# REFINING THE CONSTRUCT OF SCHOOL SAFETY: AN EXPLORATION OF CORRELATES AND CONSTRUCT VALIDITY OF SCHOOL SAFETY MEASURES

Travis C. Laxton

Jeffrey R. Sprague

University of Oregon Institute on Violence and Destructive Behavior

This paper was partially supported by the Hamilton Fish Institute on School and Community Violence, through Award No. 2005-JL-FX-0157, awarded by the Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice. Points of view or opinions in this document are those of the author and do not necessarily represent the official position of policies of the U.S. Department of Justice or The George Washington University.

**Abstract**

Although school violence has long been a problem in American public schools, recent instances of nationally publicized school shootings have pushed this issue and the need for school safety, into the focus of national attention. Parallel to and often a product of this national attention a growing body of research focused on the development and etiology of youth antisocial behavior has developed. However, its primary focus has been on a variety of individual, family, and community factors. Conspicuously absent from our knowledge is a clear understanding of the role that the school and classroom environment play in the development and exacerbation of antisocial behavior. Similarly, little systematic attention has been paid to the measurement of school climate and the relationships between school climate indices and school disorder and violence. Consequently, little is known on the topic of measuring school climate and safety. The present study explored the internal consistency of and relationships between four measures including: The Oregon School Safety Survey (OSSS), the Effective Behavior Support Survey 1.5 (EBS), The School-Wide Evaluation Tool (SET), and the Oregon School Climate and Safety Survey (OSCSS). Alpha coefficients were computed for each of the scales of the above four measures. The majority of scales demonstrated acceptable reliability, but some were unacceptably low. Exploratory factor analyses were conducted on all of the above measures with the exception of the OSCSS. The OSSS demonstrated a clear two-factor orthogonal solution, the EBS demonstrated a clear four-factor oblique solution, and the SET demonstrated a four-factor oblique solution that was not clearly interpretable. Each of the derived factors was correlated. Several of the factors of the EBS demonstrated significant correlations with SET and OSSS factors. The SET and OSSS did not correlate significantly across any factors. Finally, using the OSCSS scales, three measures of school safety/disorder were predicted using standard multiple regression procedures. Significant multiple regression equations were found for each of the three measures of school safety/disorder including: Student Victimization, Student Rebellious Behavior, and Drug Availability.

Table of contents

Chapter Page

List of Tables

**Introduction**

Although widely considered an American problem, school shootings have become an international epidemic affecting such varied countries as Scotland, America, Yemen, Germany, Sweden, Canada, and most recently Bosnia-Herzegovina. In the United States school violence has been pushed to the forefront of national attention in the wake of such dramatic and nationally publicized school shootings as Thurston High School in May 1998 and Columbine High School in April 1999. The most recent school shooting occurred at Red Lake High School in Minnesota on March 21st 2005 leaving 10 dead including the perpetrator (CBSNews.com, 2005). Consequently school safety has become a primary concern to parents, schools, and the society at large (Gallup Poll Analysis, 2001). Similarly, students express concern for their own safety in school. In a recent survey of middle and high school students more than one in three said they don't feel safe at school (National School Safety Center, 2001).

*Social and Political Responses*

Despite the recent increase in public interest, school violence has been a concern of policy makers and researchers since the early 1970s. A review of national interest in and social policy toward school violence from the 1970s to the present will provide an historical understanding of the school violence phenomenon in American public schools.

In 1971 the Senate Judiciary subcommittee to investigate juvenile delinquency commissioned a study of school violence and vandalism. The resulting 1975 report entitled, "Our nation's schools-a report card: 'A' in school violence and vandalism", portrayed many American public schools as dangerous places and sparked the concern of national policy makers (Bay, 1975). Three years later the National Institute of Education completed a landmark national study of school violence and victimization (National Institute of Education, 1978). This report published some of the first shocking statistics on violence in American public schools indicating that nearly 282,000 students and 5,200 teachers were physically assaulted in secondary schools each month (Elliot, Hamburg & Williams, 1998). In response to these revelations of the violent state of many American public schools and the concern that they invoked, national policy makers drafted Federal initiatives intended to reduce juvenile delinquency and promote safe and orderly schools.

The first of these was the Juvenile Justice and Delinquency Prevention Act of 1974 (JJDP). The JJDP Act was intended primarily to reform the juvenile justice system's procedures for dealing with and addressing the needs of youth. The Act also included a block grant program to distribute Federal funds to the states. The purpose of these funds was to help in implementing the congressionally established standards, and to support new programs designed to prevent and control delinquency and to improve the juvenile justice system (Governor's Juvenile Justice Advisory Committee, 1999). Although the JJDP Act was not specifically focused on schools and the delinquency and violence problems they faced it did lay the foundation for a new approach to juvenile delinquency and violence. More specifically, the Act changed the ways in which society addressed and understood the problems of juvenile delinquency focusing more on the prevention of and diversion away form a criminal lifestyle rather than punishment and incarceration after the fact.

In the 1992 reauthorization of the JJDP Act, Congress further reinforced the need for and the ability to engage in delinquency prevention through the addition of the Title V Community Prevention Grants Program. Title V established incentive grants for local delinquency prevention programs intended to fund collaborative, community-initiated and community-based delinquency prevention efforts. The Community Prevention Grants Program was founded on a research-based framework that focuses on the reduction of risk and the enhancement of protective factors. It is guided by six fundamental principals intended to prevent youth from entering into the juvenile justice system (Hsia & Bownes, 1998). Title V further enhanced the ability of and emphasized the need for local community involvement in controlling and preventing youth delinquency and violence.

Parallel to the efforts focused on community-based delinquency prevention were efforts focused on community- and school-based drug abuse prevention. In response to concerns about rising youth drug abuse Congress enacted the Drug-Free Schools and Communities Act (DFSCA) of 1986 designed to establish drug abuse education and prevention programs. Similar to the JDDP Act, the Drug-Free Schools and Communities Act also included Federal funds for states, schools, and communities to develop or expand drug prevention programs under Subtitle B of Title IV (Ringwalt, et al., 1994). Thus schools became recognized at the Federal level as an important context for prevention efforts, at least with regard to drug abuse.

Between 1974 and 1992 there was an increasing focus on delinquency prevention at the community level and drug abuse prevention at the community level and within the schools. In 1994, with the reauthorization of the DFSCA as the Safe and Drug-Free Schools and Community Act (SDFSCA) under the Elementary and Secondary Education Act (ESEA) of 1994, the connection was made between drug abuse and violence. SDFSCA expanded the scope of the original Act to include authorization of violence prevention activities fueled in large part by growing concerns about the safety of schools and the perceived need to expand the types of intervention and prevention activities engaged in by schools. The intent of the revised Act was to aid and encourage schools to develop integrated programs that addressed a variety of student risk factors contributing to alcohol and drug abuse *as well* as violent behavior (Iowa Department of Education, 2003).

The problems of violence and drug use in the schools were also recognized in a parallel measure, the Goals 2000: Educate America Act signed into law in 1994 establishing a set of eight national education goals. Although primarily focused on academic outcomes goal six stated “By the year 2000 every school in the United States will be free of drugs, violence, and the unauthorized presence of firearms and alcohol and will offer a disciplined environment conducive to learning” (North Central Regional Educational Laboratory, 2003)

Over the past decade the SDFSCA has been reauthorized and amended several times. In the 1998 reauthorization a set of guiding principles known as the Principles of Effectiveness were developed to improve program accountability and out of concerns for program effectiveness and responsibility. The Principles of Effectiveness required schools to conduct needs assessments of drug and violence problems, develop measurable goals, implement research-based programs or promising approaches, and to assess progress toward meeting established goals. Most recently the Safe and Drug Free Schools and Communities Act was reauthorized as Title IV, Part A of the “No Child Left Behind” Act of 2001: 21$^{st}$ Century Schools – Safe and Drug-Free Schools and Communities and increased both state accountability and flexibility in the expenditure of Federal funds (Iowa Department of Education, 2003).

Over the past three decades, beginning in the early 1970’s, there has been a shift in focus from dealing with juvenile delinquency, violence and drug abuse problems through the juvenile justice system in a reactive, punitive manner to addressing these problems proactively using preventive community-initiated and community-based approaches. Most recently these efforts have come to focus on the schools and address the related problems of youth drug abuse, delinquency and violence within a prevention focused risk and protective factors framework. In part these efforts have come to focus on the schools because the problems, in particular school violence, have been perceived as occurring to a significant degree within the schools. Similarly, well-developed research literatures focused on these problems have taught us that such problems have their etiology in childhood, often in the elementary school years, and that the social institutions that are best positioned to prevent the development of drug abuse, delinquency and violent behavior are our schools (Walker, et al., 1996; Walker, Steiber, & Bullis, 1997; Walker & Sprague, 1999).

Recently there has also been a growing understanding and recognition of the roles that schools play in the development of drug abuse, violence and delinquency. Schools are not simply places that these problems happen neither are they inert agents in the development of these problems. Instead, schools often play an active role in the development of such problems moderating or exacerbating the risk factors that students bring with them to school (Walker, et al., 1996).

If schools are to effectively achieve society's goals of preventing youth violence and drug abuse and providing safe and drug free schools they must recognize and actively address the factors within their control that cause, or at least increase school violence and behavior problems. Additionally, recent changes to Federal legislation have increased accountability in the expenditure of Federal funds. Recipients are now required to conduct a needs assessment, to develop needs-based goals, and to monitor progress toward those goals.

Only recently have schools begun to develop and implement comprehensive prevention programs (Corvo, 1997; Larson, 1994; Orphinas et al., 1996). Conspicuously absent from such programs has been a clear articulation, analysis, or explanation of the factors responsible for increasing or decreasing rates of school violence and student problem behaviors (Shoffner & Vacc, 1999). In consideration of the recent increase in accountability attached to Federal funds, and the consequent need to develop and implement effective intervention programs, schools face a critical challenge. Schools must begin to accurately assess and target for change the specific factors that need to be addressed in prevention programs designed to achieve safer schools. However, before such efforts can be successful we must develop reliable, valid, and efficient systems of school-based measurement designed to provide a needs assessment and an indication of the specific factors that should be targeted for change.

The purpose of the present study was to contribute to the development of such a system of school-based measurement. The study examined the psychometric properties of three currently used measures of school function and risk and their validity in relation to school safety. These measures included the School Wide Evaluation Tool (SET), the Effective Behavior Support Survey 1.5 (EBS 1.5), and the Oregon School Safety Survey. An attempt was made, based on the results of the initial analyses, to determine the optimal combination of measures and variables for assessing school safety and the factors that account for it. The specific research questions addressed in this study included:

1. What are the factor structures and internal consistency coefficients of the SET, EBS 1.5, OSSS, and OSCSS?
2. What are the zero-order correlation coefficients between the SET, EBS 1.5, OSSS, and OSCSS at the full-scale and subscale levels?

3. What are the zero-order correlation coefficients between the community and school archival data, SET, EBS 1.5, OSSS, and OSCSS full-scale and subscale scores and the school safety measure?
4. What combination of community and school archival variables, measures, and subscales best predicts the construct of school safety/disorder?

# CHAPTER 2
## Literature Review

Having briefly examined the history of our social and political responses to juvenile delinquency, drug abuse and school violence, at least two questions remain: How effective have we been in addressing the problems of our youth as noted above? Where are we now in dealing with these problems? A more specific and appropriate question might be to ask how far we have come in achieving the sixth goal of the Goals 2000: Educate America Act. We will now examine a variety of statistics and social indicators, both past and present, in an attempt to answer these questions. Trends, or changes over time, will also be explored in an effort to better understand what lies ahead. The literature will also be reviewed focusing on individual student, community and school factors associated with school violence and disorder. Finally, a number of limitations in the current measurement of school violence will be reviewed.

### *Prevalence and trends in juvenile delinquency and school violence*

The public's perception of juvenile crime and delinquency, particularly in recent years, is driven largely by media accounts of rare and tragic events, which although significant and disturbing may not represent actual trends and problems faced by the majority of communities. A better representation of the trends and problems posed by juvenile crime and delinquency, both in the society at large and in the schools, is provided by an examination of the abundant government and private statistics gathered each year.

Due largely to the media tendency to focus on extreme forms of violence, murder has become synonymous with interpersonal violence. Murder rate statistics have come to quantify the problem of interpersonal violence in America. Although murder is a significant and troubling form of violence its occurrence is relatively infrequent compared to other less extreme but more common forms of violence. It is important to consider the problem of interpersonal violence as more than murder rate statistics.

Murder among youth, both as perpetrators and victims, is a significant problem. In the United States homicide is one of the leading causes of juvenile deaths. In 1997, the National Center for Health Statistics listed homicide as the fourth leading cause of death for children between the ages of 1 and 4, the third for youth between the ages of 5 and 14, and the second leading cause of death for those between the ages of 15 and 24 years. Compared to other industrialized nations, the United State's 1993 child homicide rate per 100,000 children under the age of 15 was five times the number of other industrialized nations' rates combined (2.57 vs. 0.51). The rate of child homicides involving a firearm was 16 times greater in the U.S. than the other countries combined (0.94 vs. 0.06) (Snyder & Sickmund, 1999).

Despite these dire statistics, examination of homicide trends in the United States reveals that the homicide rate has declined in recent years. In 1993 the number of juveniles murdered peaked at 2,900, or about 4 murders for every 100,000 youth under the age of 18. By 1997 this number had dropped to 2,100, or about 3 murders per 100,000 juveniles. However, in contrast to the pattern of decline in all murders in the United States the number of 1997 juvenile homicides remained significantly above the mid-1980's levels when about 1,600 juveniles were murdered annually (Snyder & Sickmund, 1999).

Although juvenile homicide is a major concern to the nation at large, such events are relatively rare in public schools. It has been estimated that youth face a less than one in a million chance of suffering a school-associated violent death. During the first half of the 1997-1998 school year

(July 1, 1997 – December 31, 1997) less than one percent of the more than 2,500 nationwide child homicides were committed at school (on school property, at a school sponsored event, or on the way to or from school or a school sponsored event) (U.S. Department of Education & U.S. Department of Justice, 1999).

Nonfatal violent crimes are much more common among youth. In 1999 students between the ages of 12 and 18 were 2.6 times more likely to be victims of nonfatal serious violent crimes (e.g. rape, sexual assault, robbery, and aggravated assault) away from school (476,000) as they were at school (186,000) (Kaufman, et al., 2001). Examination of serious violent crime trends against youth between the ages of 12 and 18, both at and away from school, indicates that students have consistently suffered more serious violent crimes away from school between 1992 and 1999. However, when serious violent crime and simple assault are considered together the difference between victimization rates at and away from school decreases substantially (Snyder & Sickmund, 1999).

Although serious violent crime and murder rates have been decreasing in recent years, the prevalence rates of some types of crimes at school have not. Between 1993 and 1999 the percentage of $9^{th}$ through $12^{th}$ grade students who were injured or threatened with a weapon on school property within the past 12 months remained relatively constant between 7 and 8 percent. However, the percentage of $9^{th}$ through $12^{th}$ grade students who reported carrying a weapon on school property within the past 30 days fell from 12 percent in 1993 to 7 percent in 1999 (Kaufman, et al., 2001). Teachers are also victimized at school. On average, each year from 1993 to 1997 teachers reported 131,400 violent crimes and 222,800 thefts against teachers at school, which equates to 31 violent crimes and 53 thefts for every 1,000 teachers (U.S. Department of Education and U.S. Department of Justice, 1999).

More common forms of victimization experienced by students at school include bullying and harassment, verbal abuse, and theft. A recent survey found that almost 30 percent of youth surveyed indicated that that they had been involved in bullying as either bully or victim. More than 16 percent of students reported being bullied at least occasionally during the school term, while 8 percent reported bullying or being bullied at least weekly (Nansel, et al., 2001). Another survey of teens found that 42 percent witnessed bullying or taunting among other students at least once per day, and an additional 26 percent report seeing bullying incidents at least once per week (National Crime Prevention Council, 2001). More than two-thirds of 14- to 17-year-old students surveyed indicated that there is a group of students at their school that sometimes or frequently intimidate others (cited in NSSC Review of School Safety Research, 2001). Students also frequently experience verbal harassment. A 1999 survey of 12- to 18-year-old students found that 13 percent reported someone at school had used hate-related words against them in the past six months (Kaufman, et al., 2001).

Theft is by far the most common school crime. Between 1992 and 1996 12- to 18-year-old students were the victims of substantially more theft crimes at school than away from school with an estimated 2.1 million in-school thefts reported in 1996. Reported theft both at school and away from school has been in decline since about 1993 dropping 21 and 16 percent between 1992 and 1996 at-school and away from school, respectively. The discrepancy between theft at- and away from school has also lessened although substantially more thefts continue to be reported at school than away from school (Snyder & Sickmund, 1999).

As may be noted in the preceding statistics, students suffer a variety of forms of victimization both at- and away from school ranging from the very serious (e.g. murder, rape, assault) to the less serious (e.g. theft, bullying, verbal abuse). It appears that the more serious crimes are substantially less frequent in general and tend to occur at higher rates out of school than in school. In fact the most serious, murder, is exceedingly rare at school and occurs at a much higher rate out of school than in school. In contrast, the less serious forms of victimization appear to be much more prevalent and may occur more frequently at- than away from school.

The range of youth violence evidences different trends. It appears that the majority of serious crimes against youth, both at and away from school, have substantially declined over the past decade while many of the more prevalent, less serious crimes have remained stable or declined to a lesser extent. When the total number of crimes against students at and away from school are considered, students suffer slightly more victimizations at school than away from school. Between 1992 and 1996 the total number of crimes against students was slightly higher at- than away from school ranging between just over 150 and 100 crimes per 1,000 students, respectively. Between 1996 and 1997 the number of crimes against students at school dropped to just below the number away from school with both hovering just above 100 total crimes per 1,000 students (U.S. Department of Education & U.S. Department of Justice, 1999).

If we define school safety, as many have done, as a lack of serious violent crime then, based on the preceding statistics, schools are safe places to be. However, if we define school safety more broadly, including the range of violent events that occur in schools the question of how safe our schools are is less clear. *As noted above, students have tended to suffer slightly more total victimizations at school than away from school*. Commenting on crimes committed against youth at school Snyder and Sickmund (1995) noted, “There is no comparable place where crimes against adults were so concentrated” (p.16). Clearly there are differences of opinion on how safe our schools are depending in large part on how school violence and safety are defined. We have chosen a broader definition of school safety that will be elaborated on in a later section and thus would argue that although the majority of schools are not particularly dangerous places to be they are also not particularly safe.

### ***Public Perceptions***

Although statistics provide a more objective assessment of the crime, and violence problems faced by schools and communities also important is the subjective impression of individuals. Perception and fear are important to the extent that students behave in a manner consistent with their perception (Pearson and Toby, 1991). As fear of crime and violence increase in schools and communities, confidence in the ability of adults and social institutions to provide a safe environment diminishes (Welsh, Greene, and Jenkins, 1999). Potential student responses may include carrying weapons to school, fighting or putting on a tough “front” to manage others’ impressions and thwart attacks, and retaliation against perceived transgressors (Lockwood, 1997).

