2007

Predicting Repeated Sexual Behavior Problems among Youth: A Hierarchically Optimal Classification Tree Analysis

David J. Morgan

LOYOLA UNIVERSITY CHICAGO

PREDICTING REPEATED SEXUAL BEHAVIOR PROBLEMS AMONG YOUTH: A HIERARCHICALLY OPTIMAL CLASSIFICATION TREE ANALYSIS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

PROGRAM IN CLINICAL PSYCHOLOGY

BY

DAVID J. MORGAN

CHICAGO, ILLINOIS

AUGUST 2007

ACKNOWLEDGEMENTS

There are so many people in my graduate career who have been invaluable in their support, guidance, and advice. While I cannot possibly name them all here, a few deserve special note. Scott Leon, PhD, my chair, rescued me from the depths of dissertation limbo and presented me with an incredible opportunity that I could never replicate. Scott's support and insight in no small part made this dissertation what it is here, and it has been a pleasure being his first dissertation student. Fred Bryant, PhD, has been far more than my dissertation statistics advisor. Fred has been my advisor on life, teaching, and savoring, and will continue to be for as long as he'll have me as his student. Steve Spaccarelli, PhD, provided the data set for this project and met with me countless times to not only catch me up on the dataset, but to provide feedback on my work from his perspective as a clinician working with youth with SBP. Stacy Frazier, PhD, has been never-ending in her support for me and didn't hesitate to say yes to being on my committee even with her packed schedule at the Institute for Juvenile Research. Pat Rupert, PhD, was my first research advisor and taught me how to really read an article, and her perspective from someone outside the child field made this a more accessible piece of work.

To my mother, Marie Sadler Morgan, who always believed the PhD was possible.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

This study examined characteristics of youth with sexual behavior problems (*N* = 251) followed over approximately two years and applied Classification Tree Analysis (CTA) via Optimal Data Analysis (ODA) to produce a hierarchically optimal classification model to predict recurrence of SBP over time. This study was designed to address problems with clinical judgment leading to inaccurate diagnosis and in some cases, incorrect treatment or placement of the youth by ascertaining if an actuarial model designed to optimize classification accuracy might be more helpful in directing diagnosis and treatment for these youth. Results suggest initial event severity influences the potential for problem recurrence, with lower-severity (i.e., less invasive) SBP actually being more likely to recur than higher-severity SBP. The factor of Narcissism by self-report on the Antisocial Process Screening Device and the placement at time of initial event also influenced the prediction of whether or not SBP would recur. Classification performance was fair but lacks the specificity necessary for clinical usage. The classification accuracy of the model could be improved significantly in future studies with a larger sample size and a longer time of follow-up.

CHAPTER 1

INTRODUCTION

A great deal of time, energy, and money have been devoted to differentiating between juveniles who are most likely to develop chronic behavioral problems versus those who present little or lesser risk of long-term problem behavior. As Lynam (1996) eloquently stated, "The golden grail of high-risk research has been the identification of the minority of children who are most likely to persist in their antisocial behavior from among the multitude of children who engage in some antisocial acts" (p. 211). In the case of youth who exhibit sexual behavior problems (SBP; Silovsky & Niec, 2002), this issue has become especially salient recently, as evidenced by an increase in studies conducted to ascertain the characteristics of these juveniles, particularly characteristics related to chronicity (Prentky et al., 1989; Rasmussen, 1999; Spaccarelli, Bowden, Coatsworth, & Kim, 1997). The current study represents an attempt to better identify the exact pathways to repeated sexual behavior problems among sexually aggressive youth using new techniques. In this manuscript, "youth" is used to refer to children and adolescents between the ages of 10 and 18.

The Gravity of the Problem

Sexual behavior problems (SBP) amongst youth is an issue of considerable import to mental health professionals, parents, and society-at-large. One estimate of the cost of sexual abuse of youth alone (not including adult victims) places the cost of victimization at $99,000 (in 1993 dollars) per incident,

including attempts (Miller, Cohen, & Wiersema, 1996). The cost to society for treating these offenders is similarly high: In 2003, the Illinois Department of Children and Family Services received an annual budget of 1.4 billion dollars (Illinois Department of Children and Family Services, 2003). While no specific statistics are available, it is reasonable to speculate that a sizable portion of this budget is spent on youth who have been the victims or perpetrators of sexual aggression.

Definitions. Sexual behaviors are defined as problematic when they (a) occur at greater frequency or at a much earlier age than would be developmentally expected; (b) interfere with the youth's development; (c) occur with the use of coercion, intimidation, or force; (d) are associated with emotional distress (for the youth with SBP or other youth involved); and/or (e) reoccur in secrecy after intervention by caregivers (Silovsky & Niec, 2002). While Ryan (1997a) suggests that SBP (or sexually abusive behavior as Ryan terms the more intrusive behaviors) are any sexual interactions perpetrated against the victim's will, without consent, or in an aggressive, exploitative, manipulative, or threatening manner, this definition presents several quandaries for defining the population: a) When listing these definitional components, it is not clear if any or all of these are sufficient by themselves to fulfill the definitional requirement by themselves, b) some acts are not necessarily aggressive, exploitative, manipulative, or threatening (although this many terms appears to be an attempt to capture any possible event at the expense of specificity), and are not necessarily against the other person's will or without consent because many

young persons cannot give consent, and may be willing to engage in behavior that may be harmful as they do not understand the consequences, and c) the terms aggressive, exploitative, manipulative, or threatening require a large degree of subjective speculation as to the individual's intent, which is sometimes not available for analysis. There is also the issue of the intent possibly being more important than the outcome in this definition. For example, if a youth intended to be exploitative but was not successful, would this be SBP? Although the Silovsky and Niec (2002) definition avoids many of the pitfalls of the Ryan (1997a) definition, it still presents some challenges in application. For example, some behaviors (e.g., rape) might not occur very frequently. More troublesome is defining the developmental onset of a behavior that is problematic at any age. It might be the case that like conduct disorder, earlier onset is worse (Broidy et al., 2003). While these inclusive definitions might help identify juveniles at-risk or who are not yet powerful or skilled enough to carry out harmful acts, they are sometimes at odds with the legal system definitions, which carry a great deal of weight. In this manuscript, SBP will be used with the understanding that it refers to a subset of SBP--those behaviors that meet legal criteria for sex offenses. However, this usage is not meant to imply that offenses not meeting legal standards for prosecution are any less worrisome.

Legal standards of sexual aggression are often the sole arbiters of sexually aggressive juveniles' residential and therapeutic futures. Receiving a SBP label can greatly diminish a youth's chances of being placed in a foster family, let alone a well-matched family, because it seems likely many foster parents are wary of

juveniles with the SBP label (Swan, 1997). Furthermore, courts or child protective services can order therapeutic placement via residential treatment, or at least regular treatment sessions with a mental health provider.

The costs of inaccuracy

Mistakes in predicting re-offending come at a high cost. As mentioned earlier, there is significant cost to society when sexual behavior problems occur. However, lack of specificity in these decisions has similarly high costs. Critics of labeling dating back to Szasz (1963) and Rosenhan (1973) have pointed out that once applied, labels tend to "stick" and result in promoting a role of illness amongst people who share a particular negative mental status label. This occurs in a number of unfortunate ways among juveniles who exhibit SBP. Juveniles who have exhibited SBP and are identified as such are usually involved in the judicial or child protective services system in some way, resulting in stigma from the community and peers who discover this involvement. Juveniles who are in foster care may find their foster status endangered, and may have to be reassigned to a new family as many foster parents may be reluctant to harbor a youth who may exhibit sexual behavior problems while in their care. For juveniles at risk to exhibit further SBP, this potential instability of care that results from the very incident that warranted the care in the first place poses a great threat to their mental health via further instability of contingencies likely to manage the SBP.

This instability of placements often appears to be presumed in the literature, based on the logic that youth who require substantially more services

and have the potential to violate other youth in the home would be harder to place. Edmond, Auslander, Elze, McMillen, and Thompson (2002) found this to be the case among a sample of sexually abused girls in foster care, many of whom exhibited sexual behavior problems. Girls with this type of history had significantly more placements, and were more likely to be living in congregate-type placements as opposed to family-type placements. Barth and Berry (1989) also found that youth who are in need of more intensive and structured care are more likely to be placed in congregate living settings, indicating they are harder to place in traditional family-type foster homes. This may especially true for girls who exhibit SBP, because there is growing concern among caregivers that these girls may lodge false accusations of abuse against the caregivers themselves (Swan, 1997). Thus, the classification of youth as individuals with SBP can carry heavy consequences, and this appears to be a case where maximal prediction accuracy is required to prevent future harm and deleterious environmental experiences. Increased prediction accuracy would also lead to fewer unnecessary applications of more restrictive treatments, such as residential care.

The Olmstead Act, part of the Americans with Disabilities Act (ADA), mandates agencies and care providers to deliver services in "the most integrated setting appropriate to the needs of qualified individuals with disabilities" (Centers for Medicare and Medicaid Services, 2004). This particular clause, often called the "least restrictive environment" clause, is a crucial element of sound mental health policy for youth, including youth with SBP. Youth with SBP can be unnecessarily stigmatized by treatment options that are more isolating and

intensive than necessary as a result of inaccurate classification. As SBP tend to be more dramatic or shocking than more conventional forms of disruptive behavior, care providers may be wary to risk allowing less restrictive treatments if the risk of such assignment is possible recurrence of the behavior. Although we have impressive tools to evaluate construct validity of instruments, decisional cut points still depend upon a qualitative assessment of the costs of false negatives versus false positives (Rice & Harris, 1995). Many clinicians and caregivers may feel that the risk of failing to identify SBP is more costly than the risk of placing a youth in an overly restrictive environment. In the current litigious climate, this state of affairs is hardly surprising. However, if the ADA mandates placement in the least restrictive environment, then the logical outgrowth of that mandate is a requirement that youth be accurately assessed to correctly place them in the least restrictive environment for their needs. The myriad outcomes possible for youth with varying degrees of SBP thus has become a policy issue in need of considerable attention.

Many of the above outcomes (e.g., residential care) often hinge on whether or not a youth enters the legal system for the offense in question. Thus, legal definitions of SBP are at least useful, if not as valid as some would like for research purposes (especially universal and targeted prevention research). SBP by legal reporting standards includes forcible rape, forcible sodomy, sexual assault with an object, and forcible fondling (Snyder, 2000). While this definition may be considered too inclusive by some, it should be noted that many incidents of sexual assault go unreported, as suggested when examining data from past

surveys such as the National Crime Victimization Survey (NCVS; Snyder, 2000). Gathering data on the prevalence of SBP has presented a number of challenges to researchers in the past (Snyder, 2000). Differing methods of surveillance have constituted a large portion of these challenges in ascertaining prevalence. Specifically, the NCVS, which assesses victimization regardless of reporting of the act to authorities, and the Uniform Crime Reporting Program (UCR), which is based entirely on reported offenses, have not used similar methodology to gather prevalence information. The UCR before roughly 1991 did not contain any information about offenses other than forcible rape. After 1991, however, states began reporting using a new system under the UCR called the National Incident-Based Reporting System, or NIBRS. This new system provides a great deal more information about victims, offenders, and the offenses committed than before. However, it should be noted that a common factor amongst past and current definitions of SBP amongst youth in the legal system has been coercion; in other words, the victim of the act was psychologically or physically coerced.

Twelve states reported data to the NIBRS between 1991 and 1996, using the previously mentioned four categories. One recent study of the 1991-1996 NIBRS data suggests that SBP amongst youth is not as rare as once assumed. It should be noted that in these data, estimates of perpetrator age are complicated by the fact that they are based on victims' estimates of the age of the perpetrator. However, it seems likely that most victims can differentiate between juveniles and adults in general, with probably more difficulty at the boundaries (e.g., a

victim attempting to guess whether a perpetrator was 17 or 19). One conservative estimate of the percentage of sex crimes committed by juveniles suggests that approximately 23% of all reported, charged, and convicted sexual assaults (encompassing forcible rape, forcible sodomy, sexual assault with object, and forcible fondling) fall under the category of SBP (Snyder, 2000), although self-reports of SBP have placed this proportion between 56 to 57 percent of all reported sexual assaults on boys (Rogers & Tremaine, 1984; Showers, Farber, Joseph, Oshins, & Johnson, 1983), suggesting that not all incidents are reported and prosecuted. In 1996, 307,000 incidents of forcible rape and other sexual assault occurred in the United States among victims age 12 or above, with this number representing data from only twelve states.

There are no national-level data available on the prevalence of sex crimes amongst victims under age 12. However, it is safe to assume that the number of incidents would be far higher if these younger children were included. Several studies have found that among juvenile perpetrators, approximately 65% prefer victims younger than 12 years of age. Adults tend to offend against this age range of victims in similar proportion (Snyder, 2000). Prevalence can be estimated assuming adults offend equally often against children under 12 and victims over 12 as the lower confidence limit, and assuming the proportions are .7 and .3 for the same victims, respectively, for the upper confidence limit. Using these numbers, it is estimated that juveniles perpetrated a total number of sexual offenses between 151,686 and 226,677 in 1996, with between 53,090 and 79,337 being offenses against victims 12 and older, and 98,596 and 147,340 being

against victims 11 and younger. Once again, it merits a reminder that these estimates are extremely conservative as they represent data from only 12 states in the United States.

Although there are large numbers of juveniles who have exhibited SBP, the number of acts committed in a year may be somewhat misleading. It should be noted that in most samples of antisocial boys, approximately 5% of the serious offenders in a cohort account for more than 50% of violent crime in that group (Skilling, Quinsey, & Craig, 2001). This may very well be the case with juveniles with SBP. This claim is difficult to evaluate because most juveniles with SBP exhibited a number of SBP before being identified for the first time (Ryan, 1997b). Also, the degree of severity of behavior necessary to warrant legal attention has changed over the years; 20 years ago and prior, a number of courts only considered forcible rape in the "serious" category, thus warranting judicial attention and intervention (National Council of Juvenile and Family Court Judges, 1984). Thus, it is not clear that all offenses were treated the same way and received the same degree of penalty and/or treatment. Finally, it is likely that detection methods (e.g., parental monitoring) for these behaviors are not perfectly sensitive, suggesting that even after the increased scrutiny after a documented offense (which, as noted above, may differ from system to system), some offenders may re-offend without being caught. In contrast, Dawes (1996) disputes claims that small proportions of groups commit a large number of the acts that determine group membership (e.g., serial rapists) on the grounds that those who commit offenses multiple times are under more scrutiny and thus

simply get caught more. However, it is difficult to believe that there are no different types of offenders (e.g., serious/chronic vs. minor), or that all offenders are repeat offenders.

Regardless of the actual prevalence, even the lowest known prevalence in the literature suggests that SBP amongst juveniles merits more attention than is currently accorded for a number of reasons. First, as noted above, the costs of SBP to society and victims is extraordinarily high. Second, the cost of treating youth with SBP is also high, although far cheaper than incarceration (Lane Council of Governments, 2003). While no one has estimated the lifetime cost of treatment for this problem (as the lifetime trajectories of various offending patterns remains unknown), it seems likely that untreated or ineffectively treated juveniles exhibiting SBP will incur significant costs for society via utilization of government-funded mental health resources and possible incarceration costs, not to mention the cost to the offender and the offender's family in both financial and emotional terms.

Finally, it is often overlooked that life in the community-at-large is extremely difficult for sex offenders with child victims, especially as adults, given the high probability of being caught, being registered and identified in the community (if convicted as adults), and living with the innumerable consequences of the devastating actions of sexual violation of children, including guilt, anxiety, and the high probability of aggressive retaliatory action on the part of the community. There has been a recent explosion of interest in child molestation in the media, with numerous internet and television media providers

conducting their own investigations to expose child molesters in the community, including organizations such as perverted-justice.com (2006), and Dateline NBC with their popular "To Catch a Predator" series (2006). Of course, there has always been negative attention and tragic consequences with respect to child molestation. As Goddard & Saunders (2001) point out, the media have attempted to "out" child molesters for over a century. In one instance of apparent mistaken identity, Bowcott and Clouston (1997) describe a girl's death in a fire after the house in which she was staying was ignited by individuals who thought it was occupied by a child molester. These incidents merely underscore the public perception that child molesters (to name only one type of sex offender) are incapable of change. This is a perception that can be refuted only through research that directly answers two crucial questions: 1) At what rate do sex offenders re-offend, and 2) What variables predict with reasonable sensitivity and specificity which of these sex offenders will exhibit chronic sexual behavior problems?

Currently, the public perception seems to focus only on sensitivity at the considerable expense of specificity. Specificity of prediction is crucial to the aim of changing public sentiment regarding sex offenders who exhibit little risk of re-offending. Specificity of prediction, if powerful enough to change public perception of the non-chronic offenders, may offer these individuals better chances for recovery and integration into the communities in which they live. The costs of poor prediction are considerable indeed if we doom a large

proportion of recovering offenders to fail via indiscriminate labeling as severe risks.

Predicting Dangerousness and Recidivism

In the past, clinicians have often used clinical prediction to determine dangerousness or risk of committing an undesirable act again (David Faust, personal communication, September 28, 2004). Clinical prediction consists of judgments based on clinical perceptions of the client or offender, sometimes termed "clinical impressions," "professional judgment," or "expert opinion," to name only a few (Garb, 1998). In child psychology and psychiatry practice parameters for SBP, "There are no specific empirical measures or psychometric tests that can identify, diagnose, or classify sexual abusers" (Shaw & The Workgroup on Quality Issues, 1999, p. 3). This is especially unfortunate given that in the same practice parameters document, the recommendation is presented: "The clinician should be prepared to address. . .the risk of repeating the sexually aggressive behavior." Recommendations such as this imply that clinicians should currently use data gleaned from interviews and instruments in some piecemeal way, if at all. It seems likely, then, that many clinicians are "flying by the seats of their pants" with respect to predicting the risk of future SBP amongst youth. This practice is troubling because of the problems associated with clinical judgment.

