and gender suggest the necessity of thoroughly examining the factor structure of the YLS/CMI by group and comparing alternative risk models. Moreover, for some groups there is a strong

possibility that relevant risk factors are being excluded, which merely weighting YLS/CMI scores would fail to address.

Study Limitations

This study did not control for socio-economic status and other macro-level variables that mitigate and exacerbate delinquency risk. Moreover, this study only examined the risk-recidivism relationship for African American and White young offenders. While risk assessment data was available for Hispanic youth, sample size constraints did not allow for validating the instrument with this group or other ethnic and racial groups.

Future Study

A future direction of research should include accounting for the multi-level nature of criminogenic risk by accounting for possible differences in structural factors that naturally increase likelihood of identified recidivism like policing levels, surveillance, and criminal justice policy. By conducting hierarchical linear modeling with neighborhood variables this could be accomplished. Moreover clustering youth across the subscales by subgroup can also identify differences in patterns of risk that could explain differing recidivism rates by group within risk levels. Research in this regard will reduce over and

under-predictions of criminal activity for female and African-American subgroups.

Chapter 9

Conclusions

The YLS/CMI is a powerful instrument capable of classifying youth based on likelihood of re-offense. The instrument has both dynamic and static risk factors which are generally accepted as strong predictors of re-offense. Moreover, in classifying youth the instrument allows for identifying youth that would benefit from diversion or being completely refracted from the system. In making these classification nearly half of the offender population can be diverted. Using this classification system benefits the juvenile justice system because it can result in lower case-load sizes, freeing up of resources that can be better used with moderate and high risk offenders, and provides specific criminogenic areas in a youth life that can be targeted with interventions leading better program efficacy and program planning. Consequently, this study further evidences why juvenile justice officials are so excited about the potential of risk assessments as a tool of delinquency prevention and rehabilitation.

That said, justice officials are also excited about risk assessments potential to reduce disparities in formally adjudicated crime due to system biases against

certain racial and gender groups. This study, however, shows that such excitement is premature. While there is a general risk-recidivism trend for African Americans, Whites, boys and girls where recidivism rates increase as risk level increases; the degree to which those increases occur differ by gender and racial group. Girls are less likely to offend and are less likely to be classified as high risk offenders. However, general trends in juvenile justice could mitigate the benefits of risk assessment with this group by widening the adjudication net for girls. For example, by expanding the use of truancy court and targeting status offenses a larger proportion of girls will fall under the purview of the court for first time and minor offenses. Thus, increasing the over-all pool of one-time offenders as compared to that of boys. The YLS/CMI in that case over-predicts offending in this group in that the low risk group of offenders is actually different than the low risk group of offenders for boys. One can then question whether there are any iatrogenic effects in exposing girls to moderate and high risk offenders in the system.

Furthermore, this study indicated the YLS/CMI under-predicted re-offense for boys, especially African American boys. African American low risk boys were comparable in proportion to other low risk groups; however they re-

offended at a substantially higher rate than the other groups. One can conclude one of two things from this trend. Either African American boys truly offend more than other groups or they are under greater surveillance than other groups due to first time offending putting them under the purview of the juvenile court. This study did not attempt to determine differences in actual deviance rates and formally prosecuted delinquency rates. Instead it determined whether the youth exhibited similar risk levels across racial and gender groups. The results indicated that African American boys did in fact exhibit similar levels of low criminogenic risk.

In exhibiting similar risk but dissimilar offending rates, item by item analysis and subscale ROC analyses, indicated that there were substantial differences in how youth were assigned to the respective risk levels. Boys quite simply scored higher on the YLS/CMI items than girls. Moreover, each subgroup differed on the set of subscales that could individually predict re-offense for their subgroup. This suggests the possibility that a number of factors were included for some groups that were not relevant. Further, several factors were excluded for some that were relevant. We can conclude than that simply weighting the YLS/CMI by race or gender would not improve

case-management or program planning due to incomplete/faulty information derived from the YLS/CMI. Instead the justice officials would benefit from tracking the pathway from deviance to formally adjudicated delinquency. Differences in those pathways to contact with law enforcement can have considerable impact on the rates of "crime" and exhibited risk. Moreover, the juvenile justice system would also benefit from expanding assessment beyond micro-level and individual based risk factors to macro-level and system based risk factors, which contribute offense rates. This can be accomplished by examining neighborhood and school effects in multi-level analysis.