A number of recent polls have surveyed the public on their perceptions of safety and the problems faced by the public schools. One annual poll asked respondents to identify the most pressing educational problems faced by the public schools. Lack of financial support was ranked first followed by lack of discipline in the schools. Overcrowding, fighting/violence/gangs, and drugs were ranked in the top five (Rose & Gallup, 2000). A second survey found that despite the decline in violence in recent years, 71% of respondents stated they believed that a school shooting was likely in their community. Similarly, despite a 40% decrease in school associated violent deaths between the 1997 - 98 and 1998 - 99 school years, there was a 49% increase between 1998 and 1999 in the number of respondents fearing a school shooting in their community (Ross, 1999).

Recently the Gallup Poll (Polling the Nations database, 2003) assessed perceptions of safety in the public schools. The number of respondents reporting that they feared for their oldest child's physical safety in 2001 ranged from 32% to 45%. Similarly, 31% of respondents reported frequently or occasionally fearing that their school-aged child might be physically harmed while attending school. Only 81% of respondents in 2001 reported feeling that the schools in their communities were "very safe" or "somewhat safe". Middle- and high school students have also been asked to report on their feelings of safety at school. One survey found that more than one third of students reported not feeling safe at school (Josephson Institute on ethics, 2001) while in a second national sample between 2% and 16% of high school students surveyed reported feeling too unsafe to go to school with the range due to variation among the states (YRBS, 1999).

Despite the recent decline in number of violent crimes against youth, both in- and away from schools, substantial numbers of students and parents continue to perceive schools as dangerous and unsafe places. Furthermore, in light of the statistics cited above, the fear of harm is greater than and thus disproportionate to the actual risk. The fear of harm appears to be a phenomenon at least partially distinct from actual rates of victimization and violence. However, the lack of correspondence between public perception of risk (fear) and actual risk does not negate the importance of perception. Instead it suggests that perception is an important psychological component of safety, and that both the perception of risk and actual risk need to be included in any serious discussion of school safety or school violence. As previously noted, low perception of safety may foster behaviors that increase actual risk.

## Multilevel School Violence Risk and Protective Factors

Why do students engage in acts of violence and other deviant behaviors both in and out of school? Research indicates that such behaviors are typically not random. Researchers in the fields of psychology, sociology, criminology, and public health have developed a large literature delineating specific antecedents and correlates of youth antisocial and problem behavior (Derzon, 2001; Farrington, 1995; Loeber & Dishon, 1983). Furthermore, these antecedents and correlates apply to a wide range of problem behaviors. Dryfoos (1990) summarized the correlates across a range of youth problem behaviors, and found more similarities than dissimilarities in antecedents. Thus there appear to be an important set of antecedent variables associated with a wide range of youth problem behaviors. Knowledge of such factors does not allow for perfect prediction of the occurrence of individual acts however, such knowledge does improve our ability to predict the likelihood of acts of violence and deviant behavior at the group level, and thus intervene to prevent and reduce school violence.

We will examine this literature focusing on individual student-, community-, and school-level factors associated with school violence. The risk for school violence is determined at multiple levels including the individual characteristics of students, the characteristics of the community in which the students reside and the school is located, and the characteristics of the school itself. It is important to note that although three distinct sets of factors related to school violence have been specified, they are not independent of each other and are often highly interrelated. Individual student characteristics often are heavily influenced by the communities in which they reside. Similarly, schools are not independent of their communities but are largely a product of the communities in which they are embedded. The characteristics and needs of students who attend a school often influence the way a school is structured and operates. It is important to keep in mind the reciprocal nature of the relationships between individual student-, community-, and school-level factors as they are discussed.

The review of school-violence associated variables will be conceptualized using a risk and protective factors model. The risk and protective factors approach to prevention emerged during the 1990's based on the understanding that a variety of problem behaviors are highly related and share common antecedents. The proponents of such an approach argue that the common antecedents are risk and protective factors identified in research as correlates of problem behavior (for a review, see Hawkins, Catalano, & Miller, 1992). Based on this conceptualization it is reasoned that prevention efforts directly targeting these risk and protective factors will reduce problem behavior (Gottfredson, 2001).

Risk factors have been defined as individual or environmental characteristics that, when present, indicate an increased likelihood that individuals will develop a disorder or disease (Garmezy, 1993). Risk factors may act both directly and indirectly through a mediational chain of events and factors to increase the probability of disorder within an individual. The presence of one or more risk factors does not indicate that disorder is inevitable but only that there is an increase likelihood of disorder. Similarly, a greater number of risk factors represent a relatively greater risk of disorder. Protective factors are individual or environmental characteristics that, when present, indicate a reduced likelihood of disorder. Protective factors are thought to reduce the risk of disorder by mediating or mitigating the negative effects of risk factors (Gottfredson, 2001). Protective factors may function by directly reducing dysfunction, by reducing the effect of a risk factor, through a disruption of the mediational chain thorough which the risk factor produces its negative effects, or by preventing the initial occurrence of the risk factor (Coie et al., 1993).

## Individual Student Characteristics

Individuals commit acts of violence and problem behavior. Consequently, the safety of schools and risk for violence is largely determined by the behavior of students. In fact, some of the strongest correlates of school disorder are the characteristics of the school's population (Gottfredson, 2001). A large body of research supports a connection between certain individual characteristics and a variety of problem behaviors.

Early aggression and conduct problems have been shown to strongly predict later aggression and conduct problems. White, Moffitt, Earls, Robins, and Silva (1990) classified children in the Dunedin Multidisciplinary Health and Development Study based on the presence of diagnosed antisocial and nonantisocial disorder. They then sought to retrospectively predict membership in each group based on early childhood traits. As early as age three, it was possible to discriminate antisocial disordered from nondisordered and other disordered youth with an 81% correct classification accuracy, based on five preschool variables. These included, in order of importance: parent-reported behavior problems at age 5, being difficult to manage at age three, demonstrated externalizing behavior at age 3, and poor performance on two tests of motor skills, including the Draw-A-Man test and the McCarthy Motor Scales. Some of the most powerful early risk factors for later violence between the ages of 15 to 18 are involvement in general offenses and substance use before age 12. General offenses include serious but not necessarily violent acts such as grand theft, burglary, extortion, and felony conviction (U.S. Department of Health and Human Services, 2001).

Hyperactivity and attention problems also appear to be implicated in the etiology of antisocial behavior. Elementary school children between the ages of 8 and 10 years, who were rated by their teachers as hyperactive, impulsive, restless, and as having poor concentration skills, have been shown to be at increased risk of later conviction for delinquency. Interestingly, hyperactivity in 8 to 10 year-olds predicted later juvenile convictions independently of conduct problems during the same time-period. Being rated as daring by both parents and peers between the ages of 8 and 10 also predicted future juvenile convictions (Farrington, 1995). Based on a community study in New Zealand, Moffitt and Silva (1988) found that almost 60% of those with a childhood diagnosis of ADD had become delinquent by age 13. Further analysis revealed that, consistent with previously cited research, ADD delinquents relative to non-ADD delinquents scored significantly lower on measures of verbal and visual integration. In summary, children who display a hyperactive, impulsive, and inattentive constellation, as well as comorbid conduct problems early in life have shown a significantly increased risk of later delinquency (Loeber, 1990).

Individual student attitudes and beliefs have also been demonstrated to increase the risk of antisocial behavior. Lockwood (1997) interviewed middle school students enrolled in an alternative school located in an area of high violence. The interviews focused on incidents of violence the students had been involved in. He discovered a well-developed and coherent belief system favoring the use of violence that students used to justify their acts of violence. Students also evidenced a preference for violent retaliation over other forms of redress, a strong belief in punishment, and sensitivity to perceived injustice and mistreatment.

A number of studies have investigated the link between delinquent or antisocial beliefs and later behavior. Thornberry, Lizotte, Krohn, Farnworth, and Jang (1991) found a reciprocal relationship between delinquent beliefs and delinquent behaviors: delinquent beliefs were found to later predict both delinquent behavior and association with antisocial peers. Similarly, early delinquent behavior predicted later expression of delinquent beliefs. Studying a large sample of adolescents using a seventeen-item self-report measure of beliefs in rules Gottfredson and Gottfredson (1992) found a negative association between belief in the validity of rules and three separate measures of substance use. Recently a series of studies (Welsh, Greene, & Jenkins, 1999; Welsh, 2000; Welsh, 2001) were conducted to examine the relationships between a number of individual student characteristics and what the authors have termed “school disorder”, a construct based in sociological theory. School disorder was defined in the studies using a number of related constructs including: student offending, misconduct, victimization, and avoidance and self-reported teacher victimization. Consistent across studies was the finding that individual student belief in the validity of the school rules was significantly negatively associated with multiple measures of school disorder. As student self-reported belief in the school rules increased various measures of school disorder decreased.

One of the strongest risk factors for engaging in antisocial behavior is association with antisocial peers. A large body of research demonstrates a strong and consistent relationship between peer association and problem behavior, both in childhood and adolescence. Peer rejection in elementary school predicts both later association with antisocial peers and an antisocial behavior pattern during adolescence (Dodge et al. 2003; Laird, Jordan, Dodge, Pettit & Bates, 2001). Dishion, Patterson, Stoolmiller, and Skinner (1991) found that popular and unpopular elementary school children associate in separate groups. When rejected children were placed together, problem behaviors increased. In contrast, when rejected children were placed with non-rejected children, problem behaviors were within normal limits. Association with antisocial peers during adolescence appears to play a deviancy-training role. Dishion, Spracklen, Anderson, and Patterson (1996) videotaped and then analyzed the conversations of adolescent male dyads. Dyads were mixed and consisted of either two delinquent subjects or one delinquent and the other nondelinquent subjects. The conversations were coded on the dimensions of topic (normative vs. rule-breaking talk) and the reactions of the listener (laugh vs pause). Results indicated a significant relationship between rule-breaking talk and a response of laughter in the delinquent dyads whereas in the mixed dyads there emerged a significant relationship between normative talk and laughter. A longitudinal analysis was conducted over the two years following the initial assessment to determine the relationship between dyadic conversation and behavior. Youth who had engaged in rule-breaking talk were significantly more likely to evidence an increase in self-reported delinquency, even after controlling for their prior level of delinquency. The authors interpreted this finding as an indication of a deviancy training effect among delinquent peers. Consistent with these findings Gottfredson (2001) notes that the largest correlate of adolescent problem behavior is exposure to, or association with delinquent peers. Studies have also investigated risk factors related to drug use. Sullivan and Farrell (1999) sought to discover specific risk and protective factors for drug use in a sample of African-American 8th graders enrolled in an urban public school. Peer influences including peer pressure and peer models for drug use emerged as some of the strongest risk factors for later drug use.

A number of school-related attitudes and behaviors have been associated with a variety of student behavioral outcomes including delinquency, violence, and drug use. Specific school-related attitudes and behaviors associated with such outcomes include attachment and commitment to school and academic performance. Attachment and commitment are key components of Hirschi's (1969) social control theory of delinquency. Attachment refers to the emotional bond between an individual and other individuals or social institutions such as schools. Commitment refers to the psychological investment in the pursuit of specific goals such as educational or occupational goals. According to social control theory, individuals with strong attachment and commitment to conventional institutions, such as schools, and goals are less likely to engage in delinquent activities than those with weak bonds. A number of studies have investigated the relationships between attachment and commitment to school and a variety of student behavioral outcomes.

Najaka, Gottfredson, and Wilson (2001) recently used the technique of meta-analysis to investigate the relationship between selected risk factors and problem behavior. The authors analyzed the results of a total of 87 studies that examined school-based interventions, programs, or procedures intended to reduce problem behavior or to affect the presumed causal factors of problem behavior. Distributions of effect sizes were derived for each risk factor and problem behavior. Bivariate regression equations based on the full distribution of effect sizes were used to indicate the relationships between specific risk factors and problem behavior constructs. Results indicated a strong positive relationship ($r = .86$) between improvements in school bonding, defined by measures of attachment and commitment to school, and improvements in problem behavior. School bonding has also been associated with drug use.

Sullivan and Farrell (1999), studying the effects of selected risk and protective factors on subsequent drug use in African-American adolescents, found that commitment to school served as a protective factor by moderating the effects of risk factors on a measure of composite drug use. Similarly, school attendance was found to be a protective factor in relation to some but not all of the drugs in the composite drug use factor. Thus school attendance appears to be less robust as a protective factor than does commitment to school. Similar results were found in a study of 7 suburban middle schools that examined the relationship between student self-report measures of school bonding, commitment to school, perceived school climate, and student problem behavior (Simons-Morten, Crump, Hayne & Saylor, 1999). School bonding, commitment, and perceived school climate were all found to be positively related to each other and inversely related to student self-reports of problem behavior. Kandel, Simcha-Fagan and Davies (1986) conducted a longitudinal study examining the interrelationships and predictors of involvement in delinquent activities and illicit drug use over a nine-year interval. Risk factors were measured during adolescence (age 15 - 16) and related to outcomes in young adulthood (age 24-25). Commitment to school was negatively related to later self-reported theft and fighting as well as Marijuana use. A related variable, high school drop out, was also found to be positively associated with self-reported fighting.

Studies have also linked academic achievement with student problem behaviors. Reviewing the literature, Gottfredson (2001) noted a moderate link between school grades and delinquency, having a police record, having a court record, and engaging in serious juvenile delinquency and drug use. A number of the studies cited above also found links between academic achievement and student outcomes. Najaka, Gottfredson, and Wilson (2001) found a significant but moderate relationship ($r = .33$) between improvements in academic performance and improvements in problem behavior. These results were consistent across grade level and risk status and thus although moderate were robust. Related to student academic achievement Sullivan and Farrell (1999) found that parent's positive expectations for academic achievement served as a protective factor for later drug use.

In summary a substantial body of research has demonstrated significant links between a number of individual student characteristics and a variety of student problem behaviors. Early aggression and conduct problems, as early as age five, strongly predict later violent and antisocial behavior. Hyperactivity and attention problems are also associated with later juvenile convictions and antisocial behavior. This finding may be due in part to the high comorbidity of conduct problems with hyperactivity. However, hyperactivity predicts later juvenile arrest independently of comorbid conduct problems and thus appears to contribute risk independent of conduct problems. Individual student beliefs have also been linked to later antisocial and violent behavior and to substance use. Specific individual beliefs that serve as risk factors for later antisocial behaviors include a belief system favoring the use of violence, and a lack of belief in the validity of the school rules. One of the strongest risk factors for engaging in antisocial behavior is association with antisocial peers. It has been hypothesized that associating with antisocial peers results in deviancy-training the consequence of which is that youth learn to be deviant and are reinforced for engaging in deviant acts. A number of school-related attitudes and behaviors have also been identified as risk factors including a lack of attachment and commitment to school, and poor academic performance.

## Effects of Community Characteristics

Because schools exist within the larger context of communities, their composition and functions are largely interdependent with and determined by the communities that surround them. Similarly students live in the communities served by schools and student characteristics, as has been shown in the previous section, both strongly influence schools and are themselves heavily influenced by their communities. Thus community characteristics directly influence schools through such avenues as funding levels and resources and indirectly through the characteristics of the students they send to the schools. Researchers have long been interested in the relationship between community characteristics, crime and violence thus a large literature has been developed in this area. More recently, attention has been focused on the relationships between community characteristics and school functioning.

Plybon and Kliewer (2002) employed data from a cross-sectional investigation of the impact of violence exposure in predominantly African-American urban youth to study the relationship between a number of community and family risk factors and child outcomes. Census data were used to determine neighborhood risk factors including: poverty, family structure, crowding, and neighborhood stability. Police records were used to assess personal victimization rates. Mothers were also asked to report on their perception of neighborhood safety and on their children's exposure to violence in the community. A number of family environment characteristics were also assessed.

In an effort to derive neighborhood types, census and police data were cluster analyzed resulting in the specification of three neighborhood types based on two dimensions: low poverty-low crime (LPLC), high poverty-moderate crime (HPMC), and moderate poverty-high crime (MPHC). Analysis of Variance was then used to examine differences in child externalizing behavior problems across neighborhood types. A number of significant differences between neighborhood types on the outcome variable emerged. Children in the LPLC neighborhoods evidenced the lowest rates and children in the HPHC neighborhoods the highest rates of externalizing behavior problems with the difference between the two significant. Family stress levels also significantly differed across the neighborhood types with families in the HPMC neighborhoods experiencing the highest levels of stress. Additional analyses indicated that family stress moderated the relationship between neighborhood type and child outcomes with higher family stress related to poorer child outcomes. Family cohesion was found to strongly buffer the effects of neighborhood disadvantage on child externalizing difficulties, but only in high poverty neighborhoods. The authors concluded that poverty, in conjunction with other neighborhood risk factors, serves as a risk factor for negative child outcomes. The study clearly indicates neighborhood differences in child externalizing behavior based on the presence and level of specific risk factors.

Similar results were found in a study of the relationship between neighborhood characteristics and rates of violent crime (Smith & Jarjoura, 1988). Based on interviews with a random sample of 11,419 individuals from 57 neighborhoods located in three states researchers examined the relationships between neighborhood crime rates and neighborhood characteristics. Neighborhood crime included violent crime and burglary rates and neighborhood characteristics included the percentage of low-income households, residential mobility, racial heterogeneity, percentage of single parent households, population density, and the percentage of the population between the ages of 12 and 20. Results indicated a positive relationship between residential mobility and violent crime rates, but only in poorer neighborhoods. The interaction between residential mobility and low-income households was robust remaining significant even after controlling for other neighborhood attributes. Similarly, racial heterogeneity was positively associated with increases in violent crime independent of both residential mobility and percentage of low-income households. Neighborhood population density was also found to positively predict violent crime rates, as did the percentages of youth and single parent households. These three characteristics were also associated with higher rates of burglary.

Children's level of exposure to community violence has been associated both with externalizing behavior problems and academic functioning. Schwartz and Gorman (2003) studied the relationship between community violence exposure and children's school-related difficulties in a sample of 237 urban third, fourth, and fifth graders. Measures included student self-report and teacher report on students as well as student academic functioning defined by GPA and SAT-9 mathematics score. A significant positive relationship emerged between community violence exposure and three measures of peer and teacher rated aggression and hyperactivity. A significant negative relationship was found between community violence exposure and academic functioning. Thus children's exposure to community violence appears to put them at risk for increased behavioral difficulties and lower academic performance.

Colder, Mott, Levy, and Flay (2000) recently examined the relationship between perceived neighborhood danger and childhood aggression in a sample of 735 predominantly African-American fifth graders in 12 elementary schools. Students were asked to self-report the lifetime occurrence of their aggressive behavior and were assessed on their positive beliefs about aggression. Parents and students were asked to indicate the level of perceived danger in the neighborhood. Boys were found to be significantly higher in both aggression and positive beliefs about aggression. High levels of perceived neighborhood danger were significantly positively associated with both positive beliefs about aggression and levels of reported aggression. Structural equation modeling assessed both the direct and indirect relationships between perceived neighborhood danger and aggression. A direct relationship between perceived neighborhood danger and aggression was found, as was an indirect relationship through positive beliefs about aggression. Results of the study indicate that student- and parent-perceived neighborhood danger is associated with both positive beliefs about aggression and actual levels of self-reported aggression. Furthermore, the relationship between perceived neighborhood danger and aggression appears to be at least partially mediated by positive beliefs about aggression.