Many clinicians and researchers alike appear to miss the point in the arguments against utilizing clinical judgment in their decision-making processes, illustrated by studies that demonstrate fair to good levels of accuracy for clinical

judgment (e.g., Gardner, Lidz, Mulvey, & Shaw, 1999). The point, as argued by Meehl in 1954, is that mechanical prediction will almost without exception be *superior*. While in some cases clinical prediction may equal mechanical prediction in accuracy, this is rarely true when the clinical predictions are based on clinical interview data (Grove, Zald, Lebow, Snitz, & Nelson, 2000). It is also clearly the case that mechanical prediction methods present fewer opportunities for bias to enter the diagnostic process (Garb, 1998). Also, as the American Psychological Association mandates via the ethics code (American Psychological Association, 2002), psychologists should carefully evaluate the validity of their assessments and tests, with the apparent intention being to improve diagnostic accuracy and thus treatment. Part of this careful evaluation should be a frank consideration of the dangers of using clinical prediction to make diagnostic or treatment decisions related to estimates of who will continue to exhibit problem behaviors and who will not. This consideration is further complicated by the concern that base rates of SBP may be too low to accurately predict (Prentky et al., 2000).

While base rates of SBP in the legal system have sometimes been quite low in studies evaluating possible factors related to chronic problems (Prentky et al., 2000), there is reason to believe that the recurrence figures may be much higher than some suggest. For example, while Prentky et al. reported a 11% recurrence rate in their sample, other researchers have reported rates ranging from 37% to 75% (Awad & Saunders, 1991; Becker, Cunningham-Rouleau, & Kaplan, 1986; Fehrenbach, Smith, Monastersky, & Deisher, 1986). As Prentky et al. note, these differences in reported recurrence rates are a hallmark of the

literature, and suggest that methodological differences may be playing at least a partial role in these drastic reported differences.

It seems likely persistent SBP may be a serious problem, indeed given the inherent bias in this type of research toward overlooking incidences of recurrence due to lack of sensitive measurement (i.e., individuals are not always caught),. More specifically, if one considers the above estimated prevalence numbers, even a 10% recurrence rate poses significant risk of psychological, physical, and financial harm to a large number of people. Using even a 10% recurrence rate and an estimate that of the 151,000 (extremely) conservatively estimated SBP incidents, and if each youth with SBP exhibits an average of five of these problems, we arrive at a calculated 3,200 youth who are likely to exhibit repeat SBP. While these arithmetic exercises are highly speculative, they are nonetheless illustrative: 3,200 disenfranchised youth who pose significant risk to others represents a serious problem by any governmental yardstick.

Problems associated with the current methods of prediction

While clinical judgment is probably the least valid way to predict future events (Grove, Zald, Lebow, Snitz, & Nelson, 2000), current actuarial methods for predicting recurrence of SBP among juveniles are lacking in predictive power. The best predictor of behavior is behavioral history, and this is not an exception in the field of SBP research (Grubin, 1999). However, reports of SBP are naturally history-based. History-based methods of actuarial prediction are not particularly useful in predicting future problems because *all* of the individuals have already exhibited the problem of interest, and the lack of variability in

history amongst some populations would mean that every youth would receive exactly the same risk rating as a result of similar history, which hardly seems useful. Also, the history used to predict behavior is static until another offense happens, or a good deal of time has elapsed, making these types of prediction strategies relatively insensitive to change that could come as a result of treatment. Of all these youth exhibiting serious SBP, society and governments are searching for strategies to accurately predict who will continue exhibit problems despite being already identified, and who is significantly less likely to exhibit future SBP. These new models may help predict the intensity of treatment required to prevent sex offending, as well.

Attempts to predict SBP

Until recently, there had been few attempts to predict recurrence of SBP amongst juveniles with respect to actuarial methods. Past attempts to classify youth who have exhibited SBP have focused on a multitude of variables measured in myriad ways, from sex roles, attitudes, and peer norms (Ageton, 1983), to penile plethysmography to assess sexual deviance (Seto, Lalumière, & Blanchard, 2000). A number of factors have been added to various predictive models to increase predictive power, such as psychopathy (Frick, O'Brien, Wootton, & McBurnett, 1994; Hare, 1991), as well as victim age and sex in relation to age of the offending youth (Prentky et al., 2000; Snyder, 2000), sometimes used to suggest typological distinctions by offender/victim age match as presented by Prentky et al (2000). Interestingly, from the earliest studies to

more recent examinations of SBP amongst juveniles, several common findings have emerged.

Delinquency The first commonality of the studies of SBP amongst juveniles is that delinquency is the single most important predictor of SBP recurrence (Ageton, 1983; Awad & Saunders, 1991; Awad, Saunders, & Levene, 1984; Becker, Kaplan, Cunningham-Rathner, & Kavoussi, 1986; Fehrenbach et al., 1986; Spaccarelli et al., 1997). How delinquency is related to SBP recurrence may require some explanation from related bodies of research: specifically, conduct problems and psychopathy. Conduct Disorder (CD) is a specific term for a syndrome of delinquency in the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text revision, American Psychiatric Association, 2000). While estimates vary, it has long been established that some youth with conduct problems exhibit similar problems in adulthood, and some do not. Estimates of CD to adult antisocial personality disorder (APD) and proportions of those with APD who exhibit psychopathy suggest that roughly a quarter of youth who exhibit CD go on to exhibit an especially pervasive pattern of antisocial behavior in adulthood, commonly referred to as psychopathy (Broidy et al., 2003; Frick, Bodin, & Barry, 2000; Kazdin, 1995; Lynam, 1996).

Frick (2000) suggests that there are three fundamental constructs that constitute the syndrome of psychopathy: impulsivity, narcissism, and callous-unemotional traits. Frick points out that it is hypothesized that many constructs in psychology exhibit multifinality, with the same construct potentially leading to a number of different disorders. Thus, impulsivity is not a hallmark of

psychopathy in and of itself; for example, for quite some time, Attention Deficit Hyperactivity Disorder (ADHD) has shared the construct of impulsivity with other disorders (American Psychiatric Association, 2000). However, the unique combination of the above three traits have demonstrated reasonable validity as a measure of psychopathy (Frick, Bodin, & Barry, 2000). The same reasoning may very well apply to the concept of repeat (and thus likely continuing into adulthood) SBP among youth. More specifically, while many behaviors that can constitute delinquency are related to CD, they may also be related to serious SBP among juveniles. However, these delinquent behaviors shared with CD are not by themselves sufficient to lead to a SBP outcome. Instead, these delinquent behaviors need to be accompanied by other constructs deemed integral to the concept of persistent youth SBP.

Sexual deviance Aside from the likely inclusion of delinquency, however, other constructs must be integral to a prediction algorithm of SBP. One posited construct is sexual deviance, especially sexual interest in children. These interests, commonly termed pedophilic interests, have been demonstrated to discriminate between offenders against children and those with sex interests in individuals of similar age when assessed via penile plethysmography (Seto, Lalumière, & Blanchard, 2000). This line of reasoning operates upon the presumption that pedophilic interests predict persistent offending patterns for juveniles who have SBP. This seems to be the case; Hanson and Bussière (1998) found in their meta-analysis of 61 follow-up studies of repeat sex offenders that the strongest single predictor of sexual recidivism (weighted average $r = .32$, 95% C.I. = .29 - .35)

was deviant sexual interest, often assessed by phallometric assessment (plethysmography) or by record review of past offenses. Another large predictor of recidivism in this meta-analysis was the presence of antisocial personality disorder (APD), the logical extension of delinquency into adulthood. While not all individuals with APD exhibit psychopathy, a substantial proportion of this group do exhibit this syndrome (Frick et al., 1994). Perhaps if the studies in the meta-analysis only included individuals with APD who met psychopathy criteria, the predictive relationship between APD and recidivism would have been even stronger. This pattern of evidence strongly suggests that both psychopathy and deviant sexual interest are reasonable variables to include in a model to predict repeat SBP among juveniles suggestive of chronic SBP.

There are several caveats to the deviant sexual interest-recidivism relationship, however. First, while plethysmography has demonstrated modest predictive validity with respect to recidivism, the same is not true of self- or therapist report of deviant sexual interest (Kahn & Chambers, 1991; Prentky & Knight, 1993; Prentky et al., 2000). The second caveat concerns the age at which plethysmography becomes an accurate reflection of deviant sexual interest. Amongst adult sex offenders, any sexual interest in children is clearly deviant. However, it is not clear from a developmental standpoint a) when youth begin thinking about sex, or whom they think about, and b) when sexual interests solidify to the point where they are substantially predictive of sexual interests into adolescence and adulthood.

Developmental theories that do address sexual development seem to posit that mature sexual interest comes around early adolescence (Grotevant, 1997), and while the field of sex preference development is still in its infancy (which should be differentiated from gender identity development, which while comprehensive, usually only addresses sexuality tangentially if at all), it seems likely that sexual preferences are hardly solidified early in adolescence, except perhaps in cases of early undesirable sexual contact that shape subsequent sexual behavior. So, it is not altogether clear that pedophilic preference is synonymous with sexual deviance *throughout* the developmental span of childhood and adolescence. Finally, it should be noted that penile plethysmography is difficult to administer on a large scale because it requires a structured lab environment and involves exposing offenders to stimuli they should be avoiding, an ethically questionable practice at best (Worling & Curwen, 2000).

The state of the art in actuarial prediction of SBP

Prentky et al. (2000) and Gretton et al. (2001) have conducted the most comprehensive evaluations of actuarial methods to predict SBP recurrence amongst youth to date, and thus, these studies deserve special attention. While both studies were archival in method, past research has suggested that reliable and valid assessments can be made from file information alone, provided the file information is reasonably comprehensive (Gretton, 1998; Rice & Harris, 1997). Gretton et al. attempted to predict SBP using the Hare Psychopathy Checklist: Youth Version (PCL-YV; Forth, Kosson, & Hare, 2003), while Prentky et al. used the Adolescent Psychopathy Taxon Scale (Harris, Rice, & Quinsey, 1994).

Both studies used some measure of deviant sexual arousal. Prentky et al. (2000) used deviant sexual history and "evidence of sexual preoccupation" (p. 79) to code for sexual deviance, while Gretton et al. (2001) used plethysmographic evidence for the same purpose. While the studies are similar in most respects, it should be noted that the Prentky et al. study suffers from the problems associated with using historical measures of deviance (insensitivity to change), as well as questionable validity from this type of deviance data.

Prentky et al. seem to propose typologies of juvenile sex offenders in their study by classifying their sample into rationally derived categories, including: Child Molester, Rapist, Sexually Reactive Child, Fondler, Paraphilic Offender, and Unclassifiable. Over two-thirds of their sample fell into the category of "Child Molester", defined as a youth with all victims age 11 or younger, with the offender at least five years older than the oldest victim. As with many typological classifications, such as those in the DSM-IV (4th ed., text revision, American Psychiatric Association, 2000), there is the problem of the arbitrary nature of the criteria for inclusion in the category. More specifically, why should the age gap between offender and victim be five years, and not six? Is there any theoretical or at least empirical explanation for why this particular age gap represents a class of some kind? These questions are not answered by the Prentky et al. study. There is also the problem that these categories are rationally derived with no theoretical basis, and thus the results are likely sample-dependent. More importantly, without any theoretical basis for these categories, it becomes very difficult to

design research that is able to examine the etiology of such a category of behavior when the category itself is so vaguely defined.

Prentky et al. (2000) encountered difficulty obtaining significant results because their sample exhibited such a low rate of recidivism, which the authors attribute to the short follow-up period (Prentky, Lee, Knight, Cerce, 1997). However, the factor that appeared to best differentiate individuals with chronic SBP from those who did not exhibit SBP was the antisocial/delinquent factor, while their sexual deviance factor did not seem to be helpful in predicting problem recurrence. Interestingly, after categorizing their sample by the previously mentioned offender types (e.g., Child Molester), the authors do not report whether the various hypothesized factors (e.g., antisocial/delinquent) predicted recurrence by offender type. This would seem to be an especially important question given that it appears that the general category of child molesters seems to be the most chronic in SBP, as discussed previously (Hanson and Bussière, 1998). The authors were probably unable to conduct these analyses due to the low rate of recurrence in their sample overall, making it impossible to further divide the already small group of individuals with chronic problems into categories of offense. This information would be helpful in assessing whether or not these more serious problems can indeed be identified on the basis of the two primary variables of psychopathy and deviant sexual interest.

Gretton et al. (2001), on the other hand, found similar results with respect to psychopathy, with the finding that high levels of psychopathy as measured via the PCL:YV were associated with higher levels of recurrence, but

the results were significant, probably as a result of the larger sample size relative to the Prentky et al. study. High levels of sexual deviance as measured via penile plethysmography were also associated with recurrence risk, in contrast to the Prentky et al. study, which used history-based measures of sexual deviance. However, the authors note that these factors, while seemingly powerful predictors of recurrence, do not exhibit a great deal of specificity with respect to sexual vs. nonsexual offense recurrence, an odd finding given that sexual deviance and psychopathy *together* predicted recurrence far better than just psychopathy alone. The authors speculate that it may be that their sample was comprised of so-called "versatile offenders" who are not so much sex offenders per se as they are broad criminal offenders who are diverse in their offense profiles. This would seem to be an argument that offenders who commit sex crimes in addition to regular crimes are qualitatively different from those youth who commit primarily sex crimes alone.. Youth who commit sex crimes alone appear to be very rare, as delinquency (which is associated with both) is obviously predictive of trouble with the law other than just sex offenses. Another explanation for these odd results could be that the Gretton et al. sample may very well have been comprised of youth who were sexually recidivating in a manner predicted by the high psychopathy and deviance scores, but that as a result of local variability in law enforcement efforts, varying local standards, and attention to sex offenses, these offenders were more likely to be caught for non-sex offenses. Gretton et al. note that in many cases, it was not clear from the records whether the subsequent offenses were sexual in nature, given the lack of

police records for their study. A replication of the Gretton et al. study with a more diverse sample from different geographic regions (with similar enforcement) as well as more specific offense data would shed some light on this question.

The authors also entertain the notion as noted previously that perhaps sexual deviance is less important amongst juveniles (perhaps as a function of early and incomplete development of sexual interests) than the psychopathy they may exhibit, which suggests that any SBP treatment program must treat the pervasive pattern of psychopathy in addition to any SBP problems to effect lasting change.

Theoretical models of chronic juvenile sexual behavior problems

To date, no actuarial models of repeat or chronic SBP have been based upon any theoretical foundation. However, several theories have been advanced in the literature. Ryan (1997a) developed a model of juvenile sex offending that is based largely on the premise that sexually abused youth become abusers themselves. While this may explain the etiology and variability in outcome for a small proportion of juvenile sex offenders, the majority of juvenile sex offenders report having experienced physical and emotional abuse at the hands of caregivers, but not necessarily sexual abuse (Knight & Prentky, 1993; Prentky et al., 2000; Rasmussen, 1999). The large number of juvenile sex offenders who apparently have not suffered sexual abuse calls into question an etiological model of chronic SBP that presupposes sexual victimization in the history of the offender.

In summary, there are only a few variables that appear to consistently relate to repeat SBP amongst youth in a systematic way. One of these variables is psychopathy, which is a complex phenomenon made up of multiple constructs. Of the individual constructs comprising psychopathy, impulsivity has also been suggested as a contributor to continued SBP (Gretton et al, 2001). As Frick (2000) has pointed out, impulsivity is a phenomenon that clearly exhibits multifinality. With this argument in mind, it seems likely that impulsivity amongst those with sexual behavior problems can present a problem even in the absence of psychopathy, but this potential hypothesis awaits empirical testing. Finally, the measurement of sexual deviance seems to be a construct that deserves considerable attention in a field concerned with sexual behavior that is clearly deviant. However, as noted, the measurement of sexual deviance is fraught with challenges as a result of ethical concerns about plethysmography and the apparent difficulties in obtaining valid self- or therapist reports of deviance. The constructs of psychopathy, impulsivity, and sexual deviance seem central to the assessment and treatment of juvenile sex offenders, especially with respect to prediction of re-offending.

Theory and the current study. Grubin (1999) has suggested a model for juvenile sex offending based on the factors already discussed, as well as some new constructs that have received little attention. Grubin posits that three constructs account for the outcome of sex offending: 1) deviant sexual fantasy, 2) interpersonal/social and emotion isolation, and 3) deficits in either the cognitive or affective components of empathy. Grubin proposes this model can

provide a starting point for examining the origins and maintenance of sexual offending. Grubin emphasizes that some of these constructs need more exploration and are open to some interpretation, especially with regard to isolation, which Grubin presents as a construct in need of research, with the primary questions relating to whether isolation is a preference, or an indicator of the underlying disorder. Grubin appears to be presenting this model as a potential etiological study of adult offenders, but there is not any specific attention to developmental processes in the etiological proposal. This inevitably begs the question of the representations of isolation and other constructs among juveniles.

While the literature on the importance of attachment is inconsistent with regard to its power to affect later development (Schneider, Atkinson, & Tardiff, 2001), this construct may have particular relevance for certain populations such as foster youth because of the common instability of their placements (Edmond et al., 2002). Changes in foster parents can often mean changes in schools and peers, so attachment to parent figures may represent a protective or risk factor, with positive attachment representing low isolation (and perhaps might thus serve a protective role), while low attachment would be construed as high isolation and thus may place particular offenders at greater risk for re-offending. The ecological considerations do not stop here, as there are often predictable reasons for low or high levels of attachment, such as the relative level of parenting skill exhibited by caregivers. Thus, it would probably be helpful to evaluate isolation from family in the contexts of attachment and parenting

practices to gain a better understanding of the youth's microsystem (Bronfenbrenner, 1977). Ideally, the concept of isolation should integrate other ecological levels, such as peers, school, and other community structures.