Literature Cited

Akers, R. (2000). *Criminological Theories: Introduction, Evaluation, and Application.* Los Angeles: Roxbury.

Albonetti, C. A. (1987). Prosecutorial discretion: The effects of uncertainty. Law & Society Review, 21, 291-313.

Andrews, D., Kiessling, J., Robinson, D.,& Mickus, S. (1986). The risk principle of case classification: An outcome evaluation with young adult probationers. *Canadian Journal of Criminology, 28,* 377-384.

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

Bandura, A., & McDonald, F. (1963). The influence of social reinforcement and the behavior of models in shaping children's moral judgments. *Journal of Abnormal and Social Psychology, 67,* 274.

Barnoski, R. Assessing Risk for Re-offense: Validating the Washington State Juvenile Court Assessment. *Olympia: Washington State Institute for Public Policy, 2004.*

Belknap, J. (1996). *The Invisible Woman: Gender, Crime, and Justice.* Cincinnati: Wadsworth.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18-32). Thousand Oaks, CA: Sage.

Braithwaite, John. (1989). *Crime, Shame, and Reintegration.* Cambridge: Cambridge University Press.

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design.* Cambridge, MA: Harvard University Press.

Burgess, E.W. (1928.) "Factors Determining Success or Failure on Parole." In A.A. Bruce (ed.). *The Working of the Indeterminate Sentence Law and the Parole System in Illinois* (pp. 205-249). Springfield, IL:

Illinois Committee on Indeterminate-Sentence Law and Parole.

Burke, R.(2001) *An Introduction to Criminological Theory.* Willan Publishing, Devon.

Catchpole, R. E. & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior, 30(6),* 688-708.

Chesney-Lind, M. & Shelden, R. G (2004). *Girls, Delinquency and Juvenile Justice* (3rd ed.). Belmont,CA: Wadsworth.

Bureau of Justice Statistics. Sourcebook Criminal Justice Statistics, 2002.30th Ed. (Pub NCJ 203301) Washington DC: BJS, 2004. Available at www.albany.edu/sourcebook.

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

Courtwright, D. T. (1996). The drug war's hidden toll. *Issues in Science and Technology,* 73.

Cullen, F. T., & Agnew, R. (2006). *Criminological theory past to present: essential readings. (Third edition).* LA: Roxbury Publishing Company.

DeComo, R. E., Wagner, D., & Baird, S. C. (1994). *Validation of the Oregon risk assessment tools used for classification of offenders on supervision.* San Francisco: National Council on Crime and Delinquency.

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

Duncan, R., Kennedy,W., & Patrick, C. (1995). Four-factor model of recidivism in male juvenile offenders. *Journal of Clinical Child Psychology, 24,* 250-257.

Eck, J. E., & Weisburd, D. (1996). *Crime in Place.* Monsey: Criminal Justice Press.

Elliot, D., Huizinga, D., & Ageton, S. (1985). *Explaining delinquency and drug use.* Beverly Hills. CA: Sage.

Erikson, E. H. (1963). Childhood and society. Norton 2d.ed.

Fellner, J., & Mauer, M. (October 1998). "Losing the Vote: The Impact of Felony Disenfranchisement Laws in the United States," Human Rights Watch and The Sentencing Project.

Flores, A., Travis, L. F., and Latessa, E. J. (2004). Case Classification for Juvenile Corrections: An assessement of the YLS/CMI. NCJRS Report.

Frank, S. J., Paul, J. S., Marks, M., and Van Egeren, L. (2000). *Journal American Academy of Child Adolescent Psychiatry, 39(10)*, 1300-1308.