The above cited studies examined the links between specific neighborhood characteristics and violence and other problem behaviors within the neighborhood. Although the studies did not explicitly link community characteristics with school violence and other problems, the link may be assumed. If neighborhoods produce residents who engage in such behaviors in their communities those same individuals are likely to attend or interact with the schools in their communities in a similar manner and thus bring aggression, problem behaviors and crime to their schools. As previously noted, schools reflect the communities that surround them. A number of studies have specifically examined the relationships between community characteristics and specific school characteristics and the impact of communities on school function and problems.

One such study investigated the effects of crime and violence in neighborhoods and schools on in-school student behaviors and performance using a large nationally representative sample of middle- and senior-high school students (Bowen & Bowen, 1999). Constructs assessed included neighborhood danger based on student self-report measures of negative peer culture and student's experiences with crime and violence, and student self-report measures of perceived school danger. Outcome measures included school attendance, a grade index, and a scale the authors labeled trouble avoidance, which assessed three school-related problem behaviors (lower score indicates fewer problem behaviors). A comparison of means indicated that males reported significantly more personal experiences with neighborhood and school danger than females. Similarly, high school students reported significantly higher mean scores on neighborhood peer culture and personal threats and school crime and violence compared with middle school students.

Hierarchical multiple regression analyses were used to determine the relationships between the outcome measures and student demographics, neighborhood danger and school danger. Student demographic variables significantly predicted trouble avoidance accounting for 5.2% of the variance. The addition of neighborhood peer culture and neighborhood personal threats significantly contributed to the prediction of trouble avoidance, explaining an additional 11.9% and resulting in a little more than 17% of the total variance being accounted for. These results indicated that increased neighborhood danger was associated with lower trouble avoidance. Further analyses were used to examine the unique contributions of the neighborhood and school measures to trouble avoidance. The unique contributions of the neighborhood and school measures to the prediction of trouble avoidance were 5.7% and 5.5% of the variance, respectively. These results indicate that the combination of student demographic and neighborhood danger variables account for more of the variance in school problem behavior then school danger measures although these also accounted for a significant portion of the variance. Neighborhood danger and student demographics were also significantly associated with student academic performance indicating that as neighborhood danger increases student grades decrease.

Employing a sample of 58,000 students in 120 Boston public schools Hellman and Beaton (1986) examined the role of community and school characteristics in patterns of violence in urban public middle- and high schools. School violence was defined using student suspension rates. Analyses indicated that more than one third of student suspensions in Boston public schools were for violent acts. Suspension rates also demonstrated a strong positive relationship with community crime and a moderate positive relationship with calls for police service. Consistent with the previously cited research this relationship was interpreted by the authors as an indication of the interconnectedness of school violence and disruption problems and crime and disruption in the community. Significant relationships between school suspension rates and community characteristics were also found. A significant positive relationship emerged between suspension rates and poor or overcrowded housing, population density, and community instability. Suspension rates were negatively associated with the percentage of owner-occupied residences in communities suggesting a stabilizing and thus protective effect. Separate analyses indicated a relationship between community structure measures and community crime rates similar to that found with school suspensions. Taken together the results of this study suggest that community characteristics associated with community crime and violence are also associated with school suspensions, one measure of school violence and disruption. Thus there appears to be a strong community contribution to school crime, violence, and disruption.

In summary, both community and school crime and violence rates are significantly associated with a common set of community characteristics including: community poverty, poor quality and overcrowded housing, percentage of single parent households, neighborhood racial heterogeneity, population density, residential mobility, and percentage minority population. Evidence also suggests that community crime and violence rates are strongly correlated with the student behavior problems, crimes and violence that schools within those communities experience. Thus it appears that community characteristics contribute substantial variance to the prediction of school violence and serve as risk factors for school violence and crime.

**Effects of School Characteristics**

Thus far in this document the literature on risk- and protective factors at the individual student and community levels has been reviewed. It appears that students may present with inherent characteristics that put them at risk for engaging in violent and antisocial behaviors. Similarly, a number of community characteristics place students and schools at risk for acts of violence, aggression and crime. When students enter schools, and community members interact with schools they bring with them accumulated risk factors and behavior patterns. These in turn interact with the characteristics and environment of the school which buffer or exacerbate the risk factors and behavior patterns contributed by the individual and community. In brief, school environments interact with community and individual factors in a manner that either increases risk or serves to protect against risk and thereby at least partially determines the behaviors of individual students that constitute school violence. We will now explore the research that has examined the specific contributions of schools to school violence and antisocial behavior.

Student-school bonding has been associated with a number of adolescent problem behaviors. Social bonding or attachment is a central tenet of social control theory, which suggests that an emotional attachment to conventional social institutions, such as family and school, commitment to conventional pathways to achievement, and belief in the legitimacy of social rules and order form a social bond between individuals and normative social institutions that prevents deviant behavior (Hirschi, 1969).

Simons-Morton, Crump, Hayne, and Saylor (1999) investigated the link between student-school bonding and adolescent problem behavior using a sample of sixth, seventh, and eight graders in seven suburban middle schools. Students were administered self-report measures assessing student problem behaviors, school bonding, perceived school climate, and student perception of their adjustment to school. Results indicated that the measures of school bonding, perceived school climate, and school adjustment were positively correlated with each other but inversely correlated with the measure of problem behavior. A number of sex and grade differences between the three measures and problem behavior emerged. Reports of school bonding, climate, and adjustment were all significantly higher in females than in males, and higher among sixth relative to both seventh and eight graders. Similarly, seventh graders scored significantly higher than eight graders on the three measures. The level of reported problem behaviors was significantly higher among males than females. There were also significant differences in the levels of reported problem behaviors across the grade levels such that eighth graders reported significantly more problem behaviors than seventh graders who reported significantly higher levels of problem behavior relative to sixth graders. These results indicate that student-school bonding, student rated school climate, and student adjustment are all significantly associated with rates of student problem behaviors. Furthermore the level of student problem behavior increases with grade level and is disproportionately accounted for by male students.

Similar results were found in a study of the relationship between school bonding and delinquency focusing on differences between African American and Caucasian students (Cernkovich & Giordano, 1992). The authors administered a multidimensional measure of school bonding to a sample of 942 12- to 19-year old students. The measure included scales that assessed school attachment, attachment to teachers, school commitment, perceived risk of arrest to educational and occupational opportunities, amount of involvement in school activities, parental interest in and support for school-related activities, and perceived opportunity for future success. Students were also asked to self-report on their delinquent behaviors.

Initial correlation analyses indicated that males reported significantly higher delinquency rates than females. This finding was consistent across the two racial groups for both males and females. Similarly, African American and Caucasian students demonstrated comparable levels of school bonding suggesting that African American students are at least as strongly bonded to the educational system as are Caucasian students. Additional analyses examined the relationships between school bonding factors, SES, age, school context and delinquency by regressing delinquency on each of these variables.

A number of significant race and sex differences were found in the prediction of delinquency. Most notably, school commitment was significantly associated with delinquency rates for males from both racial groups indicating that as commitment to school increases delinquency rates decrease. School attachment was also negatively associated with delinquency rates, however only for African American males. Significant negative and positive correlations for perceived risk of arrest and age, respectively, emerged in association with delinquency rates, but only for white females. The regression model accounted for between 16.8% and 11.9% of the variation in delinquency rates depending on the sex and race of the subgroup.

Despite these differences in prediction the authors concluded that there were no significant race or sex subgroup differences in the overall regression model's ability to predict delinquency. Overall, attachment to teachers, perceived opportunity, sex, and commitment to school were found to be the strongest predictors of delinquency. The authors also concluded that the effects of school bonding on delinquency rates are essentially the same across racial groups.

Student perceptions of school safety have also been associated with specific school factors. Employing a sample of tenth graders Lee and Croninger (1996) studied the relationship between student perceptions of school safety and several school characteristics. Students were asked to rate their level of agreement with the statement, "I don't feel safe in this school" which was used as the primary outcome measure. The researchers validated the student school safety perception measure by correlating it with reported victimization rate resulting in a significant positive relationship. School characteristics were also found to significantly predict student safety perceptions. Student demographics exerted a strong effect indicating that students in schools with low average student socioeconomic status and high minority concentrations felt significantly less safe. The effects of student demographics were so strong that when included in the prediction equation with other school structural characteristics, including school size, and urban/suburban location they drowned out the effects of the structural characteristics accounting for 29% of the between school differences in student safety perception. Positive student-teacher relationships were also strongly predictive of student perceived safety such that an increase in positive student-teacher relationships also increased student perceptions of safety.

School characteristics have also demonstrated a protective function against known risk factors. Based on a longitudinal study of nine high schools, Crosnoe, Erickson, and Dornbusch (2002) examined the protective functions of selected school and family factors against the risk of having deviant friends. Students were administered surveys that assessed their frequency of involvement in various delinquent acts and substance use. They were also administered a survey that assessed hypothesized protective factors including: parental monitoring and involvement, household organization, bonding to teachers, academic achievement, and orientation to school, which measured student engagement in and commitment to learning.

Controlling for demographic characteristics the authors conducted separate regression analyses examining the relationships between the deviancy measure and the risk and protective factors. Having delinquent friends was significantly associated with delinquent behavior in both boys and girls. A number of the school factors were found to function as protective factors for delinquency including academic achievement, school orientation, and teacher bonding. These factors were also found to protect against later use of alcohol and other illegal drugs although effects varied by sex. The results of these two studies suggest that school characteristics including student demographics, student-teacher bonding, academic achievement levels, and student engagement in and commitment to learning may act to both increase risk and buffer the effects of other risk factors associated with significant negative student outcomes.

In a series of publications Welsh and colleagues (Welsh, Green & Jenkins, 1999; Welsh, 2000; Welsh, 2001) explored the results of a large-scale study of Philadelphia public schools. The study was designed to explore the relationships between school, community, and individual student characteristics and what the authors called "school disorder" in a sample of eleven middle schools. School disorder was defined using student self-report measures of Misconduct, Offending, Victimization, Perceived Safety, and Avoidance. The authors also used student self-report measures to assess the psychosocial climate of the school and individual school-related student characteristics. A number of community characteristics were also included in the analyses.

Analyses explored the relationships between the psychosocial climate and student characteristic constructs and each of the school disorder scales. A majority of the psychosocial climate scales including: Respect for Students, Fairness of Rules, Clarity of Rules, Planning and Action, and Student Influence were significantly associated with the level of school disorder. Specifically, levels of disorder were lower in schools where students reported higher levels of respect for students, greater fairness and clarity of the school rules, greater levels of student involvement, and more planning and action on the part of the teachers and administration. These scales measure student perceptions of their school climate.

Specific student characteristics were also significantly associated with levels of school disorder including: Belief in Rules, School Effort, and Positive Peer Associations. Levels of disorder were found to be lower in schools where students reported higher levels of belief in the validity of the school rules, greater school effort, and higher levels of association with prosocial peers. Age, sex and race were also significantly associated with levels of disorder suggesting that schools whose populations are characterized by larger numbers of male students, nonwhite students, and older students experience greater levels of school disorder.

In summary, the rates of school violence and other problem behaviors appear to be affected both by the psychosocial climate of a school and the demographic and school-related characteristics of the student population. Schools which demonstrate respect for the students, develop a set of fair and consistent rules that are clearly stated, allow for student influence in the operation of the school, and actively engage in planning and action experience lower levels of disorder. Similarly, fewer problems occur in schools where students believe in the validity of the school rules, are committed to and put forth effort to succeeding socially and academically in school, and are attached to and experience positive relationships with prosocial others including peers and teachers. The demographic composition of schools also affects school outcomes. Schools with larger percentages of minority students, male students, and older students are at greater risk for acts of aggression, violence and crime.

***School Safety Measurement***

Although well-developed measures designed to assess youth violence currently exist, few specifically assess school violence as a distinct phenomenon. The majority of these measures are focused on assessing either the prevalence and frequency of specific antisocial and violent acts within a sample of youth, or on an individual's current antisocial and violent behaviors or potential for such behaviors at a later time. Only a few measures have been developed to assess community and school organizational and cultural features that may serve as risk and protective factors for youth violence (for a review see Minogue, Kingery & Murphy, 1999).

Such an emphasis on assessing the prevalence and frequency of antisocial and violent acts in samples of youth and in individuals is understandable given the long standing clinical and research interest in youth violence, broadly defined. Clinicians have been interested in diagnosing and treating individuals who exhibit patterns of violence and antisocial behavior. Researchers have been interested in studying the prevalence, frequency, and antecedents of such behaviors in groups. However, only recently have we come to view school violence as a problem unique to the school setting rather than as simply violence that happens in schools (Furlong & Morrison, 2000) and therefore requiring study and intervention within the context of the school.

Attempts to measure school safety and violence have only recently begun and have relied heavily on instruments developed for use in non-school settings such as clinical practice and research. Although many of the instruments used demonstrate acceptable reliability and validity, the field of school violence measurement continues to suffer methodological and measurement limitations. Five such limitations will be discussed.

The first and foremost limitation of the current assessment tools is the lack of a clear, operational definition of school violence, which has been variously defined as the perpetration of violence, antisocial behavior, violence victimization, criminal behavior, school discipline/climate, and fear/worry beliefs (Furlong & Morrison, 2000). The lack of a clear, operationally defined, and mutually agreed upon definition of school violence continues to hamper the development of reliable and valid measures.

A second problem in the measurement of school violence is an overemphasis on the assessment of individual student behaviors and characteristics. More specifically, the focus is often on the assessment of rates and types of student problem behaviors and on finding students who pose the greatest risk for violent and antisocial acts. Although it is important to assess individual student behaviors and characteristics, as noted in the previously cited literature, it is also important to assess community and school characteristics and organizational features. Thus one common and important limitation of school violence measurement is an overemphasis on the assessment of individual student characteristics to the exclusion of school and community characteristics and organizational features.

Consistent with the overemphasis on the assessment of individual student characteristics is a third limitation, a paucity of empirically validated and reliable instruments designed to measure school and community characteristics related to school safety and violence. As previously shown, a variety of school and community variables contribute a significant portion of variance to the prediction of school violence and thus are important to include in any comprehensive and valid assessment of school violence and safety. Additional effort should be made to develop reliable and valid instruments that assess school and community factors associated with school violence and safety.

A fourth methodological limitation frequently results from reliance on a single source of information and a single method in the assessment of school safety. Two critical threats to validity in any investigation include a monomethod bias and a monooperation bias (Cook & Campbell, 1979). Thus the validity of an assessment is compromised when information is derived from a single method of measurement (e.g., self-report) and when a construct has been operationally defined in a single way (e.g., school violence defined only by school crime rates). Each assessment method, to varying degrees, yields different results that are not available from other sources. This suggests that various assessment methods have unique strengths and limitations that need to be recognized and taken into account when measuring specific constructs (Campbell & Fiske, 1959). Optimal assessment of a construct results from the combination of data from multiple methods and multiple operational definitions commonly referred to as a multitrait-multimethod assessment (Meyer, et al., 2001). Similarly, measurement should incorporate information from multiple perspectives. For example, students, teachers, and parents might all be queried for the unique information that each can provide (Minogue, Kinger & Murphy, 1999; Stephens, 1998). Frequently the assessment of school violence and safety is at least partially invalidated due to an over reliance on a single method of assessment (e.g., survey), a single source of data (teacher report), and a simple, single operational definition of school violence all of which, as noted above, result in threats to validity.

Finally, a fifth limitation in the measurement of school violence and safety is the lack an agreed upon measurement system that allows for the meaningful and efficient integration of information from multiple methods of assessment and multiple sources of information, as argued for above. Although it is important to measure individual student, school, and community characteristics when assessing school violence and safety using a multitrait-multimethod approach and to incorporate information from multiple sources this is not enough for several reasons.

As noted above, different methods of measurement evidence strengths and weaknesses when used to assess specific constructs. Similarly, different data sources may also differ in the value and validity of evidence provided. Without a system of measurement that recognizes the strengths and limitations of each method and source of data it is very unlikely that these strengths and limitations will be taken into account and weighted appropriately when drawing conclusions. Without a measurement system, attempts to integrate data from multiple methods and sources will also likely be unreliable, as each instance of measurement will employ a different method of integrating the data. Another reason for the importance of a system of measurement is that it allows for the meaningful integration of information and thus provides a comprehensive and integrated understanding of school violence. Without such a system to integrate information one is left with a fragmented and incomplete understanding. A final reason for the importance of a measurement system is that of efficiency. Unless the relationships between multiple measures are understood and accounted for, unnecessary redundancy in measurement may result, which is time consuming and inefficient.

In summary, school violence is a complex, multidimensional construct that results from the interplay of a variety of risk and protective factors. Important contributors include individual student-, community-, and school characteristics and organizational features. Society has become Increasingly aware of and concerned with the problem of school violence, and consequently with improving school safety. Largely through Federal legislation, policies and funding have been put in place to encourage and provide resources for schools to address youth violence in a proactive and preventive manner. However, schools are also expected to demonstrate results with the funding provided by developing measurable goals, implementing research-based programs, and monitoring progress toward meeting establish goals. Thus a necessary component of accountability is a comprehensive and valid system of measurement.

Although efforts are being made to measure school violence and safety, a number of limitations continue to plague these efforts. More specifically, no clear, operational definition has been agreed upon; there is an overemphasis on the assessment of student characteristics to the exclusion of community and school characteristics; few validated and reliable measures of school and community characteristics exist; validity of measurement results is often threatened by the over reliance on a single source and method resulting in a monomethod bias and monooperation bias; and finally, there is no measurement system that allows for the integration of information from an assessment of school violence.

This study was designed to advance the field of school violence measurement by addressing the previously noted limitations. The psychometric properties of three instruments designed to measure school and community characteristics, risk and protective factors, and organizational features were examined. The validity of these measures were evaluated in relation to school violence using a multitrait-mulitmethod design that includes multiple sources of information, multiple measurement methods, and multiple operational definitions of school violence. Finally, the interrelationships between each of the measures and the construct of school violence were assessed to determine the optimal combination of measures.

# CHAPTER 3

## Methods

### *Participants*

The study included four data samples specific to the four measures used. Because archival data collected in previous studies and data collected in the present study were combined, data samples overlap to varying degrees. I will briefly describe sample characteristics for each measure.

The Oregon School Climate and Safety Survey (OSCSS) was administered to students in two middle and two high schools located in Oregon. A total of 665 student surveys were collected. The sample was 48.3% male and 50.2% female with 1.5% unknown. Students ranged in age from 10 to 18 years with a mean of 15.51 years. The sample racial composition included 4.4% American Indian/Alaskan Native, 1.5% Asian-American or Pacifica Islander, 15.5% Spanish-American, 2% African American, 67.5% Caucasian, 7.5% Other with 1.7% unknown.

The School Wide Evaluation Tool (SET) was completed by an outside observer. A total of 246 surveys were collected from 156 schools located in Oregon (80.1%), California (19.5%), and Arizona (.4%). The schools included in the sample were located in rural, urban and metropolitan communities.

The Oregon School Safety Survey (OSSS) was completed by teachers and administrators in each school. A total of 2,668 surveys were completed in 104 schools located in Oregon, California, Minnesota, and Arizona. The schools included in the sample were located in rural, urban and metropolitan communities. Information on the total number of schools included per state was not available.

The Effective Behavior Support Survey 1.5 (EBS 1.5) was completed by teachers and administrators in each school. A total of 1,337 surveys were completed in 59 schools located in California (16%), Minnesota (.1%), New Mexico (9.6%), Oregon (67.3%), and South Dakota (7%). The schools included in the sample were located in rural, urban and metropolitan communities. Table 1 below details the number and type of schools that participated and the number of surveys completed at each level.