Grubin discusses deviant fantasies in terms of sadistic fantasies, but it seems many sex offenders may not be thought of as sadists, as they actually fantasize about loving relationships with their victims (Ryan, 1997a). Thus, it seems there is no reason why the fantasy component of the model cannot be expanded to include deviant thinking and behavior to better represent the construct of sexual deviance, which, as pointed out previously, seems to be an important component of sex offending. While penile plethysmography could be considered an excellent standard in light of the more recent research already discussed, the concerns regarding the feasibility and ethicality of the practice suggest research should look for acceptable alternative measures of sexual deviance that involve a lower level of invasiveness. Ideally, this construct of deviance should be measured in such a way that it captures a number of aspects of deviance, such as past behavior as well as deviant cognitions.

Empathy is also a reasonable construct to include in a model of sexual offending, given the devastating nature of the act, and the recommendation that treatments include empathy-building as a goal (Lane Council of Governments, 2003). Empathy has been assessed for quite some time as part of the package that represents psychopathy, with lack of empathy being an important part of the entire package. As common measures of psychopathy (e.g., Hare, 1991; Frick et al, 1994) typically provide both independent factor scores and total scores, it is

possible to evaluate the level of psychopathy as well as the presence of empathy by itself.

Impulsivity, also considered an important component of psychopathy (Frick, 2000), is, as noted previously, a construct of multifinality, meaning impulsivity is not sufficient for any one disorder in and of itself, and most important, the presence of impulsivity can lead to a number of different outcomes. Impulsivity, then, like empathy, can probably be best studied in the manner proposed for empathy above. Impulsivity may be especially important for particular types of juvenile offenders; it seems that that there is a far higher level of predatory (i.e., psychopathic) intent related to persistent rape behavior than, for example, exhibitionism, which by definition involves no particular victim selection, and the experience does not involve any apparent intent to harm. In contrast, depending on the attacker-victim relationship, it is likely rape can be much more about violent assertion of dominance and power than it is about sex, per se. The point to bringing up these distinctions is that a model of repeat SBP must take into account the fact that not all individuals with SBP will qualify as psychopaths, even on a subclinical level. Thus, other factors should serve as predictors for those individuals in the absence of psychopathy. It may be that impulsivity, the one component of psychopathy that seems to be present in all sex offenses, may be the predictor in absence of the entire package of psychopathy. At its most simplistic, sex offending could be reduced to a disorder of behavioral dyscontrol, or an impulse control problem. Most programs seem to recommend that some form of impulse control be taught to offenders to lower

the risk of future SBP (Lane Council of Governments, 2003), further suggesting that impulsivity plays a far larger role than has been previously suggested in the literature.

Exploratory variables. Past theories have often posited that youth with SBP have particularly severe sexual abuse histories, and therefore their SBP are simply "reactive" (Lane, 1997). There is also considerable support for the notion that exposure to violence and pornography may play significant roles in the development of SBP and externalizing behavior disorders in general (Gorman-Smith, Henry, & Tolan, 2004; Margolin, 2005). Thus, any research undertaken in the area of actuarial modeling of SBP should take these constructs into consideration. This may include measures of violence in the home and community, measures of sexual and physical abuse history, and exposure to pornography. Exposure to violence and pornography as well as ongoing abuse may be a function of the type of environment (e.g., residential placement, foster placement).

Current purpose. The purpose of the current study is to examine a modified and expanded version of the Grubin (1999) model using a sample of youth with SBP, with the ultimate intent of discerning the most powerful and accurate method for assessing recurrence risk. An actuarial risk assessment method with real-world clinical utility could lessen the enormous psychological, financial, and societal costs currently incurred by misclassification. As noted previously, much is at stake in the task of classifying youth who have SBP, and this study aims to propose strategies to adequately predict SBP recurrence with maximal accuracy.

Hypotheses

This study proposes to examine a classification system of children who exhibit SBP with a focus on prediction of the categorical outcome of repeat SBP (i.e., yes/no) using an approach termed classification tree analysis (CTA) via optimal data analysis (ODA; Yarnold & Soltysik, 2005). This technique is unique in several respects. First, CTA allows evaluation of predictors in a nonlinear fashion, allowing for multiple pathways to the same outcome. This presents a significant advantage over evaluations performed under the general linear model. Second, CTA via ODA allows researchers to identify the absolute best model (as compared to perfect classification) among all predictors presented for inclusion in the classification tree. The multiattribute "tree" is constructed with successive predictors which serve to classify a gradually decreasing proportion of the total sample with maximal accuracy. Once a dichotomous outcome is defined, it is not necessary to identify optimal cut points for each predictor variable. Instead, ODA finds classification rules that maximize the overall percentage of classification accuracy at each step of the analysis (Yarnold, 1996). Thus, the model allows for exploration of prediction models with no particular a priori prediction structure in place. ODA is also uniquely suited for use with constructs exhibiting multifinality. If one construct does not lead to one outcome, ODA has little difficulty allowing branching to as many possible outcomes as exist, so long as the sample is sufficiently large enough to classify into these branches.

Additional statistical procedures (discussed in Chapter 3, *Results*) provide analysis of the likelihood that the results will replicate in other samples, thus

maximizing confidence that the resultant classification tree is not due to chance characteristics of the sample used. Finally, it is notable that CTA via ODA does not require the typical statistical assumptions to be met (e.g., normality), thus making it ideal for skewed distributions, provided there is adequate sample size for all "branches" of the tree.

A serious problem with studies of the type proposed is defining problem recurrence. For example, if a youth with SBP has an additional incident, but this incident involves less serious (i.e., less coercive or invasive) behavior, should it still be considered a recurrence? In substance-abuse treatment research, it is often expected that individuals may suffer treatment setbacks at some point and engage in substance abuse, but this setback is an expected part of the process (Prochaska, DiClemente, & Norcross, 1992). The emphasis is placed on the client accepting the setback and engaging in behavior that lessens the likelihood of future treatment setbacks.. Similarly, youth with SBP may have setbacks as well, and in the same way, a single recurrence may not indicate a treatment failure. Although future studies may use slightly different criteria, it is important that the criteria for this repeat SBP are clearly defined in this study to facilitate replication. Thus, the following criteria were used for this study.

To be considered a "repeater," or a youth with continued SBP after their initial identified event, a youth in the sample had to experience more than one problem of significantly lesser seriousness than their previous incident, or one or more problems of the same or greater seriousness than their previous incidents. "Significantly lesser" seriousness was defined as a clinically significant drop in the

invasiveness of the SBP (e.g., from contact offense behavior to non-contact behavior). These cases were assessed by the current author and a licensed clinical-child psychologist working in the SBP field. There were no instances of disagreement with respect to judgment of clinically significant decline from one act to the next.

Despite the lack of necessity for specific hypotheses regarding the prediction model (given the proposed statistical procedures), there are a number of hypotheses that will affect what variables are included in the model. These specific hypotheses are outlined below.

Prediction model for SBP Youth. The optimal model for youth will include psychopathy for some, but only impulsivity for others, as it seems SBP youth may be actually made up of two subgroups: the psychopathic youth, who have more sadistic aims, and the youth who may be more benign in their fantasies and seem to possess more empathic sense. These subgroups should be represented in two separate pathways ("branches") in the tree model. Isolation, measured via parental attachment, will be included in the model. Sexual deviance will be present as a predictor for both subgroups in the tree.

The following variables will be included to assess their influence on the model as proposed: 1) gender of subject; 2) ethnic group of subject; 3) age at disposition of initial offense; 3) severity of initial offense; 4) externalizing symptoms by caregiver report; 5) any evidence of physical and/or sexual abuse by self-report or electronic records; 6) exposure to inappropriate sexual content in the community or at home (either through witnessing of sex acts or exposure

to pornography); 7) placement at screening (e.g., residential, foster); 8) self-report of illicit substance use; and 9) time between initial and follow-up data collection.

CHAPTER 2

METHOD

Participants

Participants represented a portion of the population of identified youth with SBP in a large urban area in the Midwest. All of these youths were wards of the state at the time of data collection, as the data were collected as part of a larger effort by the state Department of Children and Family Services (DCFS) to ascertain the characteristics of this population while also gathering information on factors related to relapse risk. The sample consisted of 251 cases of youth aged 10 years and up who were identified by the state as "sexually aggressive youth" at the time, a tracking system that has since been dismantled.

A number of youth in the current sample had not been adjudicated at the time of data collection. Thus, it should not be presumed that inclusion in this data set is necessarily synonymous with adjudicated status. Nonetheless, youth in the current sample had exhibited behavior concerning enough to warrant the attention of the Illinois Department of Children and Family Services. This attention usually precludes legal involvement, as youth are often mandated to receive treatment by DCFS. This step often renders the adjudication process itself unnecessary. Only youth aged 10 and above and their caregivers were included in this study because the initial study obtained self-report only from

youth aged 10 and above, and self-report was considered important to the questions posed here. Therapist report was not considered as the number of cases missing data was high (n = 69).

The mean age of the sample at the time of disposition (initial entry into the data collection process) was 13 years and four months. Ages ranged from approximately 10 to 18 years old. Males constituted 74.5% of the sample (n = 187), while females numbered 64, or 25.5% of the sample. At younger ages, proportions of boys to girls were roughly equal, but there were significantly more boys than girls among the older (i.e., > 11) youth. The sample was primarily African American (80%), with small proportions of Caucasian (13%), Latino (5%), and multi-racial (2%) youth. These proportions roughly match the ethnic proportions for wards in substitute care in the area, indicating equal likelihood of group membership (i.e., identification as sexually aggressive youth) for all ethnic groups.

Of the youth in the sample, 41% were in foster placement, 9% were living with relatives, and 50% were in settings such as residential facilities, group homes, or shelters at the time of screening for the study. Youth were screened into the database following an index offense that was identified by state authorities as sexually deviant in some way. The investigators (Spaccarelli, Edejer, Bushell, Karaitis, & Jones, 2001) opted to code the index events by invasiveness, on an 8-point scale (see Table 1), with 1 representing sexualized behavior only (e.g., unusual or precocious behavior, but no victims), 2 as consensual sex play amongst younger children, or consensual sex amongst teens, 3 as non-contact

SBP, such as exhibitionism, voyeurism, encouraging/directing sexual contact of others, and verbal sexual harassment, 4 as non-genital fondling, 5 as genital fondling, 6 as genital contact without penetration, 7 as oral penetration, and 8 as vaginal or anal penetration. 40.2% (n = 101) of the total sample (N = 251) were found to have suffered (per DCFS file review or self report) sexual victimization themselves.. 78.9% (n = 198) were victims of physical abuse (also per DCFS records or self-report).

Repeat Status of Participants.

34 participants were ultimately classified as "repeaters," in that they engaged in the same- or higher-severity SBP again between time 1 and time 2 of data collection. As noted previously, repeaters were not classified as such if they had significantly declined in the severity of their SBP as compared to their index events. It is reasonable to question whether inclusion as a repeater could be related to a possible ceiling effect. Specifically, it is possible that youth with higher-severity initial incidents would be less likely to be included because they exhibit less room to move on the scale of severity. Youth with lower-severity index events had more room to move to higher levels of severity, meaning there is the possibility of their being more readily identified as "repeaters." This does not seem to be the case with the current data, however. Out of 70 youth who had *any* repeat events (which does not necessarily lead to classification as repeater), only 17 of those were youth with index events of 4.5 or higher in severity. This is especially notable given that these youth likely received a high degree of attention and support after engaging in such severe initial SBP.

Further, out of the remaining 53 youth, their repeat events tended to be unequivocal in their status as judged by the severity of the repeat event. The mean and median of the change in severity of these 53 youth was nearly zero, with a standard deviation of 1.9. Examination of this distribution revealed that in nearly equal proportions, some youth engaged in significantly more severe behavior, some engaged in the same severity of behavior, and some engaged in significantly less severe behavior. In summary, while nearly double the number of youth classified as repeaters actually engaged in repeat behavior, those who were not classified as repeaters unequivocally seemed to have improved (i.e., the severity of their repeat SBP was much lower than their initial SBP severity). This suggests that upward movement in the severity of SBP over time was not responsible for the higher proportion of lower-initial-severity SBP youth in the "repeater" group.

Table 1

Distribution of SBP in the current sample

Invasiveness	*N*	*% of total N (251)*	*Cumulative %*
Sexualized Behavior Only (1)	12	4.8%	4.8%
Consensual Sex, Teens (2)	21	8.4%	13.2%
Consensual Sex Play, Pre-Teens (3)	21	8.4%	21.6%
Non-Contact Behaviors (3)	31	12.4%	34%
Non-Genital Fondling (4)	68	27.1%	61.1%

Genital Contact Without Penetration (5)	33	13.1%	74.2%
Genital Contact With Attempted Penetration (6)	8	3.2%	77.4%
Oral Penetration (7)	22	8.8%	86.2%
Vaginal/Anal Penetration (8)	30	12.0%	98.2%
No SBP noted (apparent misclassification)	5	1.8%	100%

Materials

The present data were collected as part of a longitudinal study of youth with sexual behavior problems (Spaccarelli et al., 2001). In an effort to gather the most comprehensive information available, the investigators gathered multi-informant report via in-person and computer-administered interviews from caregivers, youth, and the therapists of those youth, as well as records review from various state databases. The battery of measures included measures of: oral and reading comprehension, youth welfare history, family attachment, witnessing physical conflict, experience of physical conflict, witnessing community violence, psychological symptoms, sexual learning and knowledge, sexual attitudes, impulsive and aggressive personality traits, a validity scale, treatment rejection, sexual activity, drug use, knowledge of others' victimization, sexual victimization, sexual perpetration, sexual deviance, and exposure to pornography. All measures were gathered at time one and time two of the larger study. However, the current

study used only time one data, as the aim was to predict time two outcomes based on information available at time one.

This study used the following measures from the immediately preceding list, as well as DCFS records: Family attachment, psychological symptoms, sexual deviance, impulsive and aggressive personality traits (i.e., psychopathy), self-report validity, experience of physical conflict and sexual victimization, youth drug use, and exposure to pornography and other inappropriate sexual material The excluded measures were used for other research purposes and will not be discussed here. See Table 2 for a brief description of each of these included measures.

Table 2

Included Measures

Construct	*Measure name*	*Reporter*	*Number of items*
Attachment (via positive parenting practices)	Parenting Practices Measure (Tolan, Gorman-Smith, & Tully)	Y	5*
Exposure to pornography	Developed for study (Spaccarelli et al., 2001)	C, Y	33, 2
Psychological symptoms	Conners' Rating Scale (Conners, 1990)	C	20*

Psychopathy	Antisocial Process Screening Device (Frick et al., 1992)	Y	16*
Sexual behavior and deviance	Adapted from the Child Sexual Behavior Inventory (Adapted from the Multiphasic Sex Inventory; Nicholas & Molinder, 1984)	C, Y	43, 13
Validity	Marlowe-Crowne Social Desirability Scale, short form (Reynolds, 1982)	Y	13
Victimization (physical)	Conflict Tactics Scale (Straus et al., 1996)	Y	18*
Victimization (sexual)	Developed for study (Spaccarelli et al., 2001) and electronic records	Y	3
Youth drug use	Developed for study (Spaccarelli et al., 2001)	Y	9

*Represents an abbreviated version of the parent instrument.
C = caregiver, Y = youth.

Family attachment. Family attachment as a construct actually contains both parenting practices and parent-youth bond as measured by the Parenting Practices Measure (PPM) used by Gorman-Smith, Tolan, Zelli, and Huesmann (1996). The PPM measures four constructs: positive parenting, discipline effectiveness, avoidance of discipline, and extent of monitoring and involvement in the youth's life. Respondents use a 5-point Likert scale with the following descriptors: *Never*, *rarely*, *sometimes*, *often*, and *almost always*. Each item is scored 1 to 5, with the sum of items for each factor used as the subscale score. This measure is designed for multi-informant report from both caregiver and youth; however, the current study used only youth report The scales that constitute the PPM have demonstrated good to excellent reliability and validity, and have been subjected to confirmatory factor analysis to validate the structure as designed (Gorman-Smith, Tolan, Zelli, & Huesmann, 1996). The positive parenting scale constitutes a measure of attachment from youth report.

Psychological symptoms. Psychological symptoms of the youth as rated by caregivers were measured by four scales of the 48-item Conners' Rating Scale (Conners, 1990). These scales were the ADHD Total scale, the Hyperactivity scale, and an abbreviated Conduct Problems scale. The Conners' scales have demonstrated good to excellent validity and reliability (Sattler, 2001), and yielded data on conduct problems, impulsive/hyperactive behavior, anxiety, and psychosomatic problems.

Impulsive and Aggressive Personality Traits. To measure the self-reported traits of impulsivity, callous-unemotional traits, and narcissism as they relate to

psychopathy, this study utilized a modified form of the Antisocial Process Screening Device (APSD; Frick, O'Brien, Wootton, & McBurnett, 1994; Marsee, Silverthorn, & Frick, 2005). The APSD used in this study contained 16 items aggregated into three scales: Callous/Unemotional (CU) traits, impulsivity/conduct (IMP) problems, and narcissism (NAR). The original APSD contained 20 items, but recent studies have revealed that certain items do not appear to relate well to the constructs they are intended to represent (Marsee, Silverthorn, & Frick, 2005). This instrument is normally administered with a three-point Likert scale consisting of 0 (*not at all true*), 1 (*sometimes true*), or 2 (*definitely true*). However, as the current study involved the administration of a large battery of measures to youths with widely varying cognitive capacity, the study authors (Spaccarelli et al., 2001) opted to use a four-point scale as with most other instruments in the study to minimize the confusion often associated from using numerous different response scales with younger youth. The four-point scale ranged consisted of the following response options: 1 (*not at all true*), 2 (*a little true*), 3 (*mainly true*), and 4 (*very true*).