Freudenberg, N. (2002). Adverse effects of US jail and prison policies on the health and well-being of women of color. *American Journal of Public Health, 92,* 1895-1899.

Funk, S. (1999). Risk assessment for juveniles on Probation: a focus on gender. *Criminal Justice and Behavior,* 26, 1, 44-68.

Gottfredson, S. D., & Moriarity, L. J. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency, 52, 1,* 178-200.

Grinberg, I., Dawkins, M., Dawkins, M. P., and Fullilove, C. (2005). Adolescents at risk for violence: an initial validation of the Life Challenges Questionnaire and Risk Assessment Index. *Adolescence ,40,* 159.

Grove, W. M. and Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical algorithmic) prediction procedures: The Clinical-Statistical Controversy. *Psychology, Public Policy, & Law. 2(2):*293-323.

Guevara, L., Herz, D., & Spohn, C. (2006). Gender and juvenile justice decision making: What role does race play? *Feminist Criminology,* 1, 4, 258 - 282.

Harlow CW. *Profile of Jail Inmates—1996*. Washington, DC: US Dept of Justice; 1998:1-15. Bureau of Justice Statistics special report NCJ 164620.

Harris, P. (2006). What Community Supervision Officers Need to Know About Actuarial Risk Assessment and Clinical Judgment. *Federal Probation, 69* (2), 8-14.

Richard D Hartley, Sean Maddan, Cassia C Spohn. (2007). Prosecutorial Discretion: An Examination of Substantial Assistance Departures in Federal Crack-Cocaine and Powder-Cocaine Cases. *Justice Quarterly : JQ, 24*(3), 382-407.

Heider, F. (1958). *The psychology of inter-personal relations*. New York: Wiley.

Hishinuma, E. S., Johnson, R. C., Kim, P. S., Nishimura, S. T., Makini, G. K., Andrade, N. N., Yates, A., Goebert, D. A., Mark, G. Y., Mayeda, D. T., and Revilla, L. A. (2005). Prevalence and correlates of misconduct among ethnically diverse adolescents of Native Hawaiian/Part-Hawaiian and Non-Hawaiian Ancestry. *International Journal of Social Psychiatry* 51(3), 242-258.

Hodges, K., and Kim, C. (2000). Psychometic study of the child and adolescent functional assessment scale: Prediction of contact with the law and poor school attendance. *Journal of Abnormal Child Psychology,* 28,3, 287-297.

Howell, J. C. (2003). *Preventing & reducing juvenile delinquency: A comprehensive framework*. Thousand Oaks, CA: Sage.

Hubbard, D. J. & Pratt, T. (2002). A meta-analysis of the predictors of delinquency amongst girls. Journal of Offender Rehabilitation, *Vol. 34(3), 2002. Pp. 1-13.*

Huff, D., and Prell, L. The Proposed Iowa Juvenile Court Intake Risk Assessment. Division of Criminal and Juvenile Justice Planning and Statistical Analysis Center.

Jackson, P. I., & Carroll, L. (1981). Race and the war on crime: The socio-political determinants of municipal

police expenditures in 90 non-Southern US cities. American Sociological Review, 46, 290-305.

Jensen, Gary F. 2003. Social Disorganization Theory. *Encyclopedia of Criminology*. Richard A. Wright (Editor). Fitzroy Dearborn Publishers.

Jimerson, S. R., Sharkey, J. D., O'brien, K., Furlong, M. (2004). The Santa Barbara Assets and Risks Assessment to predict recidivism among male and female juveniles: An investigation of inter-rater reliability and predictive validity. *Education and Treatment of Children* 27, 4, 353-373.

Jung, S., and Rawana, E. (1999). Risk and Need Assessment of Juvenile Offenders. *Criminal Justice and Behavior,* 26, 1, 69-89.