**Table 1 Participants and Characteristics of Data Sets**

| Instrument | Elementary Schools (Surveys) | Middle Schools (Surveys) | High Schools (Surveys) | Other Schools (Surveys) | Total Surveys |
|---|---|---|---|---|---|
| Oregon School Climate & Safety Survey (OSCSS) Form A | 0 | 2 (171) | 2 (163) | 0 | 334 |
| Oregon School Climate & Safety Survey (OSCSS) Form B | 0 | 2 (172) | 2 (159) | 0 | 331 |
| School Wide Evaluation Tool (SET) | 104 (172) | 33 (49) | 14 (17) | 5 (8) | 246 |
| Oregon School Safety Survey (OSSS) | 68 (1488) | 20 (561) | 11 (524) | 5 (95) | 2668 |
| Effective Schools Battery 1.5 (EBS 1.5) | 34 (564) | 14 (283) | 8 (376) | 3 (114) | 1337 |

Due to differences in overlap between the data sets the size of the data set used in the various analyses varied. Table 2 below indicates the number of schools that overlapped other schools on the specified surveys. This indicates the number of schools included in each analysis using the set of surveys specified. It is important to note that table 2 indicates the number of schools included in various analyses and not the total number of surveys included in each analysis.

**Table 2: School Overlap by Surveys**

| | OSSS | SET | EBS | OSCSS |
|---|---|---|---|---|
| OSSS | 104 | | | |
| SET | 48 | 155 | | |
| EBS | 33 | 38 | 58 | |
| OSCSS | 2 | 2 | 0 | 4 |

***Measures***

*School-Wide Evaluation Tool (SET)*

The SET (Horner, et al, 2002) is used to assess implementation of school-wide behavior support within a school setting. It is administered by a trained observer/data collector who interviews administration, staff, and students, directly observes multiple school settings, and examines a school's behavior support documentation. The SET has 28 items organized into seven subscales that measure the seven key features of school-wide positive behavior support. The key features include: (a) school-wide behavioral expectations are defined, (b) school-wide behavioral expectations are taught to all children in the school, (c) rewards are provided for following school-wide behavioral expectations, (d) a consistently implemented continuum of consequences for problem behavior is in place, (e) problem behavior patterns are monitored and the information is used for on-going decision-making, (f) an administrator actively supports and is involved in the school-wide behavior support effort, and (g) the school district provides support to the school in the form of functional policies, staff training opportunities, and data collection options.

The SET is scored by assigning a value of 0, 1, or 2 indicating the level of implementation for each of the 28 items (0 = not implemented, 1 = partially implemented, 2 = fully implemented). Subscale summary scores are computed as a percentage of total possible points based on the points received. A total summary score is also computed based on the mean of the seven subscales.

Only one published study (Horner, et al, 2002) has reported on the psychometric properties of the SET. The authors reported subscale Alphas ranging from .63 to .92 and a full scale Alpha of .96.

*Effective Behavior Support Survey 1.5 (EBS 1.5)*

The EBS survey is used to assess the level of implementation of behavior support systems in schools. The survey examines both the current status and need for improvement of four behavior support systems including: (a) school-wide discipline system, (b) non-classroom management systems (e.g., cafeteria, hallway, playground), (c) classroom management systems, and (d) systems for individual students engaging in chronic problem behaviors. The EBS 1.5 contains 43 questions assessing each of the four systems. Respondents are asked to rate the Current Status for each feature as "In Place", "Partial in Place", or "Not in Place". For each feature respondents are also asked to indicate the Priority for Improvement as high, medium or low. Multiple school personnel complete the survey based on their experience and perceptions. The survey is scored by aggregating the responses of all respondents. Percentage of total scores is derived for each of the three response choices on both the Current Status and Priority for Improvement questions. For example, under Current Status, the total number of responses for each response option (e.g. In Place, Partial in Place, Not in Place) are divided by the total number of responses in Current Status. The same procedure is followed to score the Priority for Improvement questions. Little is known about the psychometric properties of the EBS 1.5 and no published studies have examined its psychometric properties (G. Sugai, personal communication, May 22, 2003).

*Oregon School Safety Survey (OSSS)*

The OSSS (Sprague, Colvin, & Irvin, 1995) is used to assess risk and protective factors associated with school violence and discipline problems. Risk factors include poverty, child abuse, graffiti, bullying, and deteriorating physical facilities. Protective factors include positive teacher-student relationships, parent involvement, student supervision, and high academic expectations. The survey asks respondents to rate the extent to which 17 risk and 16 protective factors exist in their school or neighborhood using a rating scale of 1 (not at all) to 4 (extensive). A range of individuals may be asked to complete the scale including administrators, teachers, parents, students and community members. The OSSS produces summary scores for the risk and protective factors. Scoring for each factor is accomplished by summing all individual questions within each factor and dividing by the total number of questions in the factor, which yields an average score for each factor. Only one published study (Sprague, Colvin, Irvin, & Stieber, 1998) has examined the psychometric properties of the OSSS. Reported Alpha Reliability coefficients for the Risk/Protect subscales are .87 and .88, respectively. A copy of the instrument is included in appendix C.

*Oregon School Climate/Safety Survey (OSCSS)*

The OSCSS measures school psychosocial climate and student characteristics associated with school violence and discipline problems. The survey was developed from a combination of scales taken from the 118-item Effective Schools Battery (Gottfredson, 1984) and the 185-item What About You (Gottfredson & Gottfredson, 1989) surveys. The survey is a self-report instrument administered to individual students and divided into two parts (A & B) to reduce the amount of time required for individuals to complete the survey. Form A consists of 50-items in eight scales assessing: Rebellious Behavior, Respect for Students, Planning and Action, Fairness of Rules, Student Influence, Belief in Rules, School Rewards, and Attachment to School. Form B consists of 50-items in seven scales assessing: Student Victimization, Student Safety, Clarity of Rules, Positive Peer Associations, Social Integration, School Effort, and Drug Availability. A random cross section of each participating school's student population will be administered the measures with half of the students completing form A and the other half form B. Scores will be aggregated across students on each of the scales to provide measures of school-level characteristics. The OSCSS was developed for the purposes of this study consequently no data on the psychometric properties of the measure are available.

## School Archival Variables

Archival data were intended to be used to assess additional school risk and protective factors and characteristics. However, due to problems with sample sizes noted below no school archival variables were include in the analyses. The school characteristics assessed using archival data originally included: Total school enrollment, average student/teacher ratio, the percentage of students receiving free and reduced lunch (poverty), average daily student attendance, average student academic performance based on standardized test performance, race, age and sex distributions of the student populations, percentage of at-risk students enrolled, percentage of student turn over, and total number of student suspensions and expulsions.

### *Research Questions*

There were two primary purposes of the study. The first purpose was to explore the psychometric properties of the SET, EBS 1.5, OSSS, and OSCSS. The second purpose was to examine the interrelationships between each of the above measures and with the outcome measure of school safety, both individually and combined. The specific research questions were as follows:

1. What are the factor structures and internal consistency coefficients of the SET, EBS 1.5, OSSS, and OSCSS?
2. What are the zero-order correlation coefficients between the SET, EBS 1.5, OSSS, and OSCSS at the full-scale and subscale levels?
3. What are the zero-order correlation coefficients between the community and school archival data, SET, EBS 1.5, OSSS, and OSCSS full-scale and subscale scores and the school safety measure?
4. What combination of community and school archival variables, measures, and subscales best predicts the construct of school safety/disorder?

Upon beginning the data analysis designed to answer questions three and four it was discovered that it was not possible to answer these questions as originally stated. This was due to the fact that the research questions required that the school serve as the unit of analysis. OSCSS survey data was available for only four schools not all of which included SET, EBS 1.5, and OSSS data. Consequently at most four schools were available for analysis and in several cases less than four schools (see table 2 above). Based on a sample of only four units of analysis it was not possible to conduct the specified correlations or multiple regression analyses specified originally. Based on discussion with my advisors it was agreed to omit question three completely and attempt to answer question four thorough other methods. It was also agreed not to include the school archival variables in the analysis again due to the fact that only four schools, at most, were available for this analysis.

The alternative method chosen to answer research question four was to conduct multiple regression using the OSCSS survey with the student, rather than the school, as the unit of analysis. The shift from the school to the student as the unit of analysis made possible multiple regression due to the substantially increased sample sizes. This analysis also made possible an at least partial answer to question four.

*Research Design*

The study used a between-subjects design with the subjects being individual schools and students. Chronbach's alpha internal consistency reliability coefficients were computed on each of the subscales of the SET, OSSS, OSCSS, and EBS in order to determine the internal consistency of these tools. The items included in each subscale in computing the alpha coefficients were those specified by each survey's authors.

Following computation of the subscale alphas the SET, OSSS, and the EBS were factor analyzed. The OSCSS was not factor analyzed due to the fact that it's factor structure has already been established by prior research. Each survey was factor analyzed using the Maximum Likelihood extraction procedure. Both Oblique and Orthogonal rotations were computed and compared using the Promax and Varimax rotation procedures, respectively. A final factor solution with rotation for each survey was selected.

Based on the results of the factor analysis factor score were computed for each survey by summing the item scores within each factor. In many cases multiple surveys were completed for each school. Because the unit of analysis is the school all surveys for a school were averaged in order to produce one summary score for each survey for each school. In order to examine the relationships between the different survey factors the summary scores for each survey factor were next correlated using the Pearson Correlation coefficient.

Finally, subscales of the OSCSS were regressed onto the three OSCSS school safety/disorder subscales. Standard multiple regression equations were computed for each of the school safety/disorder subscales including: Student Victimization, Student Rebellious Behavior, and Drub Availability. All available OSCSS subscales were entered into the equation simultaneously for each of the three analyses.

# *CHAPTER 4*
## Results

The results section is divided into four parts, each representing one of the four analytic procedures used in the study. The results are presented in the order the research questions were posed. First presented were the results of the internal consistency analyses of the subscales of the SET, EBS 1.5, and OSSS including item analysis of subscales with low alpha coefficients. Next presented were the results of the factor analyses of the above three measures. The third section is a presentation of the correlations between each of the factors derived in the previous analysis. Finally, the fourth section presents the results of the multiple regression analyses used to predict each of the school safety/disorder OSCSS subscales.

### *Internal Consistency*

To address the first research question intended to assess the internal consistency of the OSSS, EBS 1.5, SET, and the OSCSS, Chronbach alphas were computed for each at the full survey and subscale levels. There exist a number of different methods for computing internal consistency within a single measure including: Chronbach Coefficient Alpha, Kuder Richardson (KR) Formula, and Split-half Reliability Coefficient. However, Chronbach Alpha is recommended over the other two for a number of reasons. The first is that Chronbach Alpha can be applied to a broader range of data types. In contrast, KR can only be used with dichotomously scored data. A second reason relates to a potential consequence of the way in which Split-half reliability is calculated. Split-half estimates reliability by computing the correlation between two subsets of the data. Consequently reliability is a product of the way data is grouped and thus reliability may vary from group to group. In contrast, Chronbach Alpha is the mean of all possible split-half coefficients and so is not affected by the grouping of data (Yu, 2004).

Chronbach's coefficient alpha is computed as the ratio of true score variance to observed score variance, which typically ranges between 0 and 1. However, there is no lower limit to the coefficient. The closer Cronbach's alpha coefficient is to 1 the greater the internal consistency of the items in the scale. The size of the alpha coefficient is determined both by the number of items in the scale and the mean inter-item correlations. Conventions have been recommended for interpreting alpha levels: "≥ .9 – Excellent, ≥ .8 – Good, ≥ .7 - Acceptable, ≥ .6 - Questionable, ≥ .5 – Poor, and ≥ .5 – Unacceptable" (George & Mallory, 2003, p.231). One common mistake often made in interpreting high alpha coefficients is that this indicates a scale to be unidimensional. Obtaining a high alpha coefficient indicates only that the scale has good internal consistency and does not rule out the possibility that the scale is multidimensional. Factor analysis is necessary to determine the dimensionality of a scale (Gliem & Gliem, 2003; Yu, 2004).

#### *Oregon School Safety Survey (OSSS)*

The OSSS consists of 33 items divided into two scales including: Risk Factors for School Safety and Protective Factors for School Safety. A variety of school staff were asked to indicate the extent to which a set of risk and protective factors exist in their school and neighborhood using a five-point scale.

Chronbach alpha coefficients were computed on each of the two scales and on all items in the survey as a whole. These coefficients, both at the scale and survey level, were found to be consistently high with all falling in the "excellent" range ($\geq .9$). These results indicate that both of the OSSS subscales demonstrate strong internal consistency and that the instrument as a whole also demonstrates strong internal consistency among the individual items. Table 3 below contains summary statistics and coefficient alphas for the OSSS.

**Table 3**

*OSSS Subscale and Survey Chronbach Alpha Coefficients and Summary Statistics*

| Scale | α | N | M | SD | Items in Scale |
|---|---|---|---|---|---|
| Risk Factors | .90 | 2665 | 35.98 | 12.00 | 17 |
| Protective Factors | .91 | 2666 | 44.57 | 11.33 | 16 |
| Full Survey | .90 | 2664 | 80.54 | 18.11 | 33 |

*Effective Behavior Support Survey 1.5 (EBS 1.5)*

The EBS 1.5 consists of 43 questions divided into four subscales each assessing one of four school behavior support systems. Each question asks respondents to rate both the *current status* (CS) and *priority for improvement* (PI) for a feature of the behavior support system. Multiple school personnel complete the survey based on their experience and perceptions.

Chronbach's alpha coefficients were computed on each of the four subscales. Each subscale assessed one of four student behavior support systems and included responses addressing both current status and the priority for improvement. Alpha coefficients were computed on both the CS and PI responses for each subscale. Alpha coefficients were also computed on all items assessing PI and CS collapsing across the four subscales. Alpha coefficients for each of the subscales fell in the "excellent" category and ranged from .89 - .96. This indicates that all four subscales demonstrate strong internal consistency. There did not appear to be substantial differences within subscales on CS and PI alpha coefficients thus it appears that each subscale demonstrates consistent internal consistency when assessing both current status and priority for improvement. Finally, alpha coefficients for CS and PI collapsing across the subscale were both excellent and did not differ substantially from each other. This further confirms that the EBS 1.5 assesses the current status and priority for improvement of the four behavior support systems with excellent internal consistency. Table 4 below contains summary statistics and alpha coefficients for the EBS 1.5.

Table 4

*EBS 1.5 Chronbach Alpha Coefficients and Summary Statistics*

| Scale | α | N | M | SD | Items in Scale |
|---|---|---|---|---|---|
| School-Wide-CS | .82 | 1311 | 38.08 | 7.88 | 18 |
| School-Wide-PI | .91 | 1311 | 29.79 | 12.33 | 18 |
| Non Classroom Setting-CS | .86 | 1337 | 19.14 | 5.36 | 9 |
| Non Classroom Setting-PI | .92 | 1337 | 14.41 | 7.76 | 9 |
| Classroom-CS | .95 | 1337 | 24.04 | 8.83 | 11 |
| Classroom-PI | .96 | 1337 | 16.45 | 10.65 | 11 |
| Individual Student-CS | .89 | 1337 | 14.68 | 5.9 | 8 |

| | | | | | |
|---|---|---|---|---|---|
| Individual Student-PI | .93 | 1337 | 12.93 | 7.84 | 8 |
| Priority for Improvement* | .96 | 1311 | 73.26 | 34.27 | 46 |
| Current Status* | .96 | 1311 | 96.20 | 22.54 | 46 |

***Note.* CS (Current Status); PI (Priority for Improvement)**

*Includes all CS or PI item responses collapsed across the scales.

*Systems-Wide Evaluation Tool (SET)*

The SET consists of 29 questions divided into seven subscales each assessing one of seven school behavior support systems. The SET is designed as a research tool and is typically completed by an evaluator independent of the school with input from staff interviews, direct observation of the school facilities, and review of archival records and documents. Each question asks the evaluator to rate a feature of the school's behavior support systems on a 3-point scale.

Chronbach alpha coefficients were computed on each of the seven subscales and on the full survey. Alpha coefficients demonstrated a broad range from a low of .29 to a high of .91. Five of the eight alpha coefficients were in the excellent to good range indicating that four subscales and the survey as a whole demonstrate strong internal consistency. In contrast, the alpha coefficients for the subscales Expectations Defined, System of Responding to Behavior Violations, and District Level Support were in the poor to unacceptable range. These results indicate that these subscales have low internal consistency. Table 5 below contains summary statistics and coefficient alphas for the SET.

Table 5

**SET Chronbach Alpha Coefficients and Summary Statistics**

| Scale | α | N | M | SD | Items in Scale |
|---|---|---|---|---|---|
| Expectations Defined | .56 | 331 | 2.71 | 1.28 | 2 |
| Beh. Expectations Taught | .83 | 271 | 5.70 | 3.43 | 5 |
| On-Going Sys. of Reward | .91 | 438 | 3.49 | 2.53 | 3 |
| Sys. of Resp. to Beh. Viol. | .41 | 268 | 5.15 | 1.87 | 4 |
| Monitoring & Decision Making | .79 | 331 | 5.97 | 2.40 | 4 |
| Management | .85 | 329 | 11.75 | 4.48 | 8 |
| District Level Support | .29 | 142 | 3.56 | 1.53 | 3 |
| Full Survey | .90 | 127 | 35.14 | 12.07 | 29 |

As noted above, three subscales demonstrated alpha coefficients in the poor to unacceptable range. An alpha coefficient is determined by the number of items in a scale and the mean inter-item correlation. One reason for the low alpha coefficients obtained is very likely the low number of items comprising the subscales. This may especially be the case with the Expectations defined subscale, which only contains two items. A second reason for the low alpha coefficients is very likely the low inter-item correlations. If the items in a subscale do not correlate strongly with each other this will reduce the mean inter-item correlation and thus the alpha coefficient (Spector, 1992). Similarly, if the majority of items correlate strongly with each other but one or two items correlate poorly with the other items these one or two items will lower

the alpha coefficient. When items that do not correlate well with the other items in a subscale are found, deleting these items from the subscale may raise the subscale alpha coefficient. In order to explore the possibility that low inter-item correlations were accounting for the low observed alpha coefficients, item-analysis was conducted on the three lowest subscales.
The Expectations Defined subscale was examined first. This subscale contains two items both of which demonstrated identical and acceptable corrected item-total correlations (r = .40). This finding suggests that the low subscale alpha coefficient observed is most likely a result of the limited number of items in the subscale and is not caused by a low item-total correlation of any single item.

The System of Responding to Behavioral Violations subscale was examined next. This subscale contains four items with corrected item-total correlations ranging from r = .23 - .13. Question three of the subscale demonstrated a substantially lower corrected item-total correlation (r = .13) than did the other items. This finding suggests that item three may be lowering the mean item-total correlation and thus the alpha coefficient. Further examination revealed that if item three were deleted from the subscale the predicted alpha coefficient would raise to $\alpha = .44$. Although the deletion of item three of the subscale would raise the alpha coefficient it would not raise it by much and the improvement would not raise the alpha coefficient to an acceptable level. It appears more likely that the primary cause of the low subscale alpha coefficient is the low corrected item-total correlations across all of the items in the subscale. The number of items in the scale may also account for the low alpha coefficient.