The measure used in this study included four CU scale items, five IMP scale items, and seven NAR scale items. This measure most closely approximates the APSD used by the Fast Track project (Doyle & McCarty, 2000). Although the Fast Track project used the caregiver version, their proposed structure seems consistent with characteristics of the self-report instrument in this sample. Also, there is mounting evidence that self-report of psychopathy can be valid in a variety of child and adolescent populations (e.g., Caputo, Frick, & Brodsky, 1999;

Vitacco, Rogers, & Neumann, 2003). The reliabilities for the CU, IMP, and NAR scales in this study (as measured by coefficient alpha) were .80, .80, and .83, respectively (N = 251).

Validity of Self-report. The short form of the Marlowe-Crown Social Desirability Scale (Reynolds, 1982) was used for this study. Its response scale was modified to use the same 4-point Likert-type scale as the other instruments in the battery to enhance comprehension and minimize response errors based on scaling confusion. The short form of the Marlowe-Crown consists of 13 items designed to assess a respondent's tendency to answer questions in a positively biased way, or providing answers with the aim of creating a more positive impression to others at the expense of truthfulness. Individuals who score high on measures of social desirability tend to under-endorse items that reflect negative personality characteristics or behaviors (Edwards, 1953). This measure has been used extensively in numerous and varied populations with acceptable reliability and validity (Jones, Smith, & Holmes, 2004; Reynolds, 1982). In this study, the sum total of the items was used, which could be endorsed with the same scale as the APSD (ranging from 1 to 4). The Marlowe-Crown items were interspersed with the APSD items in administration to more accurately assess social desirability biases toward the instrument most likely to be sensitive to those issues (the APSD).

Sexual behavior and deviance. The self-report sexual deviance measure was adapted from the Multiphasic Sex Inventory (Nichols & Molinder, 1984) and asks true-false questions about a variety of deviant sexual behaviors, ranging

from voyeurism to sexual sadism. The current study used 13 items scored as true or false, or one and zero, respectively. A higher sum score indicates higher levels of self-reported deviance. Caregivers were administered the Child Sexual Behavior Inventory (CSBI, Friedrich et al., 1992; Friedrich et al., 2001). The caregiver measure was administered in two forms, one original and one slightly modified for caregivers of older children. The caregiver measure is distinct from the youth measure in that it asks about behavior that may be normative at certain ages, but abnormal at others, such as "touches another child's private parts" or "tries to have sexual intercourse with another child or adult." The caregiver measure was designed with a four-point Likert scale to indicate frequency of occurrence, with 0 representing *never*, 1 as *once a month or less*, 2 as *one to three times a month*, and 3 as *at least once a week*. These items were summed to produce a total caregiver score with higher scores representing sexually deviant drive, or hypersexuality.

Experience of Physical Conflict and Sexual Victimization. To assess youths' self-reported exposure to physical conflict, 18 items were adapted from the revised Conflict Tactics Scale (CTS2; Strauss, Hamby, Boney-McCoy, & Sugarman, 1996). Nine items asked about witnessing of physical violence, while the second set of nine items asked about experiencing physical violence in the home, ranging in content severity from being shoved by a caregiver to having a knife or gun used against them by a caregiver. Instead of the eight response options of the CTS2 for adults, responses were simplified to a five-point scale (1 = *never,* 2 = *only once,* 3 = *2-5 times total,* 4 = *6-10 times total, and* 5 = *more than 10 times*) to avoid

confusing youths of lower cognitive capacity. For the current study, *any* endorsement of any of the physical violence experience items was scored as a "1" representing the presence of physical abuse in the youth's history, as recommended by Straus (2006). Additionally, a youth could receive a "1" if there was any record of physical abuse in the state DCFS database. Youth with no endorsements of any of the nine experience of violence items were scored as "0", representing no reported physical abuse. Thus, a youth would be coded as having been physically abused if there was any self-report or electronic record of physical abuse.

Sexual victimization was assessed by self-report of a variety of sexual experiences in the home. The self-report measure consisted of 57 items and was constructed specifically for the larger study (Spaccarelli et al., 2001). The measure was designed to allow youth to report victimization based on their memories of discrete events. Once the nature of the event was established via initial questions, the measure then assesses the details of the event with follow-up probes (e.g., "Did that person kiss or touch your mouth or body?"). Youth were coded as having been sexually abused if they endorsed any of the following three yes-no items: 1) "Have you felt like someone made you do sexual stuff when you really didn't want to?", 2) "Other than who we just talked about, did another person make you do sexual things when you didn't want to?", and 3) "Other than those who you may have just talked about, have you ever done sexual stuff with someone much older than you (5 years or more)?".

Youth Drug and Alcohol Use. Drug and alcohol use were assessed by a series of questions developed by the study authors (Spaccarelli et al., 2001). Nine items ask participants to rate the frequency of their use of alcohol, tobacco, amphetamines, inhalants, and other various substances. Participants may respond via a five-point Likert scale ranging from *never* to *use regularly now*, scored as 1 through 5. The sum total score was used to represent global drug and alcohol use.

Exposure to Pornography and Other Inappropriate Stimuli. Exposure to pornography and other inappropriate stimuli (e.g., sex acts) was also assessed by questions designed specifically for the larger study. 33 items ask youth about exposure to body parts and sex acts through various contexts, from music up to and including personal observation. Participants may rate how often they are exposed using a five-point Likert scale ranging from *never* to *all the time*, scored as 1 through 5. The sum total was used to represent exposure to pornography and other inappropriate stimuli. A caregiver version consists of two questions asking about inappropriate stimuli (graphic television/movies and magazines with nude pictures) a youth *may have seen* while at home, using the same rating scale.

Procedure

Caregivers were contacted by study workers after being screened as eligible via a records search. Youth were required to undergo assent procedures, so not all youth necessarily reported to the study despite their caregivers' consent to participate in the study. Once at least parents (if not parents and youths) consented to participate in the study, they were scheduled for an initial interview

that lasted several hours and contained the full battery of measures listed. Youth were able to complete their measures via laptop computers to decrease the likelihood of reactivity and to provide better assurance of confidentiality. Literacy was assessed to ensure validity of self-report.

Archival data were coded by trained research team members at the undergraduate and graduate level and checked for accuracy by the investigators. These data were entered into SPSS 13.0 for Windows.. The present data were exported to ASCII format via SPSS 13.0 and then analyzed using ODA for Windows (Yarnold & Soltysik, 2005). "Jack-knife" leave-one-out analyses (LOO) were conducted to better ensure that the results of the classification tree analysis are replicable in other samples. The detailed analytic plan is outlined in the next section (Chapter 3, *Results*).

CHAPTER 3

RESULTS

Overview of data analytic strategy.

As the primary question of this study is one of prediction (i.e., repeat SBP, yes/no), a number of analytic methods would have been appropriate. However, if the question is framed specifically as whether or not a youth will continue to exhibit SBP, it becomes a question of classification. Specifically, there are two distinct classes of interest, and each youth will occupy one or the other. One class is comprised of youth whose primary criterion for membership is a lack of continued SBP after their "index offenses," or the offenses that resulted in their being classified as having SBP by the state Department of Children and Family Services (DCFS). The other class is occupied solely by youth who continue to exhibit SBP after their index offenses. Put simply, youth will either continue to exhibit SBP, or they will not. Considering outcomes in this manner, the ultimate measure of worth for any predictive model is its accuracy in predicting a youth's class membership.

While accuracy of prediction can be assessed by linear methods such as logistic regression, Classification Tree Analysis (CTA) offers a number of advantages (Feldesman, 2002; Yarnold & Soltysik, 2005): 1) there is no

requirement of normality for variable distributions, 2) missing data do not affect whether or not an individual can be included in the analysis, 3) it is robust to outliers, 4) it requires no advance variable selection, as it selects the best variables for the tree based on effect strength, 5) it can utilize any combination of categorical or continuous predictor variables, and 6) it is nonlinear, meaning the same set of linear weights is not assumed to apply to all respondents (as it is in multiple or logistic regression). This final advantage is especially useful as small groups of individuals who may represent a predictive path to an outcome (i.e., different variable combinations, different variable orders, or different decisional cutpoints predict their class membership) are allowed to have their own branch of the classification tree. In a linear analysis, these unique cases would simply represent error, such as error around a regression line. Individuals who share common paths to an outcome will share a branch of the tree. Thus, classification trees allow clear visual identification of individuals who may represent a particular type within the conceptual framework (e.g., early-onset repeat cases).

The primary purpose of the present study was to generate a classification model that identifies as accurately as possible those youths who will continue to exhibit SBP, and those who will not. Optimal Data Analysis (ODA; Yarnold & Soltysik, 2005) was used to conduct a nonlinear, hierarchically optimal CTA to construct a multiattribute "tree" model in which successive predictors serve to classify with maximum accuracy a gradually decreasing proportion of the overall sample. This technique has been utilized in various fields with excellent results, such as predicting early sexual debut among adolescents (Donenberg, Bryant,

Emerson, Wilson, & Pasch, 2003), medical decision-making (Feinglass, Yarnold, Martin, & McCarthy, 1998), and diagnosis of Attention Deficit Hyperactivity Disorder (Ostrander, Weinfurt, Yarnold, & August, 1998). CTA via ODA offers the advantage of a technique that explicitly maximizes classification accuracy, in contrast to other methods of conducting CTA, such as Relative Operating Characteristic (ROC) Analysis or Bayesian methods (Yarnold, 1996).

UniODA and the current study.

A brief explanation of how univariate ODA (herein referred to as ODA) is used in CTA may be helpful. ODA is a nonparametric analytic procedure which optimizes two-category classification (e.g., 0 or 1). Although some may view a dichotomous outcome as overly simplistic, it bears consideration that all decisions in reality are truly dichotomous when distilled to their purest form. Such decisions can be: 1) whether or not to offer intervention; 2) whether or not a diagnosis is applicable; 3) prediction of win or loss, or virtually endless other examples. With respect to the current research question, the ultimate question is one of whether or not a youth presents significant risk of developing continued SBP. All subsequent decisions such as strength or intensity of intervention, placement type, and other factors (as discussed in the *Introduction*) stem from this initial decision. CTA via ODA is the most suitable analytic strategy for this type of question.

In the case of a continuous predictor variable, ODA determines the optimal cutpoint on the distribution of an attribute variable (e.g., IQ) that yields

the best classification of individuals in the sample according to the following simple example algorithm:

IF $X < 80$ THEN class = 0; if $X \geq 80$ THEN class = 1.

ODA can also be used with categorical attribute variables, but instead of using a cutpoint, it identifies the discrete value (e.g., values representing hair color) that best classifies individuals in the sample according to this example algorithm:

IF X = brown, then class = 0; otherwise, class = 1.

If X in the above algorithms represents an individual's score on the attributes of interest, each individual can be classified by ODA according to that individual's scores on those attribute variables. For either type of variable, ODA selects the cutpoint or discrete value (for continuous or categorical variables, respectively) that optimally classifies participants, or classifies the most participants as accurately as is possible.

Selection of variables. Selecting the variables to be included in the multiattribute tree can be likened to "growing" the tree. To begin, all attribute variables of interest are analyzed in ODA to evaluate their classification performance, one at a time. The attribute variable that yields the best classification performance (evaluated via a standardized measure of effect strength) is selected as the first attribute variable that begins the tree, at the top. Yarnold and Soltysik (2005) provide the following rules for qualitatively evaluating effect strength (ES): ES values less than 25% are considered *weak*, between 25% and 50% are *moderate*, between 50% and 75% are *relatively strong*, between 75% and 90% are *strong*, and above 90% are considered *very strong*.

After the first variable has been selected, all the attribute variables are analyzed for each of the two branches of the initial node of the tree. For example, if the initial variable in the tree is "hair color," and ODA identifies the optimal classification rule to be: IF X = brown, then class = 0; otherwise, class = 1, then the list of attribute variables is run again, but this time, one analysis includes only those with brown hair, and all other hair colors are run as their own analysis. Each of these two analyses (one for the brunettes, and one for all others) will yield ES estimates for all attribute variables, and the next attribute variables in the tree are selected based on the same inclusion criterion: the attribute variable with the strongest ES is included as the next node in the tree. In this way, the "branches" of the tree are grown, and this procedure is repeated to classify a gradually decreasing proportion of individuals in the sample until all are classified as accurately as possible, or there are no more cases in one of the classes (in which case there are not enough cases to continue classification).

Significance testing in ODA. Significance testing in ODA is achieved using different methods than in more familiar analytic techniques, such as analysis of variance. Using a distribution-free permutation probability method known as Fischer's randomization procedure (Bradley, 1968), ODA provides estimated p values via Monte Carlo simulation. While this may sound complex, the method is actually quite simple: ODA randomly shuffles the class values for participants (i.e., 0 or 1). Thus, each case has an equal probability of being a 0 or 1 in this procedure. Then, ODA uses the attribute variable of interest to classify the participants based on their randomly assigned class value. ODA performs this

classification over and over again, records the levels of classification performance, and provides the number of times that the procedure yielded classification as good or better than the actual class variable under study.

The number of iterations in the Monte Carlo procedure is user-selectable in ODA; however, Yarnold and Soltysik (2005) have found that chance can almost always be accurately estimated using 10,000 iterations, as the estimated *p* value tends to stabilize at this point. Thus, if ODA is set to perform 10,000 iterations in its Monte Carlo simulation, it will do so and then report a *p* value that represents the number of times out of those 10,000 iterations (expressed as a proportion) that it was able to correctly classify the participants using the given attribute variable when they were randomly assigned to a value of the class variable. If the simulation yields a high number of iterations that were able to correctly classify as well or better than when the participants were actually classified with their real class values, such as 4000 out of 10,000 (estimated $p = .4$), this would suggest that the attribute variable in question performs at a level frequently achieved by chance. If the simulation yields a low number of occurrences in which random class assignment yields the same or better classification (e.g., 100 out of 10,000, or estimated $p = .01$), then it can be inferred that the attribute variable in question performs at a level that occurs infrequently by chance. As with other statistical procedures, $p < .05$ is used as the inclusion criterion for attribute variables in the tree.

Removal of nodes in the classification tree. After the tree has been created and all terminal nodes have been established (i.e., the tree is complete), a number of

procedures can be employed to minimize the likelihood of capitalizing on chance. The first of these procedures is called the "jackknife" or "leave-one-out" procedure (LOO). The LOO procedure removes one participant from the sample and re-runs the analysis on the remaining subsample, each time keeping track of the classification accuracy for the given attribute. This procedure is repeated until all possible subsamples have been analyzed. The results are stored and tabulated iteratively across participants. An attribute variable that is LOO-stable is one that yields the same result (i.e., the same overall accuracy, or correct classification of participants) across all possible subsamples of size N - 1. Including only nodes in the tree that are LOO-stable increases the expected cross-sample generalizability of the final classification model, or the likelihood that the model will be the same in other samples.

A second method to remove from the tree nodes that may be present due to chance is the use of a sequentially rejective Bonferroni correction procedure that controls for experiment-wise Type I error (Cook & Campbell, 1979; Ryan, 1959) and increases confidence in the effects (Klockars, Hancock, & McAweeney, 1995). One method of utilizing the Bonferroni correction for multiple comparisons while not overly reducing power is the Sidak step-down adjustment procedure (Holland & Copenhaver, 1987; Yarnold & Soltysik, 2005). In the Bonferroni-Sidak correction procedure, the number of nodes in the final tree model determines the correction, with the most stringent correction applied to the lowest p value node. Each node is subjected to this correction in ascending order of p value, with the last node being compared to the .05 criterion

(assuming all other nodes were statistically significant at their respective Bonferroni-corrected *p* values).

A final method to prune potentially irrelevant nodes from the classification model that has been suggested by Bryant (personal communication, July 5, 2006) involves utilizing conceptual parsimony in conjunction with classification accuracy. If a node adds little (or nothing) in terms of classification accuracy, then the concept of parsimony should reign and the node should be eliminated. Thus, accuracy is evaluated in light of its potential cost to conceptual parsimony, and incremental increases in accuracy that add unnecessary complexity to the model are disregarded.

Description of analyses.

For the current study, 30 variables (see Table 3) were included as attribute variables to be assessed for inclusion in the classification tree, as well as one class variable that identified participants to ODA as being either "repeaters" or "non-repeaters." Missing data points were minimal and were coded as -9 in the dataset. ODA offers the option of identifying missing data by a discrete number in the syntax, and -9 is recommended by the authors as their convention (Yarnold & Soltysik, 2005). ODA then recognizes the syntax to exclude any participants from a particular univariate analysis if their value on the attribute variable is -9. One advantage of ODA is that these participants are not fully excluded from the tree, but only from that particular node of the tree (if that variable is indeed included as a node). See Table 4 for a listing of descriptive statistics for the included variables.