Kent, S.L., & Jacobs, D. (2005). Minority threat and police strength from 1980 to 2000: A fixed-effects analysis of nonlinear and interactive effects in large cities. Criminology, 43, 731-760.

Kerstetter, W. (1990). Gateway to justice: Police and prosecutorial response to sexual assaults against women. Criminology, 81, 267-313.

Kliewer,W., Fearnow, M. D.,&Walton, M. N. (1998). Dispositional, environmental, and context-specific predictors of children's threat perceptions in everyday stressful situations. *Journal of Youth and Adolescence, 27,* 83-100.

Kliewer, W., & Kung, F. (1998). Family moderators of the relation between hassles and behavior problems in inner-city youth. *Journal of Clinical Child Psychology, 27,* 278-292.

Kliewer,W., Lepore, S. J., Oskin, D., & Johnson, P. D. (1998). The role of social and cognitive processes in children's adjustment to community violence. *Journal of Consulting and Clinical Psychology, 66,* 199-209.

Kraska, Peter B. (2006). Criminal Justice Theory: Toward Legitimacy and an Infrastructure. *Justice Quarterly,* 23 (2), 0741-8825. Retrieved September 13, 2007, from http://www.informaworld.com/10.1080/07418820600688735

Krisberg, B. & Austin, J. (1978) *The children of Ishmael. Palo Alto, CA: Mayfield.*

Krysik, J. and Lecroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12(1),* 71-81.

Lahey, B. B., Goodman, S. H.,Waldman, I. D., Bird, H., Canino, G., Jensen, P., et al. (1999). Relationship of age of onset to the type and severity of child and adolescent conduct problems. *Journal of Abnormal Child Psychology, 27,* 247-260.

Leiber, M. J. & Fox., K. C. (2005). Race and the impact of detention on juvenile justice decision making. *Crime Delinquency,* 51, 470 - 497.

Lipsey, M.W. (2005). *What works with juvenile offenders: Translating research into practice.* Presented at the Adolescent Treatment Issues Conference, Tampa, FL.

Liska, A. (1987). *Perspectives on Deviance,* 2nd ed., Prentice-Hall.

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: A Review. *Psychological Bulletin, 93, 68-99.*

Loeber, R., & Farrington, D. P. (Eds.). (1999). *Serious and violent juvenile offenders.* London: Sage.

Loftin, C., & McDowall, D. (1982). The police, crime, and economic theory: An assessment. *American Sociological Review, 47,* 393-401.

Lowenkamp, C. T., Holsinger, A., & Latessa, E. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior, 28*(5), 543-563.

Maguire, E. R. (2001). Research evidence on the factors influencing police strength in the United States. In Hiring and retention Issues in Police Agencies: Readings on the determinants of police strength, hiring and retention of officers and the federal cops program (pp. 7-25). A report to the National Institute

of Justice by Christopher S. Koper, Edward R. Maguire, and Gretchen E. Moore. Retrieved from http:// www.urban.org/pdfs/Hiring-and-Retention.pdf

Marczyk, G. R., Heilbrun, K., Lander, T., and DeMatteo, D. (2003). Predicting juvenile recidivism with the PCL:YV, MAYSI, and YLS/CMI. *International Journal of Forensic Mental Health. 2(1),* 7-18.

Mauer, M. (1999). *Race to incarcerate.* New York: New Press.

Maxfield, M.G., and C.S. Widom, "The Cycle of Violence: Revisited Six Years Later," *Archives of Pediatrics and Adolescent Medicine* 150 (1996): 390-395.

May, L. & Sharratt, S. (1994). *Applied Ethics: A Multicultural Approach,* Englewood Cliffs, NJ: Prentice-Hall.

McKinnon, L. J. (2004). Predicting risk of violence in a young offender population: The predictive validity of the PCL:YV and the YLS/CMI. *Dissertation:Lakehead University (Canada),* 54.

Mcreynolds, L. S. (2005). Prediction of juvenile recidivism: Influence of gender, psychiatric disorder, and neighborhood disadvantage.*Dissertation Abstracts International Section A: Humanities and Social Sciences* Vol 66(1-A): 358.