Finally, the District Level Support Subscale was examined. This subscale contains three items with corrected item-total correlations ranging from r = .22 - .08. Question Three of the subscale demonstrated a substantially lower corrected item-total correlation (r = .08) relative to the other items. This finding suggests that item three may be lowering the mean item-total correlation and thus the alpha coefficient. Further examination revealed that if item three were deleted from the subscale the predicted alpha coefficient would raise to $\alpha = .36$. Although the deletion of item three of the subscale would raise the alpha coefficient, the improvement would not be much and would not raise the subscale alpha coefficient to an acceptable level. It appears more likely that the primary cause of the low subscale alpha coefficient is the low corrected item-total correlations across all items of the subscale. The subscale alpha coefficient might also be improved by adding additional items with strong item-total correlations to the subscale.

*Oregon School Climate and Safety Survey (OSCSS)*

The OSCSS consists of two self-report surveys, Form A and Form B, which are administered simultaneously. The OSCSS was adapted from the Effective Schools Battery (ESB) (Gottfredson, 1984) and the What About You? (WAY) survey (Gottfredson & Gottfredson, 1989) with permission from the author. Forms A and B divide the 15 scales of the ESB and WAY into two roughly equal length forms. Forms A and B each include 50 items and seven and eight scales, respectively. Due to differences in response metrics across items, which include two-point, three-point, and four-point scales, all item raw scores were converted to z-scores before analysis. The OSCSS was administered to a sample of 665 students in two high schools and two middle schools, both in Oregon.

Chronbach alpha coefficients were computed on each of the 15 scales and on all items in each of the two surveys. Scale alpha coefficients ranged from excellent ($\alpha = .99$) to unacceptable ($\alpha = .40$) with the majority in the acceptable and above categories. Survey level alpha coefficients for both Form A and B were excellent ($\alpha$ = .91, .99, respectively). These results indicate that both Form A and Form B demonstrate excellent overall internal consistency. Examination of the individual scales indicates that 10 out of 15 scales demonstrated at least acceptable levels of internal consistency. Of the remaining five scales two demonstrated unacceptably low alpha coefficients (Student Influence $\alpha$ = .40, Planning and Action $\alpha$ = .49) and the remaining three fall in the questionable to poor range (School Rewards $\alpha$ = .63, Respect for Students $\alpha$ = .63, Fairness of Rules $\alpha$ = .58). Table 6 below contains summary statistics and coefficient alphas for the OSCSS. Included for comparison are the published Effective Schools Battery alpha coefficients.

Table 6
OSCSS Chronbach Alpha Coefficients and Summary Statistics

| Scale | Published ESB $\alpha$ | Form A/B $\alpha$ | N | M | SD | Scale Items |
|---|---|---|---|---|---|---|
| Positive Peer | .65 | .99 | 571 | -1.27 | 8.54 | 9 |
| Attachment to School* | .76 | .75 | 297 | -.03 | 5.22 | 9 |
| Belief in Rules* | .53 | .74 | 322 | .02 | 3.93 | 6 |
| Social Integration+ | .51 | .98 | 647 | -.12 | 5.77 | 6 |
| School Rewards* | .56 | .63 | 327 | -.02 | 2.72 | 4 |
| Student Safety+ | .90 | .99 | 638 | -.66 | 17.37 | 18 |
| Respect for Students* | .82 | .63 | 327 | -.01 | 2.28 | 3 |
| Clarity of Rules+ | .71 | .98 | 651 | -.06 | 3.85 | 4 |
| Student Influence* | .70 | .40 | 318 | -.04 | 2.30 | 5 |
| Planning and Action* | .83 | .49 | 315 | .00 | 2.11 | 3 |
| School Effort+ | .59 | .99 | 657 | -.04 | 5.83 | 6 |
| Fairness of Rules* | .76 | .58 | 325 | .01 | 2.21 | 3 |
| Drug Availability+ | | .96 | 6579 | -.01 | 3.80 | 4 |
| Rebellious Behavior* | | .81 | 1341 | 32.9 | 4.37 | 14 |
| Student | | .80 | 1348 | 18.3 | 2.58 | 7 |
| Form A | NA | .91 | 256 | .16 | 20.90 | 47 |
| Form B~ | NA | .99 | 551 | -8.82 | 42.57 | 46 |

*Scales include on Form A +Scales included on Form B ~ Excluded Items B1-B4 due to zero variance

As noted above, two subscales demonstrated alpha coefficients in the unacceptable range, and three demonstrate questionable to poor alpha coefficients. Alpha coefficients are determined by the number of items in a scale and the mean inter-item correlation. One possible reason for the low alpha coefficients obtained may be the low number of items comprising the subscale. This may especially be the case with the Respect for Students, Planning and Action, and Fairness of Rules subscales, each of which contain only three items. A second reason for the low alpha coefficients may be low inter-item correlation. If the items in a subscale do not correlate strongly with each other, this will reduce the mean inter-item correlation and thus the alpha coefficient. Similarly, if the majority of items correlate strongly with each other but one or two items correlate poorly with the other items these one or two items will lower the alpha coefficient. When items that do not correlate well with the other items in a subscale are found, deleting these items from the subscale may raise the subscale alpha coefficient. In order to explore the possibility that low inter-item correlations were accounting for the low observed alpha coefficients, item-analysis was conducted on the five lowest subscales.

The Student Influence subscale was examined first. The subscale contains five items all of which demonstrate similarly low corrected item-total correlations ranging from -.14 to .17. An examination of the predicted scale alpha coefficient following deletion of individual items indicated that deletion of three of the four items would have resulted in a substantially lower alpha coefficient ranging from .11 to .13. Deletion of the fourth item would not have affected the scale alpha. This item also demonstrated the lowest corrected item-total correlation. It is also noteworthy that the standard deviation of the items within the subscale was low (SD = 2.30). The low item-total correlations and the low subscale standard deviation would appear to account for the low subscale alpha. It appears likely that the low standard deviation of the subscale may be a product of the sample, which, in turn, may result in low item-total correlations and the low subscale alpha. A previously published alpha coefficient included in table 1.4 indicates an alpha coefficient of .70, which is substantially higher than that derived in the present sample. This fact suggests that the sample, and it's low standard deviation, accounts for the low subscale alpha rather than the items in the subscale themselves.

The Planning and Action subscale was next examined. The subscale includes three items all of which demonstrate similarly low corrected item-total correlations ranging from .39 -.25. An examination of the predicted scale alpha coefficient following deletion of individual items indicated that deletion of the first item would have resulted in a substantially lowered alpha (.24) while deletion of the remaining two items would not have substantially affected the subscale alpha lowering it to .43 or .48. This indicates that none of the items are substantially lowering the subscale alpha. It is noteworthy that the subscale standard deviation is quite low (SD = 2.11). The low item-total correlations and the low subscale standard deviation would appear to account for the low subscale alpha. It appears likely that the low standard deviation of the subscale may be a product of the sample, which, in turn, may result in low item-total correlations and the low subscale alpha. A previously published alpha coefficient included in table 6 indicates an alpha coefficient of .83, which is substantially higher than that derived in the present sample. This fact suggests that the sample, and its low standard deviation, accounts for the low subscale alpha rather than the items in the subscale themselves.

Finally, the three subscales with questionable to poor subscale alphas were examined. The School Rewards subscale includes four items with corrected item-total correlations ranging from .28 - .46. An examination of the predicted scale alpha coefficient following deletion of individual items indicated that deletion of none of the five items would have resulted in a higher subscale alpha. The subscale alpha that would have resulted from the deletion of any one item would have ranged from .46 - .60, thus none of the five items appear to be lowering the subscale alpha. The School Rewards subscale demonstrated a moderately low standard deviation coefficient (SD = 2.72). It is interesting to note that the published alpha coefficient (.56) for this subscale in table 1.4 is somewhat lower than the subscale alpha in the present sample. This finding suggests that the somewhat low alpha coefficient of .63 appears to be valid and due more to the items and construction of the subscale itself than to characteristics of the sample. The moderately low subscale alpha of .63 appears to be primarily a product of the low corrected item-total correlations of the items and may suggest a need for improvements to the questions and number of questions that comprise the subscale.

The Respect for Students subscale includes three items with similar corrected item-total correlations ranging from .39 - .52. An examination of the predicted scale alpha coefficient following deletion of individual items indicated that deletion of any of the three items individually would have resulted in a lowered subscale alpha ranging from .42 - .60, thus none of the three items appear to be lowering the subscale alpha. The Respect for Students subscale demonstrated a moderately low standard deviation (SD = 2.28). The somewhat low item-total correlations and the low subscale standard deviation would appear to account for the low subscale alpha. It appears likely that the low standard deviation of the subscale may be a product of the sample, which, in turn, may result in low item-total correlations and the low subscale alpha. A previously published alpha coefficient included in table 6 indicates an alpha coefficient of .82, which is moderately higher than that derived in the present sample. This fact suggests that the sample, and its low standard deviation, accounts for the low subscale alpha rather than the items in the subscale themselves. However, it should be noted that the subscale alpha derived from the present sample falls in the questionable range rather than the unacceptably low range (George & Mallory, 2003, p.231).

Finally, the Fairness of Rules subscale was examined. The subscale includes three items with corrected item-total correlations ranging from .34 - .43. An examination of the predicted scale alpha coefficient following deletion of individual items indicated that deletion of any of the three items individually would have resulted in a lowered subscale alpha ranging from .41 - .54, thus none of the three items appear to be lowering the subscale alpha. The Fairness of Rules subscale demonstrated a moderately low standard deviation (SD = 2.21). The somewhat low item-total correlations and the low subscale standard deviation would appear to account for the low subscale alpha. It appears likely that the low standard deviation of the subscale may be a product of the sample, which, in turn, may result in low item-total correlations and the low subscale alpha. A previously published alpha coefficient included in table 1.4 indicates an alpha coefficient of .76, which is moderately higher than that derived in the present sample. This fact suggests that the sample, and its low standard deviation, accounts for the low subscale alpha rather than the items in the subscale themselves. However, it should be noted that the subscale alpha derived from the present sample falls closer to the questionable range rather than the unacceptably low range (George & Mallory, 2003, p.231). Having presented the results of the internal consistency analyses of the SET, OSSS , and EBS 1.5 based on the subscales specified by the measures' authors I will now present the results of the factor analyses of each of these measures.

*Methods for Conducting Factor Analyses*

The following procedures were used in conducting exploratory factor analysis for each measure. Summary statistics were first computed and examined to determine whether the sample was appropriate for factor analysis. Summary statistics computed and examined included the following: Bartlett's Test of Sphericity, used to test the null hypothesis that the correlation matrix is an identity matrix. If the statistic is significant one may reject the null hypothesis and proceed with the factor analysis. If the statistic is not and you cannot reject the null hypothesis caution is urged in proceeding with the factor analysis (Norusis, 1994); The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) statistic was also computed. The KMO compares the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Small values for the KMO measure suggest that a factor analysis of the variables may not be appropriate, since correlations between pairs of variables cannot be explained by the other variables (Norusis, 1994). Kaiser (1974) has provided guidelines for interpreting the

KMO. Scores in the .90's are considered *marvelous*, in the .80's *meritorious*, in the .70's *middling*, in the .60's *mediocre*, in the .50's *miserable*, and values below .50 are considered *unacceptable*.

Following examination of the above noted summary statistics and a determination that the sample is appropriate for factor analysis an initial factor solution was extracted using the Maximum Likelihood method. The next step was to determine the number of meaningful factors to be retained. Hatcher (1994) has recommended the use of four procedures in a structured sequence when making a determination of the number of factors to retain. These procedures include: (1) retaining and interpreting only components with an eigenvalue greater than one; (2) examining the scree test and looking for a "break" between components with relatively large eigenvalues and those with small eigenvalues; (3) examining the percentage of total variance accounted for by each factor and deciding on a cut-off below which factors are no longer retained. For example, one might choose to retain only those factors, which account for at least five percent of the total variance; and finally (4) examining the factor solution for ease of interpretation. With the exception of the procedure number one, all of the above procedures were used in sequence to determine the number of factors to be retained. Procedure number one was not used due to the differences between Principal Components Analysis and Factor Analysis, which make the procedure inappropriate for Factor Analysis (for a discussion see Hatcher, 1994, pp.81-82).

Once the determination of the number of factors to be retained had been made the specified number of factors were extracted using the Maximum Likelihood extraction method and rotated. Two rotations were computed including Varimax and Promax. Both an orthogonal and oblique rotation were derived for the purpose of comparison and in order to determine the most appropriate rotation. A tension typically exists between the use of orthogonal and oblique rotations for a number of reasons. As Tabachnick and Fidell (1996) noted, "Orthogonal solutions offer ease of interpreting, describing, and reporting results; yet they 'strain' reality unless the researcher is convinced that the underlying processes are almost independent" (pp. 666). Similarly, although an oblique rotation may be more appropriate to the data there may be compelling reasons for the use of an orthogonal rotation, such as the need to have orthogonal factors for use in further analyses. Because of this tension both oblique and orthogonal rotations were derived and compared using a number of criteria, to determine the model which best fit the data.

A number of procedures were used in combination to determine the "best" factor solution. These included a comparison of the "goodness-of-fit" Chi-Square statistics and their significance levels. It should be noted that with large sample sizes the significance level of the Chi-Square statistic may not be a good indicator of the goodness-of-fit of the model. This is due to the fact that with large sample sizes small differences may be statistically significant but not large enough to be relevant or meaningful, thus a model may provide a good fit despite the fact that the Chi-Square statistic is significant (Hatcher, 1994). Instead of relying on the significance levels, the relative differences between the models on the Chi-Square statistic was emphasized with lower Chi-Square statistics indicating a better fit of the model.

A second procedure used was to examine the rotated factor matrices for "simple structure" and interpretability. Simple structure is defined by two characteristics. The first is that most of the variables have relatively high factor loadings on only one factor, and near-zero loadings on the other factors. The second is that most factors have relatively high factor loadings for some variables, and near-zero loadings for the remaining variables. The interpretability criteria is defined by: (a) At least three variables loading significantly on each retained factor, (b) the variables that load on a given factor share the same conceptual meaning, and (c) the variables that load on different factors seem to measure different constructs (Hatcher, 1994, pp.86, 92). For the purposes of this study items that load on a factor $\geq .30$ will be considered significant. The rationale for this is that factors that load at .30 share nine percent of their variance with the factor. Cohen (1988) has proposed that correlations between .30 and .49 be considered of "medium" size. Thus shared variance of nine percent constitutes a meaningful overlap and contribution to the factor. Only items that meet the minimum criteria of loading .30 will be interpreted.

A third procedure used to assess the models was to examine the residuals matrix of each solution to determine the magnitude of the residuals and thus the fit of the solution to the population. A good solution is indicated when the residuals, which represent the differences for each variable between the observed and reproduced correlation matrices, are small. Large residuals indicate that there is a substantial discrepancy between the observed and reproduced correlation matrices and thus the solution is a poor fit to the data (Tabachnick & Fidell, 1996, pp.648). If statistically significant residuals remain after extraction this may also indicate that another factor should be extracted (Gorsuch, 1983, pp.148).

A fourth procedure used in determining the best factor model was to examine the squared multiple correlations (SMCs) of the factors. The SMC consists of the squared correlations of factor scores predicted from scores on observed variables. The SMCs provide an estimate of the internal consistency of the solution. In a good solution, SMCs range between 0 and 1 with larger SMCs indicating more stable factor solutions. A large SMC ($\leq .70$) indicates that the observed variables account for substantial variance in the factor scores. A low SMC indicates that the factors are poorly defined by the observed variables. If an SMC is negative too many variables have been retained and if an SMC exceeds 1 the entire solution needs to be reevaluated (Tabachnick & Fidell, 1996).

Finally, a fifth procedure used to determine whether an orthogonal or oblique rotation was more appropriate was to look at the factor correlation matrix. The factor correlation matrix provides factor correlation coefficients for each factor with each other. Tabachnick and Fidell (1996, pp. 674) recommended examining the factor correlation matrix for correlations around .32 and above. When correlations exceed .32 this indicates 10% (or more) overlap in variance among the factors, which is enough to warrant oblique rotation unless there are compelling reasons for orthogonal rotation.

### *Factor Analysis*

Three measures were factor analyzed using exploratory factor analysis including the OSSS, SET, and EBS 1.5 in order to examine and establish their factor structures. The above-discussed procedures were used in sequence on all the measures unless otherwise noted, in which case the changes in the type and order of procedures will be discussed.

#### *Oregon School Safety Survey*

Summary statistics were computed to determine the appropriateness of factor analysis to the sample. Both the KMO statistic (.94) and the Bartlett's Test of Sphericity $\chi 2(528, N = 2{,}668) =$

38,198.27, $p < .01$ were found to be well within their acceptable ranges thus indicating that the sample was appropriate for factor analysis.

An initial extraction was made using the Maximum Likelihood procedure in order to determine the number of factors to retain for rotation. Using the default of eigenvalue greater than one, five factors were extracted. In order one through five the eigenvalues, and percent of the total variance, were 8.35 (25.29%), 5.95 (18.03%), 1.44 (4.37%), 1.18 (3.58%), and 1.05 (3.18%), respectively. A substantial drop in eigenvalues occurred between factors two and three. Similarly, an examination of the scree test indicated a break between factors two and three. Considered together these results suggest a two-factor solution.

Following the initial extraction and determination that two factors should be retained, a second extraction was made using the Maximum Likelihood procedure specifying a two-factor solution. Following the two-factor extraction the first and second extraction sums of squared loadings and percentage of variance accounted for were 7.62 (23.10%), and 5.52 (16.71%), respectively. Both factors combined accounted for a cumulative total of 39.80% of the variance. This solution was then rotated using both the Varimax and the Promax methods, resulting in both an orthogonal and oblique rotation. The Goodness-of-fit statistic $\chi2(463, N = 2{,}668) = 5{,}490.46$, $p < .01$ was significant suggesting that more than two factors may be needed to appropriately model the data. However, as previously noted, with large sample sizes this statistic may not be a good indicator of the goodness of fit of a factor model. In attempt to find a "good fit" additional factors were added to the model. However, even with the addition of seven factors the statistic remained significant $\chi2(318, N = 2{,}668) = 890.55$, $p < .01$. It was determined, based on the scree test and an examination of the portion of variance accounted by the factors, that a two-factor solution was appropriate. An examination of the residuals matrix using a two-factor solution also suggested that a two-factor model was appropriate. The residuals ranged from .254 to 5.66-E04 with 83 (15%) of the residuals exceeding absolute values greater than .05. These results indicate low residual values and thus a good fit between the observed correlation matrix and the reproduced correlation matrix. Similarly, examination of the SMCs for factors one and two revealed values of .93 and .91, respectively. Squared multiple correlations greater than or equal to .70 indicate that the observed variables account for substantial variance in the factor scores (Tabachnick and Fidell, 1996). Taken together these results indicate that a two-factor solution provides a good fit to the data. Table 7 below contains the factor loadings for the orthogonal rotation while table 8 contains the factor loadings for the oblique rotation.