Table 3

Included Variables

Variable Name	*Description/Source*
Age at Disposition	Age at time DCFS rendered a decision regarding SACY status, from DCFS electronic records
Age Difference Between Youth and Oldest Victim	Years of age difference between the youth participant and oldest identified victim in index event
Age Difference Between Youth and Youngest Victim	Years of age difference between the youth participant and youngest identified victim in index event
Aggression Index	Sum of all items comprising the Aggression Index from the Conners' scales
APSD Callous-Unemotional	Sum of Callous-Unemotional Scale items from the APSD
APSD Impulsivity	Sum of Impulsivity Scale items from the APSD
APSD Narcissism	Sum of Narcissism Scale items from the APSD

Combined Physical Abuse	Dichotomous indicator of any physical abuse from electronic records or self-report
Combined Sexual Abuse	Dichotomous indicator of presence of sexual abuse by either electronic records or self-report
Conduct Problems	Sum scale of conduct problems from the Conners' scales
Drug Usage	Sum of drug use items reflecting frequency of use, self-report
Ethnicity	Categorical ethnicity from self-report (African American, White, Latino, Multi-ethnic)
Exposure to Inappropriate Sexual Content in the Community	Three-level variable representing degree of exposure to inappropriate content with 0 = none, 1 = private parts of others, and 2 = sex acts, from self-report
Gender	Dichotomous gender from self-report
Impulsivity/Hyperactivity Index	Sum of Impulsivity/Hyperactivity

	Index items from Conners' scales
Overall Hyperactivity Index	Sum all items comprising the Hyperactivity Index of the Conners' scales
Physical Abuse (electronic)	Dichotomous indicator of founded physical abuse from DCFS electronic records
Physical Abuse (self-report)	Dichotomous indicator of any physical abuse as self-reported on the Conflict Tactics Scale
Placement at time of screening	Categorical variable from caregiver report reflecting the placement of the child at time of entry into the study (e.g., residential)
Pornography exposure (caregiver report)	Sum of items reflecting frequency of pornography with nude images seen by child, caregiver report
Pornography exposure (self-report)	Sum of self-report items reflecting frequency of pornography with nude images

	seen
Positive Parenting Behavior	Sum of items representing positive parenting behaviors from the Parenting Practices Scale, self-report
Repeat Status	Dichotomous class variable as coded by author, describing repeat yes or no
Severity of Initial SBP	Invasiveness of index SBP event coded from one to eight (eight being most invasive)
Sexual Abuse (electronic)	Dichotomous indicator of founded sexual abuse from DCFS electronic records
Sexual Abuse (self-report)	Dichotomous indicator of any sexual abuse as self-reported on survey
Sexual Deviance (self-report)	Sum of items from sexual deviance self-report measure
Sexual Deviance (caregiver report)	Sum of items from sexual deviance caregiver-report measure

Social Desirability	Sum of social desirability items from the Marlowe-Crown
Time Between Time One and Two Data Collections	Time in days between first and second data collections, from study records

Table 4

Descriptive Statistics

Variable Name	*Appropriate Descriptive Statistic*
Age at Disposition	*M* = 13.3 *SD* = 1.95
Age Difference Between Youth and Oldest Victim	*M* = 3.9 *SD* = 4.5
Age Difference Between Youth and Youngest Victim	*M* = 5.7 *SD* = 5.6
Aggression Index (possible range 0 - 12)	*M* = 4.3 *SD* = 3.2
APSD Callous-Unemotional (possible range 4 - 16)	*M* = 5.3 *SD* = 2.4

APSD Impulsivity (possible range 5 - 20)	*M* = 3.7 *SD* = 2.7
APSD Narcissism (possible range 7 - 28)	*M* = 5.1 *SD* = 3.6
Combined Physical Abuse	78.9 % (yes)
Combined Sexual Abuse	40.2% (yes)
Conduct Problems (possible range 0 - 24)	*M* = 10.0 *SD* = 6.0
Drug Usage (possible range 0 - 54)	*M* = 11.3 *SD* = 7.3
Ethnicity	See *Participants*
Exposure to Inappropriate Sexual Content in the Community (possible range 0 - 2)	*M* = 0.7 *SD* = 0.8
Gender	See *Participants*
Impulsivity/Hyperactivity Index (possible range 0 - 12)	*M* = 5.6 *SD* = 3.1
Overall Hyperactivity Index (possible range 0 -30)	*M* = 12.1 *SD* = 6.8
Physical Abuse (electronic)	27.8% (yes)

Physical Abuse (self-report)	70.2% (yes)
Placement at time of screening	See *Participants*
Pornography exposure (caregiver report; possible range 0-2)	*M* = 0.3 *SD* = 0.6
Pornography exposure (self-report; possible range 0 - 25)	*M* = 7.0 *SD* = 5.4
Positive Parenting Behavior (possible range 0 - 25)	*M* = 3.7 *SD* = 1.0
Repeat Status	See *Participants*
Severity of Initial SBP	See *Participants*
Sexual Abuse (electronic)	14.0% (yes)
Sexual Abuse (self-report)	40.2% (yes)
Sexual Deviance (caregiver report; possible range 0 - 129)	*M* = 10.5 *SD* = 15.6
Sexual Deviance (self-report; possible range 0 -13)	*M* = 2.0 *SD* = 2.4
Social Desirability (possible range 13-52)	*M* = 27.7 *SD* = 7.8
Time Between Time One and Two Data	*M* = 522.9

Collections (in days)	*SD* = 209.5

Analyses were conducted using the following steps. First, all 30 univariate analyses were examined to ascertain which variable exhibited the largest effect strength for sensitivity (ESS), a standardized index measuring the strength of that attribute's ability to accurately classify individuals (as discussed previously in this section under *Selection of variables*). ESS ranges from 0 to 100, with 0 representing classification accuracy expected by chance, and 100 representing perfect classification accuracy. The variable with the largest ESS was selected for inclusion in the tree at each node. Second, this variable was included only if its ESS was LOO-stable, to ensure robust generalizability of the tree. Third, the variable had to have an estimated *p* of <.05; Bonferroni-Sidak correction was used in later steps to prune nodes of the tree further. If these three conditions were met, then the attribute was included in the tree, and its branches were defined by the cutpoints or discrete values identified as optimal by ODA. This procedure was repeated until no more nodes emerged. See Table 5 for a listing of the ESS of each of the variables included in the initial analysis.

Table 5

Effect Strength for Sensitivity (ESS) in Initial Univariate Analysis

Variable Name	*ESS (%)*
Age at Disposition	4.2%
Age Difference Between Youth and Oldest	*19.0%*

Victim	
Age Difference Between Youth and Youngest Victim	12.4%
Aggression Index	*7.9%*
APSD Callous-Unemotional	*19.8%*
APSD Impulsivity	*20.9%*
APSD Narcissism	29.7%†
Combined Physical Abuse	7.7%
Combined Sexual Abuse	18.0%
Conduct Problems	*8.5%*
Drug Usage	*27.8%*
Ethnicity	16.5%*
Exposure to Inappropriate Sexual Content in the Community	*4.9%*
Gender	2.1%
Impulsivity/Hyperactivity Index	*8.7%*
Overall Hyperactivity Index	*8.4%*
Physical Abuse (electronic)	24.9%†
Physical Abuse (self-report)	3.9%
Placement at time of screening	*31.1%*
Pornography exposure (caregiver report)	*4.8%*
Pornography exposure (self-report)	*9.8%*

Positive Parenting Behavior	14.0%
Repeat Status	-
Severity of Initial SBP	32.3%†
Sexual Abuse (electronic)	5.0%
Sexual Abuse (self-report)	20.9%*
Sexual Deviance (caregiver report)	*22.3%*
Sexual Deviance (self-report)	24.5%†
Social Desirability	*19.9%*
Time Between Time One and Two Data Collections	*28.4%*

Note: Italicized values indicate ESS was not LOO-stable. P values not reported for LOO-unstable values. LOO-stability does not necessarily suggest statistical significance. Thus, many values reported here were LOO-stable but did not achieve statistical significance.
*: $p < .05$.
†: $p < .01$

The size of the final tree model is influenced by a number of parameters. The first is sample size; the more participants are available to classify, the more likely it is that additional nodes may emerge as significant predictors in the final tree model. While selection of the nodes is not influenced by sample size (selection is made by ESS, as previously noted), more participants may make it more likely a node will be LOO-stable. Other parameters that influence the potential size of the final tree are number of predictor variables included, the p value used for inclusion of nodes, and the reliability of both the predictor variables and the class variable. To obtain a tree that contains a larger number of

attribute variables and terminal nodes, a researcher would require sufficient sample size for there to be sufficient numbers of participants at various nodes of the tree as well as reliably measured attribute and class variables.

Using this strategy, the initial tree yielded a model with four attribute variables and five terminal nodes (final predicted outcomes) at the end of each branch. This initial tree is displayed in Fig. 1.

Initial classification model. The first node selected to begin the tree at the top was the severity of the index event for each participant, rated on an 8-point scale (see Table 1). Event severity yielded an ESS of 32%, a moderate effect. ODA identified a cutpoint of >4.5 to be the optimal predictor of non-repeater status. This runs counter to the proposed prediction models that contained hypotheses predicting higher likelihood of repeat behavior if initial event severity was high (suggesting more severe delinquency). Interestingly, 4.5 is exactly between "non-genital fondling" and "genital contact without penetration" in the coding system of SBP invasiveness used in this study. According to this initial node in the model, youth who were accused of genital contact SBP were less likely to exhibit continued problems than youth who were accused of non-genital contact offenses.

It is notable that other variables did exhibit similar but slightly lower ESS figures when compared to initial event severity. For example, in the initial univariate analysis, the Narcissism factor of the APSD yielded an ESS of 29.7% and was also LOO-stable. This suggests Narcissism does have a modest relationship with repeat status, and this is supported by the inclusion of

Narcissism later in the tree (as explained below). However, curious readers might note that other variables with similar (but slightly weaker) ESS numbers are not included in the model, such as physical abuse status by electronic record (ESS = 24.9%), or self-report of sexual deviance (ESS - 24.5%). While ODA does not operate with the goal of maximizing variance, it can be said that after accounting for other variables in the tree (such as initial event severity), these variables do not add any additional explanatory value (similar to when a variable with an initially significant relationship is removed from a stepwise multiple regression analysis when preceded by a variable that explains most of its variance). Thus, it should not be said that these other variables not included in the tree do not exhibit a relationship with repeat status. Instead, it should be noted that the most parsimonious model does not require these variables to explain repeat status. It is possible that with a larger sample (and thus more participants to classify), these variables might serve to classify a small proportion of the sample. However, once the higher-ESS variables are included in the model, these variables do not classify as well as their initial ESS figures might suggest.

The initial tree then yielded a node for the predicted repeaters from index event severity that contained the attribute variable of Narcissism from the Antisocial Process Screening Device, or APSD. Narcissism is one scale from the APSD that identifies youth who talk excessively about themselves or their possessions, demonstrate little regard for others' concerns, and a generally self-centered approach to problem-solving. This variable exhibited an ESS of 30%, also a moderate effect. ODA identified a relatively low cutpoint on this variable

to predict repeat status; scores on the Narcissism scale ranged from 0 to 25 (M = 5.1, SD = 3.6), and ODA found the optimal cutpoint to predict repeat SBP was a Narcissism score of > 5.5. Using this cutpoint, ODA determined that youths that were above the mean on the Narcissism scale were more likely to exhibit repeat SBP than those who were not.

The node containing the APSD Narcissism attribute variable then terminated for the non-repeater group; At this point, no more variables met the minimum $p < .05$ standard for inclusion. This terminal node yielded 74 out a total 126 non-repeaters, or 59% correct classification. For repeaters predicted from the APSD Narcissism variable, the next node included was ethnic group with an ESS of 29%. ODA predicted that Caucasian and Latino youth would engage in repeated SBP, while African American and multi-ethnic youth would not. This finding suggests that the APSD may not be invariant across all ethnic groups. However, the classification accuracy was relatively poor. While the ethnic group variable did correctly classify 89% of non-repeaters, it correctly classified only 9 out of the 22 repeaters (41% accuracy). These terminal nodes marked the termination of the left side of the tree, representing youth with relatively low-severity index events (i.e., ≤ 4.5).

Returning to the top of the tree, the initial severity attribute predicted a lack of repeat behavior for youth with event severity > 4.5. The next included variable for these youth was their placement at time of screening, with an ESS of 80%, considered very strong. ODA predicted that youth in residential settings would be repeaters, while youth in all other settings would not. "All other

settings" included specialized foster home, foster home, group home, hospital, home of relative, shelter, on the run, and unknown. Using these criteria, the model classified 70/88 non-repeaters, for 80% accuracy for that group.

It is noteworthy that there were only three repeaters total with initial severity scores higher than 4.5, and these three cases were all in residential care (resulting in 100% accuracy of prediction). However, it would be premature to draw the conclusion that residential placement is a risk factor for repeated SBP for several reasons: 1) the relationship between initial event severity and placement is based on cross-sectional correlational data, making it impossible to determine whether characteristics of the placement or the youth are responsible, and 2) 21 children with initial event severity > 4.5 were in residential treatment at the time of screening, and only three were reported to have engaged in repeated SBP. It is also notable that out of the more severe cases (i.e., > 4.5 initial severity), only three youths were repeaters out of a total of 91 youth. This suggests either a very low rate for repeat SBP in this sample or poor detection of repeat behavior. Finally, as there was no comprehensive data on previous placements, it is unknown whether all of these youth were in residential settings at the time of the index event. Youth placements at time of screening is suggestive of their placement at the time of the index event, as the average amount of time between index event and screening was relatively low (*Md* = 235 days, *SD* = 185 days). The data do confirm, however, that out of those 21 youth in residential care with high initial event severity, 13 of those 21 were in the same (residential) placements at time of disposition and time of screening. Those 13

include the 3 repeaters, suggesting that either chronic residential placement may be iatrogenic, or that there are characteristics of these youth that make them more difficult to place in typical foster opportunities.

In summary, the initial tree model revealed four variables that in conjunction predict repeat behavior with modest accuracy: 1) Severity of SBP index event, with lower severity being positively associated with repeat SBP, 2) Narcissism for youth whose initial severity ratings were 4.5 or less, with higher scores representing higher risk for repeat SBP, 3) ethnic group for youth scoring high on Narcissism, with identification as Caucasian or Latino predicting repeat SBP, and 4) placement at screening for youth whose initial severity ratings were greater than 4.5, with residential placement predicting repeat SBP. Examination of each of the five terminal nodes suggested that the tree at this point was moderately accurate overall, but that it tended to predict aggressively, often predicting repeat SBP for non-repeaters. The ability to reject non-repeaters from the repeat SBP group, or specificity for repeat SBP, was generally unacceptable in the initial classification model. Specificity for the terminal node predicting repeat with ethnic group was 60%, while specificity for repeat SBP predicted by placement was 14%. Additionally, within the group of true repeaters, ethnic group as an attribute variable would theoretically be outperformed by a coin flip.

Pruning nodes of the initial classification model. As noted previously, three separate methods were utilized to ensure generalizability of the model and confidence in effects. The first technique of verifying LOO stability of variables was actually performed as variables were inserted in the tree to save time. All

variables were LOO-stable, suggesting that all included variables would replicate in samples from a similar population. The second method, the Bonferroni-Sidak stepwise correction, was applied to correct for the number of attribute variables in the model. All attribute variables remained in the model after stepwise correction at $p < .05$, also increasing confidence in the effects of the model. The final method of conceptual parsimony in light of minimal accuracy improvement did call for removal of ethnic group from the model as an attribute variable. As discussed, classification accuracy of the repeaters was poor with ethnic group in the model, while the accuracy of repeat prediction relying on a high Narcissism score alone would yield predictive accuracy of 71%. While ethnic group may have accurately predicted a small amount more in terms of absolute number, it did so at the expense of parsimony of the model and was not impressive in terms of its proportional addition to accuracy. Finally, Caucasian and Latino youth were a small minority in the current sample; although the finding of their higher repeat risk was beyond chance, it seems likely that these youth are atypical in a number of characteristics. Thus, after pruning, the classification tree contained four terminal nodes instead of five (see Fig. 2).

Overall performance of the final model. While the performance of the individual terminal nodes of the tree is highly informative, some summary statistics are necessary to describe overall classification performance of the model as shown (Tables 6 and 7).

Table 6

Overall Classification Accuracy

Performance index	*Performance*	*Parameter*
Overall classification accuracy	170/251	(67.7%)
Sensitivity (repeat SBP)	25/34	(73.5%)
Sensitivity (non-repeat SBP)	145/217	(66.8%)
Predictive value (repeat SBP)	25/97	(25.8%)
Predictive value (non-repeat SBP)	145/154	(94.2%)

Note: Overall classification accuracy is the percentage of individuals correctly identified by the tree model. Sensitivity is a descriptive index that indicates the percentage of the actual members of a given category (e.g., youth who exhibited repeat SBP) correctly classified by the tree model. Predictive value (or specificity) is a prognostic index that indicates the percentage of the predicted classifications into a given category that were correct.

Table 7

Cross-Classification

	Youth predicted status	
Youth actual status	*Non-repeat SBP*	*Repeat SBP*
Non-repeat SBP	145	72
Repeat SBP	9	25

The final tree still represents a relatively aggressive approach to classifying potential SBP repeaters, as evidenced by its poor predictive value for repeat status. This statistic indicates that the current model will likely tend to

generate false-positive predictions with respect to repeat SBP. Sensitivity is only fair for repeat SBP, despite the overly aggressive classification. Overall, the model performs significantly better than chance, but makes a significant number of mistakes, as well. This level of misidentification would likely be unacceptable in clinical practice.

CHAPTER 4

DISCUSSION

The current study was undertaken to investigate the potential of developing a state-of-the-art actuarial prediction model of sexual behavior problems among youth. Using both youth and caretaker report when possible, theoretically related variables were examined for predictive potential using classification tree analysis (CTA) via Optimal Data Analysis (ODA; Yarnold & Soltysik, 2005).