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

Myers, H. F., Taylor, S., Alvy, K. T., Arrington, A., & Richardson, M. A. (1992).Parental and family predictors of behavior problems in inner-city Black children. *American Journal of Community Psychology, 20,* 557-576.

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

The National Center on Addiction and Substance Abuse (CASA) at Columbia University. (2004, October). *Criminal*

neglect: substance abuse, juvenile justice, and the children left behind. New York: The National Center on Addiction and Substance Abuse (CASA at Columbia University). Retrieved May 15, 2006 from http://www.nicic.org/Library/020178

National Council on Crime and Delinquency. (2001). Alameda County Placement Risk Assessment Validation. Executive Summary.

Nowicka, M. (2005). The Level of Service Inventory-Ontario Revision: A recidivism follow-up study within a sample of male young offenders. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66(5-B)*,2835.

Nunnally, J. C. (1978). *Psychometric Theory*. Mcgraw-Hill. NY.

Office of Juvenile Justice and Delinquency Prevention (2007). "OJJDP DMC, Reduction Best Practices Databases". Retrieved from: http://mpg.dsgonline.com/dmc_default.aspx

Onifade, E., Livsey, S., Davidson, W., Turke, G., Campbell, C., Mackey, M., & Malinowsky, J. (2007). *Predicting the Who, What and When of Recidivism with the Youth Level of Service Case Management Inventory*. Paper submitted for publication

Patterson, G. R., Debaryshe, B. D, & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*, 329-335.

Phillips S, Bloom B. In whose best interests? The impact of changing public policy on relatives caring for children with incarcerated parents. Child Welfare. 1998;77:531-541.

Pope, C. and W. Feyerherm. (1992). *Minorities and the Juvenile Justice System*. Rockville, MD: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention, Juvenile Justice Clearinghouse.

Quist, R. M. and Matshazi, D. G. (2000). The Child and Adolescent Functional Assessment Scale (CAFAS): A

dynamic predictor of juvenile recidivism. *Adolescence, 35(137)*, 181-192.

Rappaport, J. (1986). In praise of paradox. In E. Seidman and J. Rappaport. (eds) Redefining social problems. New York: Plenum Press

Richie BE. Challenges incarcerated women face as they return to their communities: findings from life history interviews. *Crime Delinquency*. 2001;47:368-389.

Risler, E., Sutphen, R., & Shields, J. (2000). Preliminary validation of the juvenile first offender risk assessment index. *Research on Social Work Practice,10*, 1, 111-126.

Schmidt, F., Hoge, R. D., and Gomes, L. (2005). Reliability and validity analyses of the youth level of service/case management inventory. *Criminal Justice and Behavior* 32, 3, 329-344.

Schrantz, D., & McElroy, J. (2000). *Reducing Racial Disparity in Criminal Justice: A Manual for Practitioners and Policmakers* .Sentencing Project. Washington, D.C.

Schumacher. M., & Kurz-Gwen, A. (2000). *The 8% solution: preventing serious, repeat juvenile crime*. Thousand Oaks: Sage Publications.

Schwalbe, C.S. (2007). Risk Assessment for juvenile justice: A meta-analysis. *Law and Human Behavior, 31*, 449-462.

Schwalbe, C. S., Fraser, M. W., Day, S. H., and Cooley, V. (2006). Classifying juvenile offenders according to risk of recidivism:Predictive validity, race/ethnicity, and gender. *Criminal Justice and Behavior, 33(3)*, 305-324.

Schwartz, I. M., Steketee, M. W. & Schneider,V. W. (1989). *The Incarceration of Girls: Paternalism or Crime Control?* Ann Arbor, MI: Center for the Study of Youth Policy

Seidman, E. (1987). Toward a framework for primary prevention research. In J.A. Steinberg & M.M. Silverman (Eds.), *Preventing mental disorders: A research perspective* (pp. 2-19). Washington, DC: U.S. Government Printing Office.