Table 7
OSSS Factor Matrix
(Orthogonal Rotation-Varimax)

| Survey Question | Factor 1 | Factor 2 |
|---|---|---|
| 01 | | **.60** |
| 02 | | **.66** |
| 03 | | **.52** |
| 04 | | **.66** |
| 05 | | **.69** |
| 06 | | **.64** |
| 07 | | **.60** |
| 08 | | **.54** |
| 09 | | **.52** |
| 10 | | **.44** |
| 11 | | **.50** |
| 12 | | **.49** |
| 13 | | **.72** |
| 14 | | **.67** |
| 15 | | **.73** |
| 16 | | **.64** |
| 17 | | **.46** |
| 18 | **.51** | |
| 19 | **.60** | |
| 20 | **.62** | |
| 21 | **.66** | |
| 22 | **.62** | |
| 23 | **.62** | |
| 24 | **.64** | |
| 25 | **.72** | |
| 26 | .22 | .29 |
| 27 | **.66** | |
| 28 | **.81** | |
| 29 | **.70** | |
| 30 | **.70** | |
| 31 | **.54** | |
| 32 | **.79** | |
| 33 | **.76** | |

*Note.* Values less than .20 were suppressed to

aid in interpretation.

Table 8
OSSS Pattern Matrix
(Oblique Rotation-Promax)

| Survey Question | Factor 1 | Factor 2 |
|---|---|---|
| 01 | | **.64** |
| 02 | | **.66** |
| 03 | | **.50** |
| 04 | | **.59** |
| 05 | | **.57** |
| 06 | | **.65** |
| 07 | | **.62** |
| 08 | | **.56** |
| 09 | | **.53** |
| 10 | | **.56** |
| 11 | | **.49** |
| 12 | | **.48** |
| 13 | | **.74** |
| 14 | | **.69** |
| 15 | | **.70** |
| 16 | | **.66** |
| 17 | | **.37** |
| 18 | **.46** | |
| 19 | **.53** | |
| 20 | **.55** | |
| 21 | **.62** | |
| 22 | **.60** | |
| 23 | **.60** | |
| 24 | **.64** | |
| 25 | **.68** | |
| 26 | .23 | **.34** |
| 27 | **.66** | |
| 28 | **.83** | |
| 29 | **.71** | |
| 30 | **.70** | |
| 31 | **.51** | |
| 32 | **.80** | |
| 33 | **.77** | |

*Note.* Values less than .20 were suppressed to

aid in interpretation.

An examination of the factor loadings in tables 7 and 8 indicates that both the orthogonal and oblique solutions demonstrate simple structure. The majority of survey items demonstrate high factor loadings on only one factor and near zero loadings on the other factor. It should be noted that in both tables factor loadings less than .20 were suppressed to assist in interpretation. With the exception of survey item 26, items 1 - 17 comprise one factor and items 18 - 33 comprise a second. Both solutions also meet the interpretability criteria previously discussed. Items 1 - 17 comprise one factor, which perfectly overlaps with the items assessing risk factors for school safety as specified in the survey design. Items 18 - 33 comprise a second factor, which perfectly overlaps, with the exception of item 26, with the items assessing protective factors for school safety as specified in the survey design.

Based on these findings, both the orthogonal and oblique two-factor solutions appeared to provide an appropriate fit to the data. However, an examination of the factor correlation matrix for the oblique solution indicated that the orthogonal 2-factor solution provides a better fit to the data. Table 9 represents the factor correlation matrix for the oblique factor model.

Table 9
Factor Correlation Matrix (Promax)

| Factor | 1 | 2 |
|---|---|---|
| 1 | 1.0 | .03 |
| 2 | .03 | 1.0 |

Table 9 provides the correlations between each of the two factors in the oblique rotation. As may be seen, the absolute value of the correlation between factors one and two is .03, which indicates a very weak correlation between the two factors and a shared variance of essentially 0%. This indicates that the factors are independent of one another and thus an orthogonal rotation is more appropriate than an oblique rotation. Thus a 2-factor solution with orthogonal rotation appears to produce the most appropriate factor model.

*Effective Behavior Support Survey*

The EBS Survey asks respondents to rate each item on two dimensions: Current Status and Priority for Improvement. Because of this the EBS Survey essentially consists of two identical surveys with the exception that one queries current status while the other queries priority for improvement. Consequently, both the current status and priority for improvement responses to each item could not be combined in a single factor analysis. Instead, separate factor analyses were performed on the items within each response category (e.g. Current Status or Priority for Improvement). The results of each factor analysis will be discussed separately beginning with the Current Status results.

*EBS-current status*

Summary statistics were computed to determine the appropriateness of factor analysis to the sample. Both the KMO statistic (.96) and the Bartlett's Test of Sphericity $\chi2(1035, N = 1337) = 33,494.39$, $p < .01$ were found to be well within their acceptable ranges thus indicating that the sample was appropriate for factor analysis.

An initial extraction was made using the Maximum Likelihood procedure in order to determine the number of factors to retain for rotation. Using the default of eigenvalue greater than one, seven factors were extracted. In order one through six the initial factor eigenvalues and percentage of variance accounted for were 15.04 (34.97%), 4.15 (9.16%), 2.39 (5.55%), 1.51 (3.52%), 1.20 (2.80%), and 1.07 (2.49%), respectively. A moderate drop in the percentage of variance accounted for occurred between factors three and four. However, factor four accounted for a meaningful percentage of the variance (3.52%). The fifth factor evidenced a slight drop in the percentage of variance accounted for (2.80%) relative to factor four and given the percentage of variance it accounted for appeared to be on the borderline on qualifying as an additional factor. Similar results were obtained from an examination of the scree test, which indicated a moderate break between factors three and four. Factor four and possibly factor five were observed to be on the tail end of the vertical plot of eigenvalues just above the point at which the eigenvalues leveled out into a horizontal line. These results were interpreted to indicate that factors one through three should be retained for rotation. It was less clear whether factor four should be retained.

In order to determine whether a three or four factor solution should be retained each of the above factor models were extracted and compared. Each model was compared on a number of the facets of simple structure including the number of factor loadings greater than or equal to .30 and the absolute value of the mean of factor loadings greater than or equal to .30. This was based on the assumption that factors with a greater number of factor loadings ($\geq .30$) and a higher mean factor loading would demonstrate relatively greater simple structure and thus interpretability.

The three- and four-factor solutions were compared. In the three-factor model the third factor included 10 meaningful factor loadings with a mean factor loading of .34. In the four-factor model the third factor included four substantial factor loadings with a mean factor loading of .61. The fourth factor included three substantial factor loadings with a mean factor loading of x= .31. The three- and four-factor models were further compared using the goodness-of-fit $\chi 2$ statistics. Lower $\chi 2$ statistics indicate a better fit of the factor model to the data. The three- and four-factor models produced goodness-of-fit statistics of $\chi 2(900, N = 1337) = 5276.20$, $p < .01$ and $\chi 2(857, N = 1337) = 1290.0$, $p = .00$, respectively. Although both models produced significant $\chi 2$ statistics, the four-factor model resulted in a 24% reduction in the $\chi 2$ statistic, relative to the three-factor model. This suggests that a four-factor model provides a better fit to the data than does a three-factor model. An examination of the residuals matrix using a four-factor solution also suggested that a four-factor model was appropriate. The residuals ranged from .198 to 2.60E-0four with 59 (6%) of the residuals exceeding absolute values greater than .05. These results indicate low residual values and thus a good fit between the observed correlation matrix and the reproduced correlation matrix. An examination of the SMCs for factors one through four indicated values of .98, .87, .82, and .65, respectively. Factors one through three clearly exceed the minimum recommended value of .70 (Tabachnick and Fidell, 1996). The fourth factor's SMC value is slightly below the recommended value of .70. However, taken together all of the evidence indicates that a four-factor model provides a better fit to the data than does a three-factor model. A four-factor model is also consistent with the explicit design of the EBS, which includes four scales. Based on the above-discussed comparison of different factor models and the four-scale design of the EBS, it was determined that a four-factor model should be retained for rotation.

Following the determination that four factors should be retained, a four-factor model was extracted using the Maximum Likelihood procedure. Following the four-factor extraction the first thru fourth extraction sums of squared loadings and percentage of variance accounted for were 14.01 (32.59%), 4.27 (9.92%), 1.88 (4.37%), and .93 (2.17%) of the total variance, respectively. All four factors accounted for a cumulative total of 49.05% of the variance. This solution was then rotated using both the Varimax and the Promax methods, resulting in both an orthogonal and oblique rotation. The goodness-of-fit statistic $\chi 2(857, N = 1337) = 1290.0$, $p < .01$ was significant however, as previously noted, with large sample sizes this statistic may not be a good indicator of the goodness of fit of a factor model. Table 10 below contains the factor loadings for the orthogonal rotation while table 11 contains the factor loadings for the oblique rotation.

Table 10
EBS-Current Status Factor Matrix
(Orthogonal Rotation-Varimax)

| Survey Item | Factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| CS101 | **.56** | | | |
| CS102 | **.67** | | | |
| CS103 | **.61** | | | |
| CS104 | **.61** | | | |
| CS105 | **.56** | | | |
| CS106 | **.51** | .23 | .25 | |
| CS107 | **.41** | .24 | .23 | |
| CS108 | **.40** | | | |
| CS109 | **.49** | | .23 | |
| CS110 | **.46** | | .27 | |
| CS111 | **.53** | | .32 | |
| CS112 | **.55** | | .39 | |
| CS113 | **.51** | | .29 | |
| CS114 | **.44** | | .37 | |
| CS115 | **.41** | | .35 | |
| CS201 | **.65** | | | |
| CS202 | **.62** | | | |
| CS203 | **.50** | | | |
| CS204 | **.61** | | | |
| CS205 | **.41** | | .26 | |
| CS206 | **.45** | | .25 | |
| CS207 | **.41** | | .34 | |
| CS208 | **.45** | | .44 | |
| CS209 | **.51** | | .25 | |
| CS301 | | **.87** | | |
| CS302 | .24 | **.83** | | |
| CS303 | .21 | **.86** | | |
| CS304 | .24 | **.77** | | |
| CS305 | .21 | **.79** | | |
| CS306 | .28 | **.76** | | |
| CS307 | .21 | **.73** | .23 | |
| CS308 | | **.76** | | |
| CS309 | | **.68** | .25 | |
| CS310 | .22 | **.60** | .33 | |
| CS311 | .20 | **.76** | | |
| CS401 | .27 | | **.62** | |
| CS402 | .31 | .26 | **.57** | |
| CS403 | .26 | | **.66** | |
| CS404 | .25 | | **.67** | |
| CS405 | | | **.65** | |
| CS406 | .24 | .26 | **.61** | |
| CS407 | | | **.53** | |
| CS408 | .31 | | **.70** | |
| CS116 | | | | **.70** |
| CS117 | | | | **.84** |
| CS118 | | | | **.72** |

*Note*. Values less than .20 were suppressed to

aid in interpretation.

Table 11
EBS-Current Status Pattern Matrix
(Oblique Rotation-Promax)

| Survey Item | Factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| CS101 | **.66** | | | |
| CS102 | **.79** | | | |
| CS103 | **.68** | | | |
| CS104 | **.68** | | | |
| CS105 | **.64** | | | |
| CS106 | **.50** | | | |
| CS107 | **.36** | | | |
| CS108 | **.44** | | | |
| CS109 | **.48** | | | |
| CS110 | **.44** | | | |
| CS111 | **.51** | | | |
| CS112 | **.51** | | .27 | |
| CS113 | **.51** | | | |
| CS114 | **.39** | | .28 | |
| CS115 | **.36** | | .28 | |
| CS201 | **.75** | | | |
| CS202 | **.70** | | | |
| CS203 | **.54** | | | |
| CS204 | **.67** | | | |
| CS205 | **.37** | | | |
| CS206 | **.43** | | | |
| CS207 | **.36** | | .28 | |
| CS208 | .37 | | **.37** | |
| CS209 | **.50** | | | |
| CS301 | | **.95** | | |
| CS302 | | **.88** | | |
| CS303 | | **.93** | | |
| CS304 | | **.80** | | |
| CS305 | | **.83** | | |
| CS306 | | **.77** | | |

| | | | |
|---|---|---|---|
| CS307 | **.75** | | |
| CS308 | **.80** | | |
| CS309 | **.71** | | |
| CS310 | **.58** | .25 | |
| CS311 | **.79** | | |
| CS401 | | **.67** | |
| CS402 | | **.59** | |
| CS403 | | **.73** | |
| CS404 | | **.76** | |
| CS405 | | **.75** | |
| CS406 | | **.67** | |
| CS407 | | **.59** | |
| CS408 | | **.77** | |
| CS116 | | | **.71** |
| CS117 | | | **.86** |
| CS118 | | | **.73** |

*Note*. Values less than .20 were suppressed to

aid in interpretation.

An examination of the factor loadings in table 10 indicates a lack of simple structure for the orthogonal solution. Although many items demonstrate high factor loadings on a single factor while loading minimally, if at all, on other factors, many load substantially on more than one factor. Nine variables demonstrated factor-loading coefficients of .30 or above on more than one factor. When an item loads substantially on more than one factor interpretation of the overall factor structure is made difficult. It should be noted that in tables 10 and 11 factor loadings less than .20 were suppressed to assist in interpretation.

In contrast, the factor loadings in table 11 indicate a simple structure in the oblique solution. With only one exception (CS208) all items demonstrate high factor loadings on only one factor and low, if any, loadings on other factors. Additionally, the oblique rotation improved the simple structure of the factor model by increasing single variable factor loading coefficients while decreasing the remaining factor loading coefficients across the other factors. Thus the oblique rotation improved the simple structure of the factor model by increasing single variable factor loading coefficients while decreasing the remaining factor loading coefficients across the remaining factors. Similarly, the oblique rotation also demonstrates interpretability, although to differing degrees across factors. Although all four factors have at least three variables, which load substantially, the fourth factor is comprised of only three variables. This suggests that factor four may not be adequately defined and measured. Factors two and three are represented by sets of variables that share the same conceptual meaning. However, the set of variables that load on factor one appears to represent two different constructs. Variables CS101 - CS118 were intended by the test authors to measure school-wide systems of positive behavior support while variables CS201 - CS209 were designed to measure nonclassroom setting systems of positive behavior support. Thus it appears that factor one includes two conceptually distinct sets of variables. However, it is also possible that although the test authors intended for the above set of questions to represent two distinct constructs in reality these two sets of questions actually represent a single construct. It may be that the two sets of questions assess a much broader construct which encompasses what the authors believed were two similar but distinct constructs. A reading of both sets of questions suggests that this may be the case. Both sets of questions appear to assess very similar ideas focused on school-wide positive behavior support features. The nonclassroom setting systems questions (CS201 - CS209) ask respondents to rate all nonclassroom settings including hallways, cafeteria, playground and bus. There would appear to be substantial differences across this range of nonclassroom settings and by asking respondents to rate this diverse range of settings as though they were a unified whole it appears likely that these questions, in effect, assess more school-wide systems. Had respondents been asked to rate each nonclassroom setting separately, rather than as a whole, this set of questions may have demonstrated sufficient variance to differentiate them into a separate factor. An examination of the factor correlation matrix for the oblique rotation also suggests that an oblique rotation provides the most appropriate fit to the data. Table 12 represents the factor correlation matrix for the oblique rotation.

Table 12
*Factor Correlation Matrix (Promax)*

| Factor | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1.00 | .52 | .65 | -.26 |
| 2 | .52 | 1.00 | .47 | -.16 |
| 3 | .65 | .47 | 1.00 | -.04 |
| 4 | -.26 | -.16 | -.04 | 1.00 |

As may be seen in table 12, many of the correlations between the factors exceed the recommended value of .32. An examination of the lower diagonal indicates that three of six values exceed .32 and range from an absolute value of .65 to .04. This indicates substantial correlations between several of the factors and shared variance ranging from .16% to 42%. Consequently, many of the factors do not appear to be orthogonal and thus an oblique rotation is more appropriate. The factor analytic results for the EBS Priority for Improvement responses will now be examined.

**EBS-Priority for Improvement**

Summary statistics were computed to determine the appropriateness of factor analysis to the sample. Both the KMO statistic (.98) and the Bartlett's Test of Sphericity $\chi 2(1035, N = 1337) = 46{,}805.00$, $p < .01$ were found to be well within their acceptable ranges thus indicating that the sample was appropriate for factor analysis.

An initial extraction was made using the Maximum Likelihood procedure in order to determine the number of factors to retain for rotation. Using the default of eigenvalue greater than one, five factors were extracted. In order one through five the initial eigenvalues, and percent of total variance, were 21.73 (47.23%), 2.82 (6.12%), 2.30 (5.01%), 1.52 (3.31%), and 1.15 (2.5%), respectively. A substantial drop in the percentage of variance accounted for occurred between factors one and two and a moderate drop between factors two and three, and three and four. The percentage of variance accounted for by the fifth factor was relatively small and thus it did not meaningfully contribute to the total percentage of variance. Similarly, the scree test indicated a substantial break between factors one and two, a minimal break between factors two and three, and three and four and a clear leveling off of the eigenvalues after the fourth factor. The results of the scree test along with the percentage of variance accounted for by each factor were interpreted as indicating that the first four factors should be retained for rotation.

Following the initial extraction and determination that four factors should be retained, a second extraction was made using the Maximum Likelihood procedure specifying a four-factor solution. Following the four-factor extraction the first thru fourth extraction sums of squared loadings and percentage of variance accounted for were 21.26 (46.22%), 2.50 (5.44%), 1.76 (3.83%), and .76 (1.65%) of the variance, respectively. The first thru fourth factors accounted for a cumulative total of 57.15% of the explained variance. This solution was then rotated using the Varimax and the Promax methods, resulting in both an orthogonal and oblique rotation. The Goodness-of-fit statistic $\chi 2(737, N = 1337) = 4058.11$, $p < .01$ was significant suggesting that more than four factors may be needed to appropriately model the data. However, as previously noted, with large sample sizes this statistic may not be a good indicator of the goodness of fit of a factor model. In order to test the possibility that an additional factor may be needed to provide an appropriate fit to the data a five-factor model was also computed. The resulting Goodness-of-fit statistic $\chi 2(698, N = 1337) = 3{,}192.53$, $p < .01$ decreased but remained significant. This indicates that more than five factors may be needed to reduce the Goodness-of-fit statistic to nonsignificance. However, the inclusion of five or more factors is not warranted based on the scree test and variance accounted for criteria thus a four-factor model appears to provide the most appropriate fit to the data. An examination of the residuals matrix using a four-factor solution also suggested that a four-factor model was appropriate. The residuals ranged from .137 to 1.038E-05 with 41 (4%) of the residuals exceeding absolute values greater than .05. These results indicate low residual values and thus a good fit between the observed correlation matrix and the reproduced correlation matrix. An examination of the SMCs for factors one through four indicated values of .98, .88, .81, and .66, respectively. Factors one through three clearly exceed the minimum recommended value of .70 (Tabachnick and Fidell, 1996). The fourth factor's SMC value is slightly below the recommended value of .70. However, taken together all of the evidence indicates that a four-factor model provides the best fit to the data. Table 13 below contains the factor loadings for the orthogonal rotation while table 14 contains the factor loadings for the oblique rotation.