A number of hypotheses were proposed a priori. These hypotheses concerned the variables contained in the model and its potential structure. While CTA via ODA does not function in a traditional null hypothesis significance testing framework, the original hypotheses can be evaluated with respect to their inclusion (and lack thereof) in the classification tree. While CTA can easily be limited by low sample size, a lack of inclusion in the tree can often suggest a lack of predictive value (and thus relationship) for a given variable. Inclusion in the classification tree suggests just the opposite--strong relationships or large mean differences related to the class variable.

Hypotheses related to a model of repeat SBP

Predictions for a model focused primarily on variables that should exhibit meaningful relationships with repeat SBP based on prior empirical and theoretical work. In CTA via ODA, these hypotheses can be evaluated by simply examining which variables exhibited relationships strong and stable enough to be included in the final tree model. The initial predictions for a model of SBP included the variables of psychopathy, impulsivity, deviance, isolation, and empathy. Each of these will be addressed in turn.

Psychopathy. Self-report of psychopathy as measured by the Antisocial Process Screening Device (APSD) was found to be relevant to the classification of repeat SBP youth, but under specific conditions. The Callous-Unemotional and Impulsivity scales did not demonstrate sufficient predictive power to merit inclusion in the classification tree, but the Narcissism scale was found to be predictive of repeat SBP status for certain youth. Specifically, the classification tree predicted repeat SBP status as follows: For low index event severity (i.e., ≤ 4.5) youth with a Narcissism scale score of higher than 5.5, the model predicted repeat SBP, while a Narcissism scale score less than or equal to 5.5 resulted in a predicted classification of non-repeat status. The Narcissism scale reflects a youth's absorbing self-interest that may result in disregarding the feelings of others and acting in ways that are helpful within the scope of the youth's perception, but deleterious to peers or at least peer relationships. Some items reflect what some might call "pure" narcissism

(e.g., "I think I am better than other people") while others reflect what might be termed the behavioral outcomes of that narcissism (e.g., "I get mad when someone tells me I am doing something wrong or I get punished"). This finding is new to the SBP literature, and merits careful consideration. Considering that development of the self-report APSD is ongoing to date, some background on the APSD and changes implemented in this study may be helpful in placing the current findings in context. .

Psychopathy as currently conceptualized by Frick (2000) includes the three constructs of callous-unemotional traits, impulsivity, and narcissism. While some studies have found these factors and their complete set of suggested items to be somewhat weak in relationship to each other and relevant constructs (Marsee, Silverthorn, & Frick, 2005; Murrie & Cornell, 2002; Vitacco, Rogers, & Neumann, 2003), others have found adequate to very good validity and reliability (Caputo, Frick, & Brodsky, 1999; Kruh, Frick, & Clements, 2005; Silverthorn, Frick, & Reynolds, 2001). The aforementioned studies have been conducted within distinctly different populations (e.g., clinic-referred youth versus adjudicated youth offenders), but there is no apparent relationship between population studied and reported psychometric characteristics of the APSD.

One issue of particular relevance to the current discussion is the structure and items used in each of the studies utilizing the APSD. Interestingly, although Frick proposed a three-factor model in 2000, a number

of studies since then have continued to use a two-factor model (Silverthorn, Frick, & Reynolds, 2001; Kruh, Frick, & Clements, 2005). It is unclear if this is a result of differences in utility between the two and three-factor models for the individual research designs, or if this is related to psychometric defects that arose in the analytic process of these studies. It is notable that the only study to specifically examine the factor structure of the APSD self-report version (Vitacco, Rogers, & Neumann, 2003) found that the three-factor model as originally proposed by Frick (2000) was the only structure that provided adequate fit as a confirmatory factor model. In fact, three-factor model fit in this particular study was excellent. There is one absolute commonality between all of the studies using the self-report version of the APSD: The number of items has rarely been the same, and even more rarely have the measures contained the same items.

While the current study is not a psychometric investigation of the APSD, it is important to ascertain what characteristics of the APSD lend itself to this population (especially given the inclusion of one of its scales in the final model), and if these characteristics are simply another idiosyncrasy in a large field of idiosyncratic studies, or a meaningful distinction deserving of future study. The study-specific structure and deletion of items from the APSD is scientifically problematic, and this study may add to this quandary. somewhat. However, this study does add to the aforementioned growing evidence that the three-factor structure of psychopathy is an adequate

representation of the construct. The current study used a 16-item measure, with four items deleted because of low item-scale correlations, compared to an 18-item measure used by Vitacco, Rogers, and Neumann (2003). The included items for the Callous-Unemotional scale used in this study differ considerably from the 18-item measure, while the Narcissism scale is relatively similar in number of items and content (same number of items, slightly modified wordings). The Impulsivity scale used herein is identical to the 18-item measure proposed. Perhaps the most significant change for the current study was modifying items from starting with "you" to "I" so that the social desirability items could be seamlessly integrated. Other authors have noted the somewhat accusatory tone or face validity of the original items (Rogers et al., 2002; Vitacco, Rogers, & Neumann, 2003), and speculatively, this change may have resulted in the adequate performance (in terms of internal consistency) of all the scales, an apparently rare occurrence. This change merits further investigation in future studies.

In summary, narcissism as one component of psychopathy predicts repeat status for low initial severity youth, while the other two scales of the APSD do not adequately predict repeat status and were not included in the tree model. Narcissism seems conceptually related to any type of continued antisocial behavior, although most studies to date have instead noted the importance of callous or unemotional traits in chronic delinquency amongst adults and youth (Rice, Harris, & Quinsey, 1990; Caputo, Frick, & Brodsky,

1999). In fact, Caputo et al. (1999 unexpectedly found that callous/unemotional traits differentiated sex offenders from non-sex offenders in their sample of juvenile delinquents. These differences underscore the heterogeneity Butler and Seto (2002) noted when summarizing the body of adolescent sex offender research. The unique nature of the current sample simply makes this heterogeneity more apparent.

Speculatively, the finding that narcissism is predictive of repeat SBP for some youth in this sample may be a function of the historical background of many of the current participants. As noted in the *Method* section, while many of these children had not been sexually abused, all had suffered severe neglect at the hands of caregivers, at a minimum. These youth were all wards of the state as a result of neglect or abuse, and this created a qualitatively and quantitatively different sample than any other in the literature. Certainly, other samples of youth with SBP have included youth with abusive histories; many samples came from juvenile detention facility populations. However, the wards in the current sample were *all* from seriously dysfunctional homes, often from an inner-city urban population. Narcissism may be an adaptive coping mechanism within the original family context, given that the other likely option for abused or neglected children is to consider themselves worthless or deserving of abuse and neglect.

The point to be made here is that youth exhibiting narcissism and a low level of callous-unemotional traits may represent those who have not yet

experienced life on the street or juvenile detention, optimal breeding grounds for callous-unemotional behavior. As these youth were not adjudicated, the relatively low levels of the typically predictive callous-unemotional traits is not surprising. In fact, one positive finding is that all psychopathy factor scores were similarly low. It is possible that callous-unemotional traits do not predict repeat SBP until they reach higher levels, while narcissistic tendencies predict repeat SBP at lower levels. This may also represent a significant degree of promise, as youth who are exhibiting low levels of callous-unemotional traits may be more responsive to intervention. Obviously, the presence of narcissism is not a positive sign, but in the absence of callous-unemotional traits, the prognosis for these youth may be much better than for the youth in other samples.

Impulsivity. Impulsivity in this study was to be measured primarily by the APSD, but two important subscales from the Conners' Rating Scales, parent version, also contained impulsivity subscales and were thus included. The Impulsive-Hyperactive subscale and the overall Hyperactivity Index were included in the analyses. Consistent with results from the APSD, these Conners' subscales did not predict repeat SBP. This lends credence to the notion that impulsivity is not an adequate predictor of SBP or its recurrence above and beyond the fact that it probably differentiates delinquent, ADHD, or SBP youth from normal controls, as demonstrated in other studies (Frick, 2000). The Conners' scales were obtained from caregivers and were a different

measure from the impulsivity measure for youth (APSD), and while null results of this kind do not confirm anything, it is notable that all measures of impulsivity were in relative agreement--none predict repeat SBP adequately.

Sexual behavior and deviance. Considering the oft-noted relationship between sexual deviance and SBP (Hanson & Morton-Bourgon, 2004), it was somewhat surprising that self-reported deviance did not emerge as a predictor of repeat SBP in the classification tree. Further, the level of youth sexual behavior as reported by the caregivers of youth did not demonstrate predictive value, either. It seems likely that the modified MSI in this study is not viable to assess deviance in most populations. Few paper-and-pencil measures of deviance have successfully predicted continued SBP (Kahn & Chambers, 1991; Prentky & Knight, 1993; Prentky et al., 2000), especially when compared to the gold standard of penile plethysmography (Hanson & Bussiere, 1998).). This measurement difficulty continues to present a challenge in the field of SBP research, and it seems the search must continue for measures of deviance that are reliable and valid. This is especially relevant given that penile plethysmography has become ethically objectionable, as discussed in the *Introduction* section. There is the added problem that six females in the sample exhibited persistent SBP, and they constituted a significant proportion of the overall sample (roughly 25%). As there is no analogous test to penile plethysmography for females, this presents an additional measurement challenge.

Isolation or parental attachment. Isolation (conceptualized as a lack of parental attachment) in the current study was measured via an uncommon proxy. As the study did not include a measure of attachment, per se, the logical choice was positive parenting as measured by the Parenting Practices Measure. Positive parenting has been consistently linked in the literature to positive youth mental health outcomes, and is often inversely predictive of conduct problems (Dishion & Bullock, 2002; Florsheim, Tolan, & Gorman-Smith, 1996). Positive parenting did not emerge as a significant predictor of repeat SBP. This may be related to the nature of positive parenting practices; such behaviors accumulate effects over time, affecting developmental trajectories incrementally (Dodge & Pettit, 2003).

With more youth in the sample and a longer follow-up period, it is possible differences may have emerged by setting or other fundamental environmental variables. Despite the apparent lack of power to predict repeat SBP, there was a positive finding related to the level of positive parenting reported. Specifically, most youth (roughly 63% of the sample) reported that most positive practices happened at least "sometimes," suggesting that most youth were not particularly dissatisfied with this aspect of parenting by their caregivers.

As most delinquent youth experience dysfunctional peer relations (Dodge & Pettit, 2003), it would have been helpful if a peer isolation, rejection, or relations measure were included. Understandably, the designers

of the study that produced the data used for the current study could not include every measure of interest, especially given that the battery for youth could extend to 519 questions, depending on age and response patterns (as a branching method was used in the interview software). Future studies will hopefully be able to obtain a more ecologically valid perspective on the peer relations of youth with SBP.

Empathy. Empathy, or actually the lack thereof, was measured in the current study via the Callous-Unemotional scale of the APSD. This seemed reasonable as the creators of the APSD (Frick et al., 1994) describe the Callous-Unemotional scale as a group of items reflecting a "lack of guilt, lack of empathy, and superficial charm" (p. 704). Unfortunately, as noted under the previous discussion of psychopathy, this scale did not sufficiently predict repeat SBP to be included in the classification tree model. This was puzzling as numerous studies have established the relationship between antisocial behavior and callous/unemotional traits (Frick, 1998). This may be a result of a lack of measure validity in African American youth populations, as this measure has never been tested in a population remotely similar to the sample in the current study. This is further suggested by its inclusion in the initial model to differentiate repeaters from non-repeaters, with high scores being predictive of repeat SBP only for Caucasian or Latinos, but not for African American youth, who comprised 80% of the sample. Its classification

accuracy was admittedly weak (hence its deletion from the final model), but its initial inclusion raises the question of cultural invariance.

This scale was also theoretically problematic at the outset, and this may be partly responsible for its lack of predictive utility. An examination of the face validity of the items does yield some seemingly tangential inclusions in the scale, such as "I care about how well I do in school," "I am good at keeping promises," and "I keep the same friends." As most items in the APSD are face-valid, it seems odd that these items are included in an instrument designed to capture the construct of psychopathy amongst youth in a scale designed to measure callous or unemotional traits. Interestingly, although the inclusion of these items in their intended scale has produced three-factor model fit in a confirmatory factor study (Vitacco, Rogers, & Neumann, 2003), the inclusion of two out of three of these items yielded poor internal scale consistency in the current study. These two items (the promises and friends items) were eliminated from the current instrument as a result.

Exploratory variables examined in relation to repeat SBP

Gender. There were no differences in repeat status by gender in the current sample. While some studies have found that girls differ in onset for serious conduct problems (i.e., Conduct Disorder), once problems start, their behavior and trajectories are often similar (Silverthorn, Frick, & Reynolds, 2001; Moffitt & Caspi, 2001). There were 9 female and 28 male repeaters in the current sample. As there were three times as many boys as girls in the

sample, this equates to the same rate of offending across gender. This sample appears similar to others in that there was no significant difference between girls and boys in repeat SBP in rate or risk.

Ethnic Group. Overall, this sample is similar in ethnic composition to the larger group of all youth in substitute care in the county of data collection (Spaccarelli et al., 2001) suggesting that there is no differential risk for SBP by ethnic group for youth in the DCFS system. This does not necessarily apply to all youth with SBP, as an undetermined number may have little or no involvement with the child welfare system. In conclusion, in this sample, there appears to be no differential risk of SBP identification or repeat status as predicted by the classification tree, but larger samples of non-African American youth in the DCFS system may help elucidate this finding.

Age at disposition. Participants' age at the time of disposition (generally close to age at time of incident) was not sufficiently predictive of repeat SBP to be included in the classification model. However, as with other variables lacking in predictive power, this may be a result of a scant number of repeaters, especially in the group of youth with higher-severity initial incident scores. The three youth who were correctly classified by the tree in this higher-severity branch were aged 10, 12 and 15. The 15-year-old reportedly engaged in four separate incidents of severe SBP, the first of which was rated as 8 (assault with penetration). The two younger youth had only one repeat incident each, with slightly lower severity on repeat (although not

clinically so). The older youth's repeat behavior was more severe than his initial index event SBP. Qualitatively, the older youth and the two younger youth are quite different. Whether or not this distinction would be valid in the larger sample (and population) would unfortunately require a much higher rate of repeat behavior amongst this severe initial incident group. While discouraging, it is likely that a longer follow-up period would have yielded these additional repeaters.

Severity of initial offense. Severity of initial offense was the most powerful predictor of repeat behavior and was thus placed as the first node of the classification tree model. Severity suffers from the same limitation as other history-based measures: It is not sensitive to change. It does mark a starting point for any model, but the finding that lower-severity youth were more likely to exhibit persistent SBP is counterintuitive. One explanation is that it is likely youth who engage in contact offenses (especially those involving penetration) are placed under much greater scrutiny than those who engage in mild paraphilic behavior. There may be what some clinicians term a "fear factor" (Steve Spaccarelli, personal communication, July 28, 2006) related to these types of offenses, for which the consequences for a repeat event will likely be extremely high (i.e., criminal prosecution for the child and liability for injury to caregivers and providers). These concerns are likely not as elevated when considering behaviors such as inappropriate touching or

inappropriate propositioning. This likely interacts with placement decisions, although this is not clear (to be discussed below).

Externalizing symptoms. Externalizing symptoms were measured by the impulsivity and hyperactivity scales of the Conners', already discussed above, as well as the aggression scale of the Conners'. Likelihood of repeat SBP was not substantially linked to externalizing symptoms as measured by the above three scales. On one hand, this suggests that these youth were not so-called "versatile offenders," or youth who engaged in SBP and other delinquent behaviors. However, as there was no direct measure of delinquency, this cannot be directly confirmed.

Physical and sexual abuse history. Physical and sexual abuse as measured by both self-report and electronic records failed to demonstrate substantive predictive value for repeat SBP. This seems likely a result of a lack of repeaters in the prediction model, as there are clear differences in proportion of both physical and sexual abuse. Out of the 34 repeaters, 26 (76.5%) either reported physical abuse or physical abuse was confirmed in their electronic file history. In comparison, 166 (77.6%) of the 214 non-repeaters reported physical abuse. Sexual abuse by either self-report or electronic file history was present for 19 (55.9%) of the 34 repeaters, while this was the case for 81 (37.9%) of the 214 non-repeaters. In post-hoc chi-square analyses, only the sexual abuse differences were statistically significant. However, these differences in proportion only underscore

differences, not diagnostic accuracy. In this sense, CTA via ODA was far more stringent than reporting univariate or even multivariate linear analyses, which would likely provide significant results for many variables based on the sample size alone. The small repeater sample likely prevented accurate prediction with these variables, but as with other variables, this awaits investigation with a larger group of youth exhibiting repeat SBP.

Exposure to violence and pornography. Exposure to violence by self-report (limited to abuse suffered as opposed to witnessing) and exposure to pornography by self- and caregiver report failed to predict repeat SBP with the strength necessary for inclusion in the classification model. The sum scores for both variables were less than 1.5 for both repeaters and non-repeaters. With rates this low, differences would be very difficult to detect. It is positive news that rates of reported pornography exposure and direct experienced violence were low. The larger study included a number of items related to witnessing home and community violence, so this may be the subject of another set of analyses in the future.

Placement at screening. Placement at screening was predictive of repeat status in the final model, but only for youth with relatively severe initial events scores (> 4.5). Elevated concern may have led to the repeating youths' continued placements in residential facilities, but high initial severity and the accompanying support do not seem to be associated with placement, as there were 89 youth with initial severity > 4.5, and only 18 of these were placed in

residential care. This indicates that the occurrence of severe events in all settings is quite low. Repeaters with low (≤ 4.5) severity index event scores were in a variety of settings at the time of the initial offense. Qualitative analysis of placement *after* the index event (not included in data set because of the high number of placements per child) suggests that almost without exception, if an SBP youth was not in a residential-type setting (this includes group homes, hospital, and incarceration) at the time of initial disposition, they would be soon after. This suggests that the typical response to an SBP event by DCFS is to move to a more restrictive setting, which is logical given the behavior in many cases. However, what is troubling is that a large number of repeaters (27 out of 34) were moved to residential facilities (if not already in such facilities) and tended to stay in such placements. This implies that DCFS encountered significant difficulty in finding a permanent placement for these children (evidenced by their average of 8 placements per child), or that DCFS felt the risk was too high to attempt placement in another foster family after the index event.