Sever, B. (2003). The minority population/police strength relationship: Exploring past research. *Criminal Justice Studies, 16,* 153-171.

Sharkey, J. D., Furlong, M. J., Jimerson, S. R., and O'brien, K. O. (2003). Evaluating the utility of a risk assessment to predict recidivism among male and female adolescents. *Education and Treatment of Children,* 26, 4, 467-494.

Sickmund, M (2004). Juveniles in corrections. *Juvenile Offenders and Victims, National Report Series Bulletin* (June).

Siegel, D. J. (1999). *The Developing Mind: Toward A Neurobiology of Interpersonal Experience.1-22, 276-82.*

Simourd, L., & Andrews, D. (1994). Correlations or delinquency: A look at gender differences. *Forum on Corrections Research, 6,* 26-31.

Smith, W. R. and Aloisi, M. F. (1999). Prediction of recidivism among second timers in the juvenile justice system: Efficiency in screening chronic offenders. *American Journal of Criminal Justice, 23(2),* 201-222.

Snyder, H. N., & Sickmund, M. (2006). *Juvenile Offenders and Victims: 2006 National Report.* Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Steffensmeier, D. J., & Steffensmeier, R. (1980). Trends in female delinquency: An examination of arrest, juvenile court, self-report, and field Data." *Criminology* 18: 62-85.

Streib, V. 1987. *Death Penalty for Juveniles.* Bloomington, IN: Indiana University Press.

Stucky, T.D. (2005). Local politics and police strength. Justice Quarterly, 22, 139-169.

Faye Taxman, James M Byrne, April Pattavina. (2005). Racial Disparity and the Legitimacy of the Criminal Justice System: Exploring Consequences for Deterrence. Journal of Health Care for the Poor and Underserved: Reducing HIV/AIDS and Criminal Justice Involvement in..., *16*(4), 57-77.

Thomas D Stucky (2005). Local Politics and Police Strength. *Justice Quarterly : JQ, 22*(2), 139-169.

Torbert, P. (1996). Juvenile Probation: The Workhorse of the Juvenile. *Justice System. Juvenile Justice Bulletin.* Office of Juvenile Justice and Delinquency Prevention

Turner, S., and Fain, T. (2003). Validation of the Los Angeles County probation department's risk and needs assessment instruments. Rand Institute Report.

VanVoorhis, P., & Brown, K. (1997). *Risk classification in the 1990s.*Washington, DC: National Institute of Corrections.

Von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology, 15,* 553-580.

Walsh, A. (2000). Behavior genetics and anomie strain theory. *Criminology, 38*(4).

Walters, G. D. (2006). Risk appraisal versus self report in the prediction of criminal justice outcomes. A meta-analysis. *Criminal Justice and Behavior.33(3),* 279-304.

R.L. Weisman, J.S. Lamberti, N. Price. (2004). Integrating Criminal Justice, Community Healthcare, and Support Services for Adults with Severe Mental Disorders. *Psychiatric Quarterly, 75*(1), 71-85.

Wiebush, R.C., Baird, C., Krisberg, B., and Onek, D. (1995). Risk assessment and classification for serious, violent, and chronic juvenile offenders. In

J.C. Howell, B. Krisberg, J.D. Hawkins, and J.J. Wilson (Eds.), *A Sourcebook: Serious, Violent, and Chronic Juvenile Offenders*. Thousand Oaks, CA: Sage Publications.

Wilson, J. Q. (1968). Varieties of police behavior: The management of law and order in eight communities. Cambridge, MA: Harvard University Press.

Wolfgang, M., Figlio, R. M., and Sellin, T. (1972). Delinquency in a birth cohort. Chicago. University of Chicago Press.

Wordes, M., Bynum, T. S., & Corley, C. Locking Up Youth: The Impact of Race on Detention Decisions. (1994) *Journal of Research in Crime and Delinquency*, 31, 149-165.