Table 13

EBS-Priority for Improvement
Factor Matrix
(Orthogonal Rotation-Varimax)

| Survey | Factor | | | |
|---|---|---|---|---|
| Item | 2 | 3 | 1 | 4 |
| IP101 | .32 | .27 | **.59** | .32 |
| IP102 | .34 | .29 | **.59** | .31 |
| IP103 | .25 | .25 | **.53** | .34 |
| IP104 | .26 | .33 | **.68** | .22 |
| IP105 | .24 | .37 | **.61** | |
| IP106 | .21 | .34 | **.50** | .24 |
| IP107 | .31 | .37 | **.48** | .23 |
| IP108 | .26 | .27 | **.48** | .31 |
| IP109 | .27 | .40 | **.53** | .34 |
| IP110 | .26 | .35 | **.50** | .34 |
| IP111 | .26 | .42 | **.48** | .32 |
| IP112 | .24 | .46 | **.40** | .29 |
| IP113 | .34 | .40 | **.46** | .34 |
| IP114 | .22 | **.43** | .22 | .29 |
| IP115 | | **.44** | .26 | .31 |
| IP116 | | | **-.33** | |
| IP117 | | | **-.27** | |
| IP118 | | | | |
| IP201 | .34 | .29 | .47 | **.52** |
| IP202 | .27 | .29 | .36 | **.58** |
| IP203 | .30 | .33 | .32 | **.57** |
| IP204 | .24 | .31 | .32 | **.56** |
| IP205 | .23 | .34 | .25 | **.50** |
| IP206 | .30 | .32 | .30 | **.59** |
| IP207 | .22 | .38 | | **.54** |
| IP208 | .20 | **.52** | | .43 |
| IP209 | .26 | .36 | .33 | **.59** |
| IP301 | **.79** | | .32 | .22 |
| IP302 | **.77** | .23 | .35 | |
| IP303 | **.79** | | .29 | .23 |
| IP304 | **.72** | | | .22 |
| IP305 | **.76** | | .23 | |
| IP306 | **.72** | .27 | .28 | |
| IP307 | **.71** | .29 | .26 | |
| IP308 | **.72** | .22 | | |
| IP309 | **.70** | .30 | | |
| IP310 | **.62** | .36 | | |
| IP311 | **.66** | .26 | | .27 |
| IP401 | .29 | **.65** | .26 | .22 |
| IP402 | .35 | **.62** | .25 | .21 |
| IP403 | .30 | **.72** | | |
| IP404 | .29 | **.72** | .23 | .23 |
| IP405 | .23 | **.76** | | .21 |
| IP406 | .34 | **.67** | .21 | .22 |
| IP407 | .21 | **.64** | | .23 |
| IP408 | .26 | **.76** | | .22 |

*Note*. Values less than .20 were suppressed to

aid in interpretation.

Table 14
EBS-Priority for Improvement
Pattern Matrix
(Oblique Rotation-Promax)

| Survey | Factor | | | |
|---|---|---|---|---|
| Item | 2 | 3 | 1 | 4 |
| IP101 | | | **.65** | |
| IP102 | | | **.65** | |
| IP103 | | | **.57** | .22 |
| IP104 | | | **.87** | |
| IP105 | | | **.80** | -.21 |
| IP106 | | | **.56** | |
| IP107 | | | **.50** | |
| IP108 | | | **.49** | |
| IP109 | | | **.52** | |
| IP110 | | | **.49** | .21 |
| IP111 | | .24 | **.46** | |
| IP112 | | .34 | **.35** | |
| IP113 | | | **.40** | |
| IP114 | | **.37** | | .22 |
| IP115 | | **.37** | | .23 |
| IP116 | | | **-.45** | |
| IP117 | | .24 | **-.41** | |
| IP118 | | .26 | **-.34** | |
| IP201 | | | .32 | **.52** |
| IP202 | | | | **.68** |
| IP203 | | | | **.67** |
| IP204 | | | | **.67** |
| IP205 | | | | **.59** |
| IP206 | | | | **.70** |
| IP207 | | .21 | | **.66** |
| IP208 | | **.47** | | .43 |
| IP209 | | | | **.69** |
| IP301 | **.85** | | | |
| IP302 | **.81** | | .22 | |
| IP303 | **.87** | | | |

| | | |
|---|---|---|
| IP304 | **.82** | |
| IP305 | **.85** | |
| IP306 | **.75** | |
| IP307 | **.74** | |
| IP308 | **.83** | |
| IP309 | **.79** | |
| IP310 | **.65** | .29 |
| IP311 | **.71** | |
| IP401 | | **.70** |
| IP402 | | **.65** |

| | | |
|---|---|---|
| IP403 | **.86** | |
| IP404 | **.80** | |
| IP405 | **.95** | |
| IP406 | **.74** | |
| IP407 | **.81** | -.29 |
| IP408 | **.90** | |

*Note.* Values less than .20 were suppressed to

aid in interpretation.

An examination of the factor loadings in table 13 indicates a mixed outcome with regard to simple structure in the Orthogonal rotation. Although each factor is characterized by strong factor loadings across multiple variables, there also appear to be a large number of variables which load (≥ .30) on more than one factor. Such multiple loadings of a single variable on more than one factor are referred to as double loadings, which were substantial (44) in the Orthogonal rotation. The large number of double loadings indicates a lack of simple structure and makes difficult interpretation of the factor structure. It should be noted that factor loadings less than or equal to .20 were not reported to assist in interpretation.

In contrast, the pattern of factor loadings in table 14 indicates greater simple structure in the oblique rotation. The majority of variables load highly on one and only one factor with only 2 variables loading on more than one factor. The average factor-loading coefficient across all variables, which load on a single factor, improved in the oblique rotation across all factors. The substantially reduced number of double loadings and increase in the average factor loading coefficients indicates greater simple structure in the oblique relative to the orthogonal rotation. The oblique rotation also demonstrates strong conceptual consistency among the individual items loading within each factor and results in a factor model, which closely approximates the item-scale correspondence proposed by the test authors. With few exceptions conceptually similar items cluster together in the oblique rotation. Thus it appears that an oblique rotation using the Promax method provides the best model of the data. Similarly, an examination of the factor correlation matrix for the oblique rotation also indicates that an oblique rotation is most appropriate. Table 15 represents the factor correlation matrix for the Promax rotation.

**Table 15**

*EBS-Priority for Improvement*

*Factor Correlation Matrix (Promax)*

| Factor | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1.00 | .68 | .71 | .79 |
| 2 | .68 | 1.00 | .65 | .65 |
| 3 | .71 | .65 | 1.00 | .74 |
| 4 | .79 | .65 | .74 | 1.00 |

As indicated in table 15, many of the correlations between the factors in the Promax rotation exceed the recommended value of .32. An examination of the lower diagonal of the matrix indicates that all six of the inter-factor correlations exceed .32 and range from an absolute value of .65 - .79. This indicates substantial correlations between several of the factors and shared variance ranging from 42% to 62%. Consequently, the majority of the factors do not appear to be independent of each other and thus an oblique rotation appears more appropriate and consistent with the interrelated nature of the factors.

*School-wide Evaluation Tool (SET) Version 2.0*

One significant problem encountered in factor analyzing the SET data was the large number of missing data points. Table 16 below indicates the total number of valid and missing responses across all cases for each item and the percentage of the total of missing items.

Table 16
*Total Missing and Valid Responses*
Per Item

| Item | N | | |
|---|---|---|---|
| | Valid | Missing | % Missing |
| A1 | 284 | 0 | 0% |
| A2 | 246 | 38 | 13% |
| B1 | 284 | 0 | 0% |
| B2 | 284 | 0 | 0% |
| B3 | 239 | 45 | 16% |
| B4 | 284 | 0 | 0% |
| B5 | 284 | 0 | 0% |
| C1 | 284 | 0 | 0% |
| C2 | 284 | 0 | 0% |
| C3 | 284 | 0 | 0% |
| D1 | 284 | 0 | 0% |
| D2 | 284 | 0 | 0% |
| D3 | 240 | 44 | 15% |
| D4 | 236 | 48 | 17% |
| E1 | 284 | 0 | 0% |
| E2 | 284 | 0 | 0% |
| E3 | 246 | 38 | 13% |
| E4 | 246 | 38 | 13% |
| F1 | 246 | 38 | 13% |
| F2 | 284 | 0 | 0% |
| F3 | 246 | 38 | 13% |
| F4 | 284 | 0 | 0% |
| F5 | 284 | 0 | 0% |
| F6 | 246 | 38 | 13% |
| F7 | 246 | 38 | 13% |
| F8 | 284 | 0 | 0% |
| G1 | 284 | 0 | 0% |
| G2 | 240 | 44 | 15% |
| G3 | 105 | 179 | 63% |

As may be seen in Table 16, 12 out of 29 total items (41%) demonstrate a large number of missing responses across cases. The number of missing responses on individual items as a percentage of the total possible responses range from 13% to 63%. Both the total number of items with missing data and the number of missing responses for individual items are substantial. Depending on the cause of the missing data it may potentially affect the factor structure of the solution and thus the validity and generalizability of the results. When considering the potential for bias caused by missing data one must examine both the pattern of the missing data and the amount of missing data. If the pattern of missing data is found to be nonrandom it is highly likely that the missing data present a systematic bias to the data set and thus affect the generalizability of the factor solution (Tabachnick & Fidell, 1996, p. 60).

Although it is tempting to assume that missing data are randomly distributed it is recommended that the assumption of random distribution of missing data be tested (Tabachnick & Fidell, 1996, p. 60). In order to test this assumption, "dummy" variables were created for each item with missing data across any cases. The dummy variables were used to divide cases into those with missing data on the specific item and those without missing data on the specific item. A variable which summed the responses on all items without any missing data across all cases was computed and used to compare the groups. This method was used based on the assumption that the inclusion of multiple variables would provide a robust composite for comparison, and the assumption that items without missing data were less likely to be biased than those with missing data.

An independent samples t-test was used to compare the group means of those cases with and without missing data on the composite variable on each item. Twelve items had missing data across cases thus 12 independent samples t-tests were computed. The results of all 12 t-tests were highly significant ($p<.01$) indicating that the mean on the composite variable was significantly different on all items across the two groups. The group without missing data scored higher on the composite variable relative to the group with missing data in all 12 item comparisons. This indicates that the missing data in the sample are not randomly distributed. Non-randomly distributed missing data present a problem for the data sample in that this indicates the potential for a systematic bias, which can affect the outcome of factor analysis. Because of the nonrandom distribution of missing data and the potential this poses for biasing the sample analysis of the sample without addressing this problem is not appropriate.

There are two primary methods of addressing the problems posed by missing data and the bias it may affect. The first is to remove from the analysis all cases that include missing data. This approach removes those cases with missing data and thus may reduce the potential bias of such cases to the sample. However, it also poses a number of problems. The first is that it reduces the sample size, which may affect the results of data analysis with respect to the stability of the factor solution produced (Guadagnoli & Velicer, 1988). A second problem is that by removing cases with missing data this does not necessarily unbias the sample. It is very possible that those cases with missing data on a set of items differ significantly in their answers on other items. When such a relationship exists the removal of the cases with missing items may distort the mean item responses for the remaining sample and thus significantly affect the factor solution (Tabachnick & Fidell, 1996, p.60-63).

A second method used to address the problem of missing data is to replace the missing data using one of several methods designed to estimate the missing data. Recommend methods include replacing the missing data with the grand mean of the variable from individuals who have scores on it. A similar approach is to select a random individual's score on the variable and use this score to replace the missing item. One problem with this method is that although the variable grand mean is unaffected in using both methods the variance is typically reduced in the second method and almost always in the first thus both procedures reduce correlations with other variables (Gorsuch, 1983, p. 302-303; Tabachnick & Fidell, 1996, p.63).

One procedure often used to estimate and replace missing data is regression. In regression other variables from cases without missing data are used to estimate the missing values. Regression has several advantages over replacement with the grand mean. One is that regression is not as blind as simply inserting the grand mean due to the fact that regression accounts for the relationships among variables and thus achieves a better fit. However, this is also a disadvantage because it results in scores that fit together better than they would were the data not missing (Tabachnick & Fidell, 1996, p. 64). Gorsuch (1983, p. 303) reviewed research that compared the effectiveness of regression with deleting cases with missing data, and replacing the missing scores with the variable mean. The regression procedure was found to be the most effective method for estimating what the correlation matrix would have been if all scores had been available. The regression procedure was found to be the best whether 1% or 20% of the data were missing.

Given the large amount of missing data in the SET sample it is necessary to address this problem before analysis using one of the methods detailed above. The first solution explored was that of deleting cases with missing data. One concern noted above in deleting cases is the reduction in the sample size. All cases with missing data were removed from the sample and the sample size recomputed. The sample was reduced from an initial sample size of 284 cases down to 105. This constitutes a 37% reduction in sample size. Such a large reduction in sample size strongly suggests that a bias in the sample would result from removing all cases with missing data. There also remains the possibility that 105 cases do not constitute a large enough sample size for factor analysis.

Various authors have recommended minimal absolute and ratio sample sizes based on the ratio of cases to items. Comrey and Lee (1992) provided as a guide sample sizes of 50 as very poor, 100 as poor, 200 fair, 300 as good, 500 as very good, and 1000 as excellent. Floyd and Widaman (1995) recommend a ratio of 4 - 5 cases per item. Guadagnoli and Velicer (1988) compared ratio and absolute sample sizes finding that absolute sample size was more important than the ratio of cases to items in determining stable solutions. Small sample sizes may affect the factor solution by making it unstable. Subsequent re-analyzation of small sample sizes with the addition of data may result in variables switching from one factor to another. Deletion of all cases with missing items would result, according to Comrey and Lee (1992), in a "poor" sample size. Given that the SET contains 29 items deletion of all cases with missing data would result in a ratio of 3.48 cases per item, which according to Floyd and Widerman (1995) would not constitute an acceptable sample size. Thus deletion of all cases with missing data would result in a sample size of 101, which, based on both absolute sample size and ratio criteria would be considered an unacceptable sample size.

Due to the large number of missing items and the impact that deletion of all cases with missing data would have on the sample, both in terms of the sample size and potential bias introduced, missing items were estimated and imputed. Based on the above discussion a regression method known as linear interpolation was used to estimate and impute missing data points.

Summary statistics were computed to determine the appropriateness of factor analysis to the sample. Both the KMO statistic (.92) and the Bartlett's Test of Sphericity $\chi 2(406, N = 284) = 5,623.43$, $p < .01$ were found to be well within the acceptable ranges thus indicating that the sample was appropriate for factor analysis.

An initial extraction was made using the Maximum Likelihood procedure in order to determine the number of factors to retain for rotation. Using the default of eigenvalue greater than one, five factors were extracted. In order one through five the initial eigenvalues, and percent of total variance, were 11.92 (41.11%), 3.1 (10.63%), 1.92 (6.62%), 1.40 (4.83%), and 1.02 (3.53%), respectively. A substantial drop in the percentage of variance accounted for occurred between factors one and two and two and three. A gradual drop in the percentage of variance accounted for was observed between the third and fifth factors. Factors four and five both appeared to contribute meaningful variance. Thus an examination of the percentage of variance contributed by each of the factors indicates that a four or five factor solution is most appropriate. However, the gradual decline in variance accounted for makes it difficult to differentiate between a four and five factor solution. Similar results were found from an examination of the scree test. The scree test indicated a substantial break between factors one and two, a clear break between factors two and three, although less substantial, and a gradual leveling off of eigenvalues after factor three. The results of the scree test and the percentage of variance accounted for by each factor indicated that a four- or five-factor solution should be retained for rotation. However, these results did not allow for a clear differentiation between a four or five factor solution. We will now compare in greater depth the similarities and differences between a four and five factor solution in order to select one of these two for rotation.

In order to determine whether a four or five factor solution should be retained both models were extracted and compared. Each model was compared on a number of the facets of simple structure including the number of factor loadings greater than or equal to .30 and the absolute value of the mean of factor loadings greater than or equal to .30. This was based on the assumption that factors with a greater number of factor loadings ($\geq .30$) and a higher mean factor loading would demonstrate relatively greater simple structure and thus interpretability.

The four- and five-factor models were compared. In the four-factor model the fourth factor included two meaningful factor loadings defined as factor loadings greater than or equal to .30 with a mean factor loading of .35. In the five-factor model the fifth factor included two meaningful factor loadings with a mean of .40 and the fourth factor included two meaningful factor loadings with a mean factor loading of .36. Both the four- and five-factor models lack the minimum number of significant factor loadings (three) required to constitute a factor (Hatcher, 1994, pp.86, 92). However, the four-factor model lacks the adequate number of factor loadings on only the fourth factor while the five-factor model lacks the adequate number of factor loadings on both the fourth and fifth factors. The four- and five-factor models were further compared using the goodness-of-fit $\chi 2$ statistic. Lower $\chi 2$ statistics indicate a better fit of the factor model to the data. The four- and five-factor models produced goodness-of-fit statistics of $\chi 2(296, N =284) = 832.35$, $p = .00$ and $\chi 2(271, N =284) = 637.81$, $p < .01$, respectively. Although both models produced significant $\chi 2$ statistics, the five-factor model resulted in a 22% reduction in the $\chi 2$ statistic, relative to the four-factor model. This suggests that a five-factor model provides a better fit to the data than does a four-factor model. An examination of the residuals on both factor models also indicated a better fit to the data using a five-factor model. The number of absolute value residuals exceeding .05 in the five-factor model was 50 (12%)

compared to 83 (20%) in the four-factor model. This difference in residuals also suggests a better fit to the data of the five-factor model. An examination of the SMCs for each factor in both models did not clearly differentiate between a four- and five-factor model. In both models the first three factors exceeded the minimum recommended value of .70 (Tabachnick and Fidell, 1996), as did the fourth factor in the five-factor model. However, the fourth- and fifth-factor models' last coefficients of .67 and .67, respectively, were identical and just missed the minimum recommended value. Thus SMC coefficients did not differentiate between the two factor models.

In summary, an examination of the differences in residuals and goodness of fit statistics suggests that a five-factor model is more appropriate. However, a comparison of the number of meaningful factor loadings in the four- and five-factor models indicates that only the four-factor model is likely to produce factors with the minimum of three factor loadings needed to define a factor. Despite the fact that an examination of the residuals and goodness-of-fit statistics indicates that a five-factor model provides the best fit to the data it was determined that a four-factor model would be retained for rotation. This was based on the number of factor loadings defining the last factors in each model and the fact that the fifth factor in a five-factor model would not be adequately defined.

Following the determination that four factors should be retained, a four-factor model was extracted using the Maximum Likelihood procedure. Following the four-factor extraction the first thru fourth extraction sums of squared loadings and percentage of variance accounted for were 11.52 (39.74 %), 2.66 (9.19%), 1.37 (4.73%), and .83 (2.87%) of the total variance, respectively. All four factors accounted for a cumulative total of 56.53% of the variance. This solution was then rotated using both the Varimax and the Promax methods, resulting in both an orthogonal and oblique rotation. The goodness-of-fit statistic $\chi 2(296, N = 284) = 832.35, p < .01$ was significant. This suggests that more factors may be needed to appropriately model the data, however as previously shown a five-factor model is not appropriate. Table 17 below contains the factor loadings for the oblique rotation while table 18 contains the factor loadings for the orthogonal rotation.