Overall, this raises the question of iatrogenic effects: Would these repeaters have engaged in repetitive SBP if not in residential-type facilities after the index event? It also is possible that residential facilities are more vigilant with regard to SBP, leading to the appearance of iatrogenic effects that are simply differential rates of detection. A future study should undertake the process of coding the substantial placement data from this study to

examine specific patterns of placement after the initial event, as this may yield more information than the variable of initial placement used in this study.

Self-report of drug use. Drug use did not emerge as a predictor of repeat SBP in the classification tree model. The mean for the sum score was 11.5, with a standard deviation of 4.2. Sum scores on the drug total could range from 0 to 54. Most of the sample reported infrequent (tried once or use occasionally) use of alcohol and marijuana, and the vast majority reported no experience with other substances (e.g., cocaine, heroin). 10% of the sample reported regular use of tobacco, and 5% reported regular use of alcohol. These rates are somewhat low relative to what would be expected for this population, but as the mean age was relatively young at time one, time two data may tell a different story.

Duration of time between time one and time two. This variable did not emerge as a predictor in the final tree model. Although repeat behavior was assessed via official electronic records at the end of the study, there was a remote chance that youth with a longer follow-up period would be detected more easily. Youths could have entered the study at any time during a several-year period, as the study was designed to continue for an extended period of time and could consider examining individual cohorts. However, this sample represents a reasonable cohort as the average duration of time between times one and two was roughly a year. This did vary from six months to three years, but most youth had between 12 and 24 months of follow-up time. The lack of

inclusion of this variable suggests that this follow-up duration had little, if any, effect on detection of repeat status.

Implications

The results of this research do suggest that clinicians need to take into account the initial severity of the act, but perhaps also not become complacent that more minor-severity behaviors will nor recur or be problematic in the long run given that youth with lower-severity SBP in this sample exhibited higher rates of repeat SBP. Perhaps if the same degree of support were directed toward all youth with identified SBP, recurrence would drop significantly.

It is likely that youth with SBP respond to the same interventions as conduct problems. If this is so, it would be crucial to consider this support in terms of parental monitoring and its effect on recurrent serious behavior. There is ample evidence that parents who use positive skills, monitor vigilantly, and use discipline consistently can reduce instances of undesirable behavior by large margins (Dishion & Bullock, 2002; Dodge & Pettit, 2003).

The youth in this sample exhibited little documented repeat SBP in the follow-up period of the study. It would be premature given the limited follow-up time to declare this population at low risk for recurrence. It does, however, underscore the need to appreciate the variability in developmental course for many of these youth. Most of these youth were receiving some form of treatment, and this may be partly responsible for the low recurrence rate.

While adults with SBP are often given a poorer prognosis (especially child molesters), it seems likely that with all child and adolescent conditions, there is considerable room for change, especially at early ages.

It is possible that placement has a great deal to do with recurrence of SBP, but this study poses the question of whether a more restrictive environment is better for youth who exhibit index offenses of lower severity. The need for intervention is just as pressing as with higher severity events, but a more restrictive environment may pose more risk than benefit for youth who can be exposed to other youth with more psychopathology and delinquent behavior. Departments responsible for child and family services may wish to consider more carefully the allocation of these highly restrictive and expensive environments with these lower-severity youth.

Another issue concerns the measurement of sexual deviance and sexual behavior in general. As pointed out in the introduction, few studies have found reliable and valid measures of sexual deviance other than penile plethysmography. It is possible that the lack of a measure of youth sexual deviance for caregivers resulted in the null results for deviance in this study. The caregiver measure of sexual behavior was designed to be compared to norms for various age groups, because different items reflect behaviors that may be appropriate at some developmental stages and inappropriate at others. There is little evidence for caregiver-reported youth hypersexuality (as seen among many adult sex offenders) in this study, but it is possible that the un-

normed sum score as used here is insensitive to age-related differences. It is likely that self-report will continue to be elusive in any population where delinquency is high, as is often the case in ward-of-the-state populations. However, caregiver report of deviance deserves further study.

This study also indirectly examined the sexual reactivity hypothesis. Specifically, it has often been stated (e.g., Ryan, 1997c) that SBP can usually be considered hallmarks of sexual abuse. Definitions may differ with respect to SBP, but some clinicians such as Ryan consider all SBP to be "sexually reactive behavior" as evidenced by their proposed models of SBP that inevitably include sexual abuse. This implies that all youth exhibiting SBP must have been sexually abused at some point. This notion is problematic because only 40% of the current sample had a history of sexual abuse by electronic records or self-report. While it is possible that electronic records of sexual abuse are insensitive to the true frequency of sexual abuse, there seems to be little reason to believe this is the case with self-report. The youth in this sample were apparently open and forward about sexual issues and a number of traits that might normally cause guilt or shame among youth, as evidenced by the rates of endorsement of various items on deviance and their willingness to discuss details of abuse. This is probably likely related to these children being required to discuss incidents and sexuality repeatedly with caseworkers and therapists. Further, social desirability did not correlate with crucial measures that would indicate response bias. Thus, if self-report of sexual abuse can be

regarded as reasonably accurate, there was a relatively low proportion of sexual abuse in this sample to assert that the reactivity hypothesis is true. This is crucial for therapists who may attempt to probe for sexual abuse (that may not be in a youth's history) simply on the basis of the occurrence of SBP. SBP should not be considered a hallmark of sexual abuse, and clinical judgment may cloud the accuracy of this decision considerably. Finally, youth have proven suggestible in interviewing circumstances regarding sexual abuse, and so probing for sexual abuse should probably be undertaken only when there is suspicion based on corroborating evidence of sexual abuse as opposed to SBP alone.

While the hypothesis that all SBP represent sexual reactivity is not supported by the current data, this does not mean that some youth who experience sexual abuse will not exhibit SBP. These results simply underscore the caution necessary in presuming that a child exhibiting SBP may have been sexually abused. In short, SBP is not synonymous with sexual abuse history.

Limitations and recommendations for future study

The primary limitation of this study is the same as most others: the follow-up period was too short to truly gauge risk for all participants in the sample. This flaw was not by design in the current case; DCFS (the funding organization) opted not to continue funding the research. This is especially unfortunate given that one of the primary questions of any child welfare

agency is which children are at greatest risk for psychological disorders and their associated service needs.

Another issue related to lack of follow-up is the limited number of repeaters available for study as participants. As discussed in the *Results* section, a number of variables demonstrated an initially modest relationship but were not included in the final tree when other variables of higher effect strength preceded them. This does not mean these variables hold no explanatory value; in fact, it is likely that these variables could enhance the specificity of prediction if a larger sample had been available for classification. Another study with at least as large a sample and more follow-up time would probably allow for more nodes in the model and thus would allow finer-grained distinctions between repeaters and non-repeaters of SBP.

This study also suffered limitations based in the modifications of instrumentation for the study. This was unavoidable in some cases, as no measures existed for the construct in the age group of concern. However, there is some question as to whether these results would be exactly the same if certain characteristics were changed, such as the scaling of the measures and wording of items. Including common measures with parallel multi-informant versions would better approximate the state-of-the-art in evidence-based assessment and would yield results interpretable across studies.

A final limitation concerns the sample itself. The youth in this sample represent the culmination of tremendous effort on the part of various

individuals and agencies to collect these data. The study of this population and SBP is unprecedented in the literature. While this represents a crucial first step to better understanding SBP in seriously disadvantaged populations, it also makes this sample considerably different from all other samples of youth with SBP that have relied on adjudicated samples of different ethnic make-up, clinic-referred samples of high socioeconomic strata, or adult offenders using retrospective report or file review. This difference makes it difficult to ascertain why certain striking differences (such as the lack of callous/unemotional traits to predict repeat behavior) are present in this sample and not others. Most of the children in this sample had suffered physical abuse that sometimes bordered on lethal, and virtually all were victims of neglect prior to DCFS intervention. While their history may be similar to many adjudicated populations, this study represents an opportunity to understand youth with SBP before they have moved down that unfortunate road. It is not clear how these processes affect these youth and their emerging SBP, and this is best answered with longitudinal data with numerous time points.

A number of improvements can be made to improve the predictive power of this research. Future studies should include longer follow-up periods, more time points, well-developed unmodified measures whenever possible, and coded data from various databases concerning specifics of the placement. The context in which these youth spend most of their time would

seem to be crucial to their development, and understanding this environmental context better should yield impressive insight on risk factors for recurrent SBP. It is notable that one of the most intriguing findings of this study was related to placement after the event, which was not readily available for all youth due to the need for comprehensive coding. Understanding the nature of their various placements and the quality of the environments with respect to therapy, social milieu, basic needs, and safety (to name only a few factors) will likely greatly improve our understanding of risk for persistent SBP among youth in the child welfare system, and perhaps all youth with SBP.

These recommendations are feasible, but state departments of child welfare must consider comprehensive research a funding priority for many of these recommendations to be implemented. Considering the monetary expenditures on therapeutic services each year from these agencies, it would seem that research that better targeted those services could be extremely cost-effective in very little time. This commitment to funding must be for the extended duration, which can be a challenge given that funding priorities change with state administrations and the directors of child welfare services. However, methods such as those used in this study can undoubtedly lend considerable utility to clinical decision-making for youth with SBP, an area that continues to be sorely lacking. With ultimate clinical outcomes of these youth in mind, funding long-term studies of SBP seems to become less of an

optional decision and more of a question of the timing of money spent: adjudication and treatment later, or effective treatment now.

APPENDIX A

FIGURES

FIGURE 1

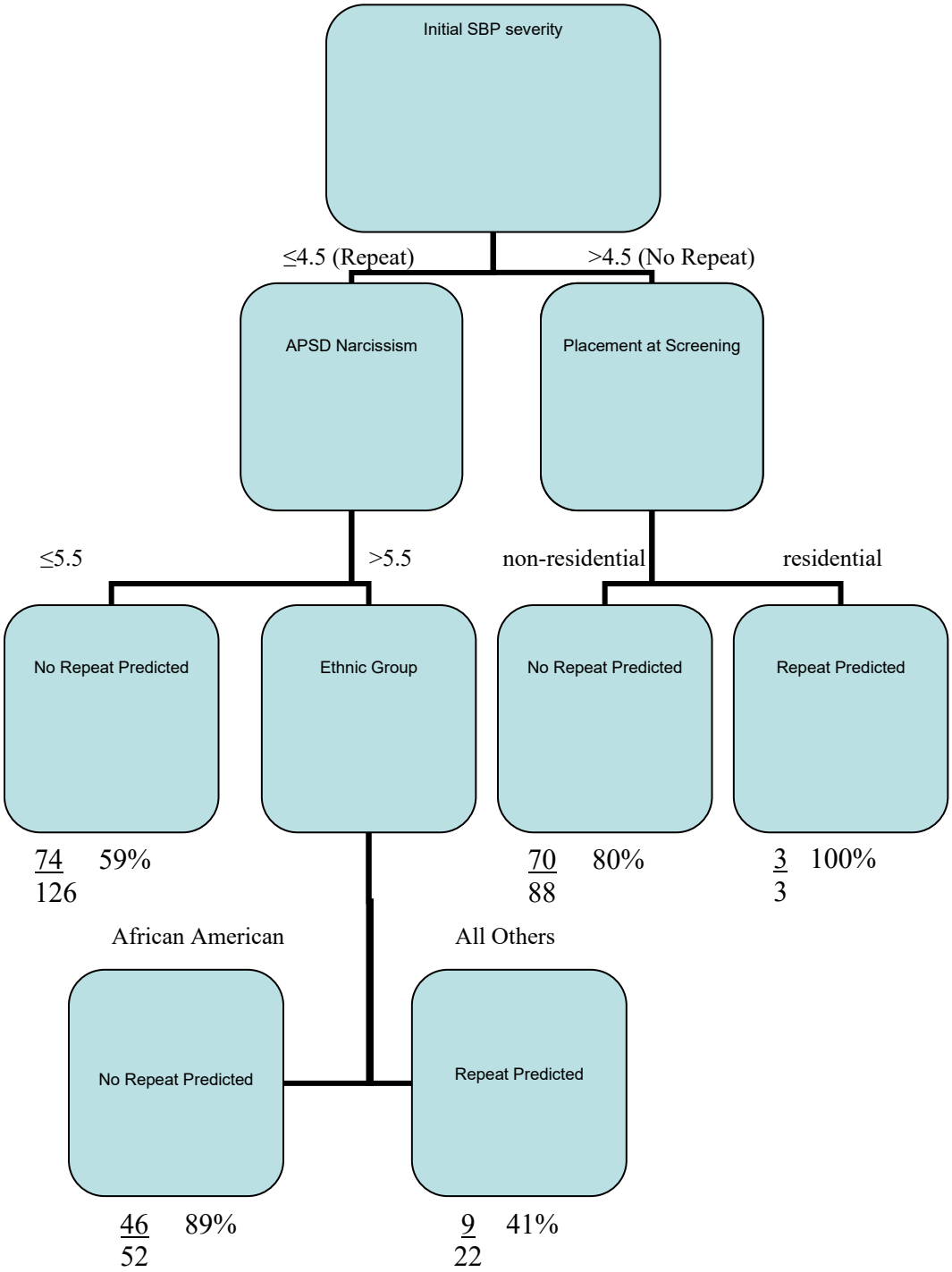

Fig.1. The Optimal Data Analysis (ODA) classification tree model for prediction of repeat sexual behavior problems (*N* = 251). All nodes of the tree were leave-one-out (LOO) stable and significant at p<.05 after Bonferroni-Sidak correction for number of analyses performed. Cutpoints on continuous variables are provided below the node on its respective branches, representing the value selected by ODA that yields optimal classification accuracy. Fractions below endpoints of the tree represent number of correct classifications out of all individuals in the class, with those fractions as percentages of correct classification directly beside them.

FIGURE 2

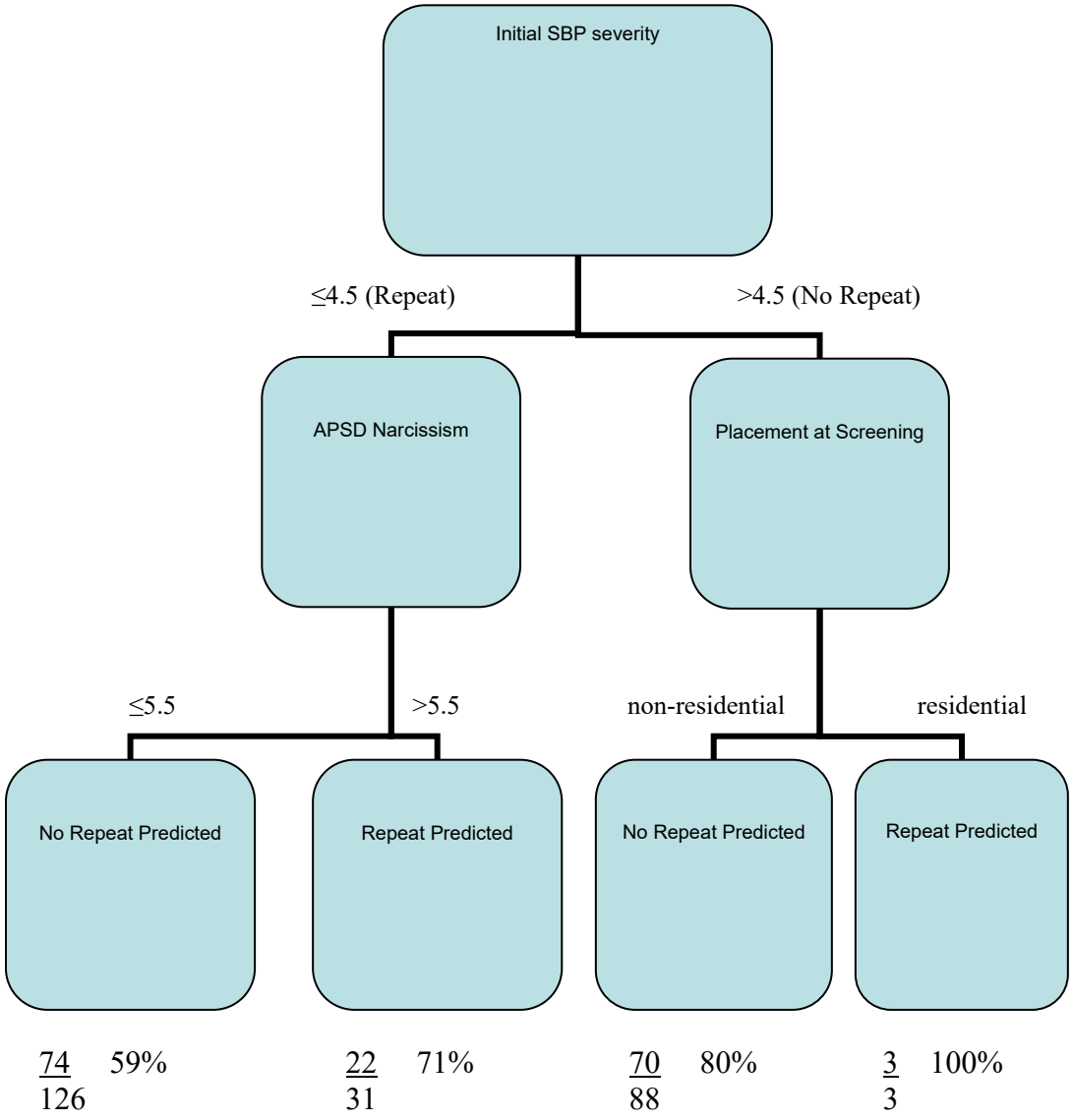

Fig.1. The Optimal Data Analysis (ODA) classification tree model for prediction of repeat sexual behavior problems (*N* = 251). All nodes of the tree were leave-one-out (LOO) stable and significant at p<.05 after Bonferroni-Sidak correction for number of analyses performed. Cutpoints on continuous variables are provided below the node on its respective branches, representing the value selected by ODA that yields optimal classification accuracy. Fractions below endpoints of the tree represent number of correct classifications out of all individuals in the class, with those fractions as percentages of correct classification directly beside them.

REFERENCES

Ageton, S. (1983). *Sexual assault among adolescents*. Lexington, MA: Heath.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed, text revision)*. Washington, DC: Author.

American Psychological Association. (2002). *Ethical principles of psychologists*. Washington, D.C.: Author.

Awad, G. A., & Saunders, E. B. (1991). Male adolescent sexual assaulters: Clinical observations. *Journal of Interpersonal Violence, 6*, 446-460.

Awad, G. A., Saunders, E. B., & Levene, J. (1984). A clinical study of male adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 28,* 105-115.

Barth, R., & Berry, M. (1989). Child abuse and welfare services. In Policy Analysis for California Education (PACE). In R. Barth & M. Berry (Eds.), *Conditions of children in California* (pp. 225-256). Berkeley, CA: PACE.

Becker, J. V., Kaplan, M. S., Cunningham-Rathner, J., & Kavoussi, R. (1986). Characteristics of adolescent incest sexual perpetrators: Preliminary findings. *Journal of Family Violence, 1*, 85-97.

Bowcott, O., & Clouston, E. (1997). Nightmare on any street. *The Guardian*, *June 10, G2*, pp. 1-3.

Bradley, J. V. (1968). *Distribution-free statistical tests*. Englewood Cliffs, NJ: Prentice-Hall.

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., Fergusson, D., Horwood, J. L., Loeber, R., Laird, R., Lynam, D. R., Moffitt, T. E., & Pettit, G. S. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology, 39*, 222-245.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513-531.

Butler, S. M., & Seto, M. C. (2002). Distinguishing two types of adolescent sex offenders. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*, 83-90.

Caputo, A. A., Frick, P. J., & Brodskey, S. L. (1999). Family violence and juvenile sex offending: The potential mediating role of psychopathic traits and negative attitudes toward women. *Criminal Justice and Behavior, 26,* 338-356.

Centers for Medicare and Medicaid Services. (2004). Retrieved October 17, 2004, from http://www.cms.hhs.gov/olmstead/default.asp.

Conners, C. K. (1990). *Conners' parent rating scales (PRS)*. Eagan, MN: Pearson Assessments.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues in field settings*. Chicago: Rand-McNalley.

Dawes, R. (1996). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.

Dishion, T. J., & Bullock, B. M. (2002). Parenting and adolescent behavior: An ecological analysis of the nurturance hypothesis. In J. G. Borkowski, S. L. Ramey, & M. Bristol-Power (Eds.), *Parenting and the child's world: Influences on academic, intellectual, and social-emotional development.* (Monographs in Parenting Series, pp. 231-249). Mahwah, NJ: Erlbaum.

Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology, 39,* 349-371.

Donenberg, G., Bryant, F. B., Emerson, E., Wilson, H. W., & Pasch, K. E. (2003). Tracing the roots of early sexual debut among adolescents in psychiatric care. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 594-608.

Doyle, S. R., & McCarty, C. A. (2000). *Psychopathy screening device* (technical report). Retrieved July 20, 2006 from http://www.fasttrackproject.org/techrept/p/psd/ psd8tech.pdf

Edmond, T., Auslander, W., Elze, D. E., McMillen, C., & Thompson, R. (2002). Differences between sexually abused and non-sexually abused adolescent girls in foster care. *Journal of Child Sexual Abuse, 11*, 73-99.

Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology, 37,* 90-93.

Evans, G. W. (2004). The environment of child poverty. *American Psychologist, 59*, 77-92.

Fehrenbach, P. A., Smith, W., Monastersky, C., & Deisher, R. W. (1986). Adolescent sexual offenders: Offender and offense characteristics. *American Journal of Orthopsychiatry, 56*, 225-233.

Feinglass, J., Yarnold, P. R., Martin, G. J., & McCarthy, W. J. (1998). A classification tree analysis of selection for discretionary treatment. *Medical Care, 36*, 740-747.

Feldesman, M. R. (2002). Classification trees as an alternative to linear discriminant analysis. *American Journal of Physical Anthropology, 119*, 257-275.

Florsheim, P., Tolan, P. H., & Gorman, D. (1996). Family processes and risk for externalizing behavior problems among African American and Hispanic boys. *Journal of Consulting and Clinical Psychology, 64*, 1222-1230.

Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *The Hare Psychopathy Checklist: Youth Version (PCL-YV)*. Toronto: Multi-Health Systems.

Frick, P. J. (1998). *Conduct disorders and severe antisocial behavior*. New York: Plenum.

Frick, P. J. (2000). The problems of internal validation without a theoretical context: The different conceptual underpinnings of psychopathy and the disruptive behavior disorder criteria. *Psychological Assessment, 12*, 451-456.

Frick, P. J., O'Brien, B. S., Wootton, J. M., & McBurnett, K. (1994). Psychopathy and conduct problems in children. *Journal of Abnormal Psychology, 103*, 700-707.

Friedrich, W. N., Fischer, J. L., Dittner, C. A., Acton, R., Berliner, L., Butler J., Damon, L., Davies, W. H., Gray, A., & Wright, J. (2001). Child Sexual Behavior Inventory: Normative, psychiatric, and sexual abuse comparisons. *Child Maltreatment, 6*, 37-49.

Friedrich, W. N., Grambsch, P., Damon, L., Hewitt, S. K., Koverola, C., Lang, R. A., Wolf, V., Broughton, D. (1992). Child Sexual Behavior Inventory: Normative and clinical comparisons. *Psychological Assessment, 4*, 303-311.

Garb, H. (1998). *Studying the clinician: Judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996). Clinical versus actuarial predictions of violence in patients with mental illnesses. *Journal of Consulting and Clinical Psychology, 64*, 602-609.

Goddard, C., & Saunders, B. J. (2001). *Child abuse and the media.* Child Abuse Prevention Issues, 14.

Gorman-Smith, D., Henry, D. B., & Tolan, P. H. (2004). Exposure to community violence and violence perpetration: The protective effects of family functioning. *Journal of Clinical Child and Adolescent Psychology, 33*, 439-449.

Gorman-Smith, D., Tolan, P. H., Zelli, A., & Huesmann, R. L. (1996). The relation of family functioning to violence among inner-city minority youths. *Journal of Family Psychology, 10*, 115-129.

Gretton, H. M. (1998). *Psychopathy and recidivism in adolescence: a ten-year retrospective follow-up*. Unpublished doctoral dissertation, University of British Columbia, Vancouver, Canada.

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

Grotevant, H. D. (1997). Identity processes: Integrating social psychological and developmental approaches. *Journal of Adolescent Research, 12*, 354-357.

Grove, W. M., Zald, D. H., LeBow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12,* 19-30.

Grubin, D. (1999). Actuarial and clinical assessment of risk in sex offenders. *Journal of Interpersonal Violence, 14*, 331-363.

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*, 348-362.

Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. Public Works and Government Services Canada: author.

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

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*, 315-335.

Holland, B. S., & Copenhaver, M. D. (1987). An improved sequentially rejective Bonferroni procedure. *Biometrics, 43*, 417-423.

Illinois Department of Children and Family Services (2003). *Budget request briefing: Fiscal year 2004*. Springfield, IL: author.

Jones, K. A., Smith, N. C., & Holmes, P. S. (2004). Anxiety symptom interpretation and performance predictions in high-anxious, low-anxious, and repressor sport performers. *Anxiety, Stress, and Coping, 17,* 187-199.

Kahn, T. J., & Chambers, H. J. (1991). Assessing reoffense risk with juvenile sexual offenders. *Child Welfare, 70*, 333-345.

Kaslow, N. J., Thompson, M. P., Brooks, A. E., & Twomey, H. B. (2000). Ratings of family functioning of suicidal and nonsuicidal African American women. *Journal of Family Psychology, 14*, 585-599.

Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence*. Thousand Oaks, CA: Sage.

Klockars, A. J., Hancock, J. R., & McSweeny, M. J. (1995). Power of unweighted and weighted versions of simultaneous and sequential multiple comparison procedures. *Psychological Bulletin, 118*, 300-307.

Knight, R. A., & Prentky, R. A. (1993). Exploring characteristics for classifying juvenile sex offenders. In H. E. Barbaree, W. L. Marshall, & S. M. Hudson (Eds.), *The juvenile sex offender* (pp. 45-83). New York: The Guilford Press.

Kruh, I. P., Frick, P. J., & Clements, C. B. (2005). Historical and personality correlates to the violence patterns of juveniles tried as adults. *Criminal Justice and Behavior, 32*, 69-96.

Lane, S. (1997). The sexual abuse cycle. In G. Ryan & S. Lane (Eds.), *Juvenile sexual offending: Causes, consequences, and correction* (pp. 77-121). San Francisco: Jossey-Bass.

Lane Council of Governments. (2003). *Managing sex offenders in the community: A national overview*. Eugene, OR: author.

Levine, M., Doueck, H. J., Freeman, J. B., & Compaan, C. (1996). African American families and child protection. *Child and Youth Services Review, 18*, 693-711.

Lynam, D. R. (1996). Early identification of chronic offenders: Who is the fledgling psychopath? *Psychological Bulletin, 120*, 209-234.

Margolin, G. (2005). Children's exposure to violence: Exploring developmental pathways to diverse outcomes. *Journal of Interpersonal Violence, 20*, 72-81.

McLoyd, V. C. (1998). Socioeconomic advantage and child development. *American Psychologist, 53*, 185-204.

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

Marsee, M. A., Silverthorn, P., & Frick. P. J. (2005). The association of psychopathic traits with aggression and delinquency in non-referred boys and girls. *Behavioral Sciences and the Law, 23*, 803-817.

Miller, T. R., Cohen, M. A., & Wiersema, B. (1996). *Victim costs and consequences: A new look*. Washington, D. C.: National Institute of Justice.

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course persistent and adolescent-limited antisocial pathways among males and females. *Development and Psychopathology, 13*, 355-375.

Murrie, D. C., & Cornell, D. G. (2002). Psychopathy screening of incarcerated juveniles: A comparison of measures. *Psychological Assessment, 14*, 390-396.

National Council of Juvenile and Family Court Judges. (1984). The juvenile court and serious offenders. *Juvenile and Family Court Journal (special issue).*

Nichols, H., & Molinder, I. (1984). *Manual for the Multiphasic Sex Inventory.* Tacoma, WA: Criminal and Victim Psychology Specialists.

Ostrander, R., Weinfurt, K. P., Yarnold, P. R., & August, G. (1998). Diagnosing attention deficit disorders using the BASC and the CBCL: Test and construct validity analyses using optimal discriminant classification trees. *Journal of Consulting and Clinical Psychology, 66*, 660-672.

Perverted-justice.com. (2006). Retrieved July 28, 2006 from http://www.perverted-justice.com.

Prentky, R. A., Harris, B., Frizzell, K., & Righthand, S. (2000). An actuarial procedure for assessing risk with juvenile sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 12*, 71-86.

Prentky, R. A., & Knight, R. A. (1993). Age of onset of sexual assault: Criminal and life history correlates. In G. C. N. Hall, R. Hirschman, J. R. Graham, & M. S. Zaragoza (Eds.), *Sexual aggression: Issues in etiology, assessment, and treatment* (pp. 43–62). Washington, DC: Taylor & Francis.

Prentky, R. A., Lee, A. F. S., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior, 21*, 635-659.

Prentky, R. A., Knight, R. A., Sims-Knight, J. E., Straus, H., Rokous, F., & Cerce, D. (1989). Developmental antecedents of sexual aggression. *Development and Psychopathology, 1*, 153-169.

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

Proctor, B. D., & Dalaker, J. (2003). *Poverty in the United States: 2002*. Washington, D.C.: United States Government Printing Office.

Rasmussen, L. A. (1999). Factors related to recidivism among juvenile sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 11*, 69-85.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crown Social Desirability Scale. *Journal of Clinical Psychology, 38*, 119-125.

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

Rogers, C., & Tremaine, T. (1984). Clinical intervention with boy victims of sexual abuse. In S. Greer & I. R. Stuart (Eds.), *Victims of sexual aggression: Men, women, and children* (pp. 26-42).). New York: Van Nostrand Reinhold.

Rogers, R., Vitacco, M. J., Jackson, R. L., Martin, M., Collins, M., & Sewell, K. W. (2002). Faking psychopathy? An examination of response styles with antisocial youth. *Journal of Personality Assessment, 78*, 31-46.

Rosenhan, D. L. (1973). In being sane in insane places. *Science, 179*, 250-258.

Ryan, G. (1997a). Sexually abusive youth: Defining the population. In G. Ryan & S. Lane (Eds.), *Juvenile sexual offending: Causes, consequences, and correction* (pp. 3-9). San Francisco: Jossey-Bass.

Ryan, G. (1997b). Incidence and prevalence of sexual offenses committed by juveniles. In G. Ryan & S. Lane (Eds.), *Juvenile sexual offending: Causes, consequences, and correction* (pp. 10-18). San Francisco: Jossey-Bass.

Ryan, G. (1997c). Theories of etiology. In G. Ryan & S. Lane (Eds.), *Juvenile sexual offending: Causes, consequences, and correction* (pp. 19-35). San Francisco: Jossey-Bass.

Ryan, T. A. (1959). Multiple comparisons in psychological research. *Psychological Bulletin, 56*, 26-47.

Sattler, J. M. (2001). *Assessment of children: Behavioral and clinical applications, (4th ed.).* La Mesa, CA: author.

Seto, M. C., Lalumière, M. L., & Blanchard, R. (2000). The discriminative validity of a phallometric test for pedophilic interests among adolescent sex offenders against children. *Psychological Assessment, 12*, 319-327.

Shaw, J. A., & The Workgroup on Quality Issues. (1999). Practice parameters for the assessment and treatment of children and adolescents who are sexually abusive of others. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 55S-66S.

Showers, J., Farber, E. D., Joseph, J. A., Oshins, L., & Johnson, C. F. (1983). The sexual victimization of boys: A three-year survey. *Health Values: Achieving High-Level Wellness, 7*, 15-18.

Silovsky, J. F., & Niec, L. (2002). Characteristics of young children with sexual behavior problems: A pilot study. *Child Maltreatment, 7*, 187-197.

Skilling, T. A., Quinsey, V. L., & Craig, W. M. (2001). Evidence of a taxon underlying serious antisocial behavior in boys. *Criminal Justice and Behavior, 28*, 450-470.

Snyder, H. N. (2000). *Sexual assault of young children as reported to law enforcement: Victim, incident, and offender characteristics*. Washington, D. C.: Bureau of Justice Statistics, U. S. Department of Justice.

Spaccarelli, S., Bowden, B., Coatsworth, J. D., & Kim, S. (1997). Psychosocial correlates of male sexual aggression in a chronic delinquent sample. *Criminal Justice and Behavior, 24*, 71-95.

Spaccarelli, S., Edejer, Bushell, Karaitis, & Jones, (2001). *Child sexual behavior problems longitudinal study codebook*. Chicago, IL: author.

Straus, M. A. (2006). *Scoring the CTS2 and the CTSPC*. Retrieved on July 20, 2006 from http://pubpages.unh.edu/%7Emas2/CTS28a3.pdf

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised Conflict Tactics Scales (CTS2): Preliminary and psychometric data. *Journal of Family Issues, 17,* 283-316.

Swan, T. (1997). Problems in caring for sexually abused girls: Care providers speak out. *Community Alternatives, 9*(1), 71-87.

Szasz, T. S. (1963). *The manufacture of madness*. New York: Harper & Row.

To catch a predator. (2006). Retrieved July 28, 2006 from http://www.msnbc.msn.com/id/10912603/.

Vitacco, M. J., Rogers, R., & Neumann, C. S. (2003). The Antisocial Process Screening Device: An examination of its construct and criterion-related validity. *Assessment, 10,* 143-150.

Worling, J. R., & Curwen, T. (2000). Adolescent sexual offender recidivism: Success of specialized treatment and implications for risk prediction. *Child Abuse & Neglect*, *24,* 965–982.

Yarnold, P. R. (1996). Discriminating geriatric and non-geriatric patients using functional status information: An example of classification tree analysis via UniODA. *Educational and Psychological Measurement, 56*, 656-667.

Yarnold, P.R. & Soltysik, R.C. (2005). *Optimal data analysis: A guidebook with software for windows*. Washington, DC: American Psychological Association Books.

VITA

David J. Morgan was born in Atlanta, Georgia and lived there until moving to Chicago, Illinois to attend graduate school in Child-Clinical Psychology at Loyola University Chicago. His research interests are primarily in the area of disruptive behavior disorders among children and the interface between neuropsychological profiles of these youth and effective treatment informed by these profiles. He received his clinical training in the variety of settings, including the Loyola University Chicago Counseling Center, a school in the Englewood area of South Chicago, a disruptive behavior disorders clinic, and a neuropsychology service at a major area hospital. He performed his general child internship at a residential, therapeutic day school, and outpatient treatment facility in the suburbs of Chicago, Illinois.