Table 17
SET-Pattern Matrix
(Oblique Rotation-Promax)

| Survey Item | Factor 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| A1 | **.81** | | | |
| A2 | | **.52** | | |
| B1 | **.67** | | -.22 | |
| B2 | **.59** | | | .26 |
| B3 | | **.55** | .28 | |
| B4 | .32 | .28 | | **.44** |
| B5 | .40 | .24 | | **.49** |
| C1 | **.87** | | | |
| C2 | **.88** | | | |
| C3 | **.81** | | | |
| D1 | **.90** | | | |
| D2 | **.74** | | | |
| D3 | | .22 | | **-.40** |
| D4 | | .32 | | **.38** |
| E1 | **.75** | | | |
| E2 | **.88** | | | |
| E3 | | **.85** | | -.28 |
| E4 | | **.72** | | |
| F1 | | **.65** | | |
| F2 | **.64** | | .38 | |
| F3 | | | **.71** | |
| F4 | **.61** | | .28 | |
| F5 | **.70** | -.30 | .53 | |
| F6 | | .26 | **.76** | |
| F7 | | **.65** | .36 | |
| F8 | **.55** | .33 | | -.25 |
| G1 | **.60** | | .20 | |
| G2 | | | **.57** | |
| G3 | | | | **-.64** |

***Note.* Values less than .20 were suppressed to** aid in interpretation.

Table 18

SET-Factor Matrix
(Orthogonal Rotation-Varimax)

| Survey Items | Factor 1 | Factor 2 | Factor 3 | Factor 4 |
|---|---|---|---|---|
| A1 | **.77** | | | |
| A2 | .23 | **.50** | | |
| B1 | **.62** | .27 | | |
| B2 | **.66** | .32 | | .32 |
| B3 | .24 | **.60** | .37 | |
| B4 | **.53** | .43 | | .46 |
| B5 | **.61** | .42 | | .50 |
| C1 | **.78** | .25 | | |
| C2 | **.79** | | | |
| C3 | **.74** | | | |
| D1 | **.78** | | | |
| D2 | **.65** | | | |
| D3 | | | | **-.32** |
| D4 | | .37 | | .37 |
| E1 | **.74** | .24 | | |
| E2 | **.76** | | | |
| E3 | .20 | **.74** | .25 | |
| E4 | | **.66** | .28 | |
| F1 | .22 | **.61** | .23 | |
| F2 | **.70** | .21 | .46 | |
| F3 | | .31 | **.67** | |
| F4 | **.64** | .23 | .37 | |
| F5 | **.69** | | .55 | |
| F6 | | .38 | **.75** | |
| F7 | | **.62** | .47 | |
| F8 | **.55** | .39 | .22 | |
| G1 | **.62** | | .29 | |
| G2 | | .29 | **.55** | |
| G3 | | | | **-.56** |

***Note*. Values less than .20 were suppressed to**
aid in interpretation.

An examination of the factor loadings in table 18 indicates poor simple structure in the Orthogonal rotation. Eleven items demonstrate substantial factor loading coefficients (>=. 30) across multiple factors and three items substantially load on three separate factors. These double loadings make difficult the interpretation of individual factors as it is unclear which factors many of the items load on. The large number of double loadings indicate a lack of simple structure and make difficult interpretation of the factor structure. The fourth factor is also poorly defined with only one item uniquely loading this factor. The remainder of items, all loading more than one factor, demonstrate relatively low factor loading coefficients. The third factor is defined by the minimum three factor loadings but also demonstrates a substantial number of double loadings, which makes difficult its interpretation. It should be noted that factor loadings less than or equal to .20 were not reported to assist in interpretation.

In contrast, examination of the oblique rotation in table 17 indicates relatively greater simple structure. Six items demonstrate substantial factor loading coefficients

($\geq$ .30) across two separate factors and one item demonstrates substantial factor loadings across three separate factors. Although these seven items with double or triple loading factor coefficients negatively impact the simple structure of the solution and thus increase the difficulty of the interpretation the simple structure of an oblique rotation is relatively greater than that of an orthogonal rotation. An oblique rotation also increases the factor loading coefficients of a majority of the items, which produces greater simple structure. Thus an oblique rotation, although not producing an ideal factor structure solution, is relatively better than an orthogonal rotation. An examination of the factor correlation matrix for the oblique rotation also indicates that an oblique rotation is most appropriate. Table 19 represents the factor correlation matrix for the Promax rotation of the SET.

Table 19
*SET Factor Correlation Matrix (Promax)*

| Factor | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 1.00 | .59 | .67 | .26 |
| 2 | .58 | 1.00 | .64 | .45 |
| 3 | .67 | .64 | 1.00 | .41 |
| 4 | .26 | .45 | .41 | 1.00 |

As may be seen in table 19, all but one of the correlations between the factors in the Promax rotation exceed the recommended value of .32. An examination of the lower diagonal of the matrix indicates that five of the inter-factor correlations exceed .32 and range from .41-.67. This indicates substantial correlations between several of the factors and shared variance ranging from 17% to 45%. Consequently, the majority of the factors do not appear to be independent of each other and thus an oblique rotation is most appropriate and consistent with the interrelated nature of the factors. Having determined the factor structures of each of the measures I will now present the results of the correlational analyses between each of the derived factors.

### *Correlations among Measures and Scales*

Investigations of the relationships among the different measures were conducted, using the Pearson Product Moment Correlation in a correlation matrix. Such an investigation will enable a determination of the differences and similarities among the different measures and the degree to which the different measures do and do not overlap each in terms of variance accounted for. However, before launching into an examination of the relationships among measures it is important to first discuss and determine what constitutes a "meaningful" correlation.

Correlation coefficients are judged based on at least two criteria. The first is statistical significance and the second is what I will refer to as the meaningfulness of the observed correlation. Statistical significance is simply a measure of the probability that the observed correlation occurred by chance rather than by a true and systematic relationship between variables. Statistical significance is the product of three factors including the size of the observed relationship, the sample size, and the variability within the sample(s). This means that the statistical significance of a specific correlation such as .25 may be statistically significant in an analysis with a large sample size and low sample variability but not significant in an analysis involving a small sample size and a moderate level of sample variability. Thus consideration of statistical significance alone is not sufficient in determining the importance and meaning of an observed correlation. An observed correlation may or may not be statistically significant depending on the three factors noted above.

The "meaningfulness" of an observed correlation is more subjective than statistical significance but takes into account the real-world value of an observed relationship. For example, an observed correlation of .25 might be highly meaningful if it describes the relationship between the ingestion of a substance and the likelihood of death within a week, even if such a relationship is not found to be statistically significant. In contrast, if the observed correlation between two instruments intended to measure the same construct is found to be only .25 this would not constitute a meaningful relationship even thought such a correlation might be statistically significant. Thus the meaningfulness of an observed correlation is context dependent but also takes into account the somewhat subjective but important real-world value of the observed relationship.

Cohen (1988) discussed the use of the percentage of variance accounted for (r2) in determining the meaningfulness of an observed relationship. He emphasized the relative nature of determining meaningful correlations and notes that the question in determining whether an observed correlation is meaningful is "relative to what?", which may not be answered concretely. He proposed a somewhat subjective rubric for categorizing observed correlations in terms of "small", "medium", and "large" effect sizes. The effect size statistic quantifies the observed difference between groups or measures as a product of the pooled samples variance. This results in a raw score being converted to standard deviation units and thus makes comparison across groups with different standard deviations possible (Lipsey, 1990). Cohen (1988) proposed that correlations between .10 and .29 be considered "small" effects, correlations between .30 and .49 be considered "medium" effects, and correlations .50 or higher be considered "large".

Before interpreting the correlations between measures it is necessary to determine what constitutes a "meaningful" correlation in the context of the present study. The study is intended to examine the nature of the relationship among a number of instruments and predictor and outcome variables. Given this purpose it would appear more important to error on the side of not interpreting statistically significant but minimally meaningful relationships in favor of interpreting relationships and differences that are broadly considered meaningful. Cohen (1988) has proposed that correlations between .30 and .49 be considered of "medium" size. A correlation of .30 would indicate shared variance between measures of 9%. A shared variance of 9% would appear, at a minimum, to constitute a meaningful relationship between measures and also meet Cohen's definition of a medium effect size. In the forgoing examination of correlations between measures only those correlations that are equal to or great than .30 will be interpreted as meaningful even if values lower than this are found to be statistically significant.

*SET Correlated with EBS*

The SET and EBS were both designed to measure highly similar constructs. The EBS was developed first to measure the level of implementation of effective behavior support practices within a school setting using staff report. Following the development of the EBS the need was perceived to develop a research instrument that could be implemented by an outside observer. This resulted in the development of the SET. We will begin our examination of the relationships between measures by examining the relationship between the SET and EBS. Table 20 below contains the correlations between the SET and two halves of the EBS (Current Status and Priority for Improvement).

Table 20

*Correlations between EBS Current Status, EBS Priority for Improvement and the SET*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. EBS_CS_1 | 1.00 | | | | | | | | | | | |
| 2. EBS_CS_2 | .77** | 1.00 | | | | | | | | | | |
| 3. EBS_CS_3 | .79** | .71** | 1.00 | | | | | | | | | |
| 4. EBS_CS_4 | -.51** | -.43** | -.25 | 1.00 | | | | | | | | |
| 5. EBS_PI_1 | -.49** | -.37** | -.23 | .94** | 1.00 | | | | | | | |
| 6. EBS_PI_2 | -.55** | -.40** | -.39** | .84** | .87** | 1.00 | | | | | | |
| 7. EBS_PI_3 | -.54** | -.39** | -.39** | .85** | .85** | .85** | 1.00 | | | | | |
| 8. EBS_PI_4 | -.65** | -.52** | -.38** | .93** | .88** | .87** | .91** | 1.00 | | | | |
| 9. SET_1 | .47** | .34** | .39* | -.25 | -.25 | -.14 | -.19 | .26 | 1.00 | | | |
| 10. SET_2 | .28 | .16 | .26 | .06 | -.08 | .07 | .04 | -.01 | .84** | 1.00 | | |
| 11. SET_3 | .09 | .24 | .02 | .00 | .00 | .20 | .09 | .05 | .70** | .70** | 1.00 | |
| 12. SET_4 | .42** | .33** | .44** | -.26 | -.40* | -.25 | -.30 | -40* | .72** | .62** | .50** | 1.00 |

***Note*. "CS" indicates Current Status; "PI" indicates Priority for Improvement**

** p≤ .01 (2-tailed) *p≤.05 (2-tailed)

An examination of the correlations between the EBS Priority for Improvement (EBS-PI) and EBS Current Status (EBS-CS) indicates statistically significant and meaningful negative correlations between each of the scales with one exception. The significant correlations range from -.38 to -.94. The strongest correlations appear to be between the EBS-CS factor 4 and all of the EBS-PI factors. The strongest correlation is between the EBS-CS factor 4 and EBS-PI factor 4. The weakest correlation is between the EBS-PI factor 1 and EBS-CS factor 3. Although the majority of EBS Current Status and Priority for Improvement factors are significantly correlated they are not so highly correlated as to suggest that both are measuring the same constructs, with the exception of EBS-CS factor 4 and all of the EBS-PI factors. The pattern of negative correlations indicates an inverse relationship between ratings on the EBS-PI and the EBS-CS. This indicates that when school staff rate the current status of implementation of positive behavior supports in their school as being high they rate the priority for improvement of positive behavior supports as low. Conversely, when staff rate the implementation of positive behavior supports in their school as being low they rate the priority for improvement of positive behavior supports as high. This relationship makes intuitive and practical sense.

The pattern of correlations between the SET and EBS-CS and EBS-PI presents a different picture. Both the SET factor 1 and SET factor 4 demonstrate significant and meaningful positive correlations with the first three factors of the EBS-CS. However, the magnitude of these correlations, which range from .39 - .47, are not large enough to indicate that the SET and EBS-CS are measuring the same construct but rather overlap each other between 15 and 22 percent. SET factors one and two do not significantly correlate with any of the EBS-CS which further suggests that the SET and EBS-CS are not measuring the same construct and overlap only marginally in the constructs they measure.

The EBS-PI factors demonstrate very little correlation with the SET. Only SET factor 1 correlates significantly and meaningfully with EBS-PI factors two and four. Both of these correlations are essentially identical at -.40. In contrast to the EBS-CS, which demonstrates at least some overlap across a number of factors, the EBS-PI essentially demonstrates little meaningful correlation with the SET. Only two EBS-PI factors correlate with the fourth SET factor. The magnitude of these correlations indicates a shared variance of 16% and suggests that the SET and EBS-PI measuring substantially different constructs. Taken together these results indicate that EBS and SET do not overlap to a substantial degree and thus do not measure the same constructs. This is noteworthy in that the SET was developed as an analogue measure to the EBS.

*SET Correlated with OSSS*

Next examined was the relationship between the SET and OSSS. Table 21 below contains the correlations between the four SET factors and the OSSS Risk and Protect factors.

Table 21

*Correlations between SET and OSSS*

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1. SET 1 | 1.00 | | | | | |
| 2. SET 2 | .84** | 1.00 | | | | |
| 3. SET 3 | .70** | .70** | 1.00 | | | |
| 4. SET 4 | .72** | .62** | .50** | 1.00 | | |
| 5. OSSS Risk | .11 | .15 | .05 | -.08 | 1.00 | |
| 6. OSSS Prot. | .10 | .11 | .18 | .19 | -.42 | 1.00 |

**** p≤ .01 (2-tailed)**

As may be seen in table 21, there are no statistically significant correlations between any of the four SET factors and the OSSS Risk and Protect factors. This indicates that the SET and OSSS are not significantly or meaningfully related across any of their factors and thus do not measure similar constructs.

In contrast, the EBS-CS and EBS-PI both demonstrated significant and meaningful correlations with the OSSS Risk and Protect factors. The pattern of these correlations is also interesting to note. Table 22 below contains the correlations between each of the EBS-CS, EBS-PI, and OSSS factors.

Table 22
*Correlations between EBS Current Status, EBS Priority for Improvement and the OSSS*

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. EBS_CS_1 | 1.00 | | | | | | | | | |
| 2. EBS_CS_2 | .77** | 1.00 | | | | | | | | |
| 3. EBS_CS_3 | .79** | .71** | 1.00 | | | | | | | |
| 4. EBS_CS_4 | -.51** | -.43** | -.25 | 1.00 | | | | | | |
| 5. EBS_PI_1 | -.49** | -.37** | -.23 | .94** | 1.00 | | | | | |
| 6. EBS_PI_2 | -.55** | -.40** | -.39** | .84** | .87** | 1.00 | | | | |
| 7. EBS_PI_3 | -.54** | -.39** | -.39** | .85** | .85** | .85** | 1.00 | | | |
| 8. EBS_PI_4 | -.65** | -.52** | -.38** | .93** | .91** | .88** | .87** | 1.00 | | |
| 9. OSSS Risk | -.22 | -.12 | -.07 | .50** | .47** | .61** | .58** | .51** | 1.00 | |
| 10. OSSS Prot. | .67** | .66** | .74** | -.28 | -.20 | -.43* | -46** | -.45** | -.42** | 1.00 |

***Note*. “CS” indicates Current Status; “PI” indicates Priority for Improvement**

** p≤ .01 (2-tailed)   *p≤.05 (2-tailed)

The EBS-CS factors indicate school staff members’ rating of the presence of positive behavior supports in their school. In contrast the EBS-PI factors indicate staff members’ priority for improving positive behavior supports in their school. An examination of the correlations in table 22 between the EBS-CS and EBS-PI factors indicates that the majority of correlations are significant, meaningful, and negative. This indicates that when staff rate the level of implementation of positive behavior supports as being high, they rate the priority for improving positive behavior supports in their school as low. Conversely, when staff rate the level of implementation of positive behavior supports as being low, they rate the priority for improving positive behavior supports in their schools as high. This pattern makes sense given that when positive behavior supports are present with a high degree of fidelity there is little need for improvement and thus staff rate this as a low priority for improvement. In contrast, when positive behavior supports are not present or are not implemented with fidelity there is a need for improvement and thus staff would rate the need for improvement higher.

An interesting pattern of correlations emerges in table 22 between the OSSS Risk and Protect factors and the EBS-CS and EBS-PI factors. With the exception of the fourth EBS-CS factor, no significant correlations exist between EBS-CS and OSSS Risk. This indicates that the presence or absence of positive behavior supports, as assessed by the EBS-CS, is not related to the rated risk factors for a school. In contrast, all of the EBS-CS factors, with the exception of factor four, are positively significantly and meaningfully related to the OSSS Protect factor. This would suggest that the presence of positive behavior supports within a school are related to protective factors in the school and that this relationship is positive. This means that as the staff rated level of implementation of positive behavior supports within a school increases so does the rating of protective factors. Conversely, as the staff rated level of implementation of positive behavior supports decreases so does the level of protective factors.

All four of the EBS-PI factors are significantly and meaningfully positively correlated with the OSSS Risk factor. Similarly, with the exception of the first EBS-IP factor, all of the EBS-PI factors are significantly and meaningfully negatively correlated with the OSSS Protect factor. Taken together these results indicate that as the priority for improving positive behavior supports in a school (EBS-PI) increases so does the risk while protective factors decrease. Thus in relation to priority for improvement of positive behavior supports the OSSS risk and protect factors are positively and negatively related, respectively, both to each other and to the priority for improvement of positive behavior supports. Only in relation to priority for improvement do the OSSS risk and protect factors appear to be inversely related to each other. In the staff rating of the current level of implementation of positive behavior supports in their school the OSSS risk and protect factors do not appear to be inversely related and only the OSSS protect is significantly related to the current level of implementation of positive behavior supports in a school. The final section includes the results of the multiple regression analyses designed to determine the optimal combination of OSCSS subscales in predicting each of the three school safety variables.

### *Multiple Regression of OSCSS*

To address the fourth research question and investigate the relationships between student reported beliefs, behaviors, and perceptions and the School Safety construct multiple regression was conducted on OSCSS data. The School Safety construct was defined using three OSCSS scales: Rebellious Behavior, Drug Availability, and Student Victimization. As previously noted the OSCSS was administered to students in two forms, form A and form B with each student completing only one of the two forms. Form A included nine scales: Respect for Students, Fairness of Rules, Student Influence, Planning and Action, Drug Availability, Rebellious Behavior, Attachment to School, Belief in Rules, and School Rewards. Form B included six scales: Clarity of Rules, Student Victimization, Positive Peer Association, School Effort, Social Integration, and Student Safety. Because the OSCSS scales were divided between two student surveys and students completed only one survey each two distinct nonoverlapping samples were analyzed, one for each form. Because there were two different forms and samples it was only possible to regress scales that shared the same form. For example, the scales included on Form A were regressed onto the School Safety scales of Rebellious Behavior but not onto the Student Victimization and Drug Availability scales because these were included on Form B. Similarly, only the scales on Form B were regressed on Student Victimization and Drug Availability.

#### *Search for and Removal of Multivariate Outliers*

Three multiple regression equations were computed for each of the three School Safety scales using two samples, one for each of the two student surveys. Before analysis each of the two samples were examined for multivariate outliers. Multivariate outliers are cases with extreme values on one or a combination of variables that distort statistics. Multivariate outliers often lead to both Type I and Type II errors and frequently lead to results that do not generalize due to their being overly determined by outliers (Tabachnick & Fidell, 1996, pp.66). Consequently, multivariate outliers pose significant problems in multiple regression analyses and should be eliminated or transformed. In order to identify multivariate outliers Mahalanobis distances statistics were run for each of the cases in both samples. Mahalanobis distance is the distance of a case from a centroid of the remaining cases where the centroid is the point created by the means of all variables and is distributed as a chi square variable (Tabachnick & Fidell, 1996, pp.178, 67). Table 23 below indicates the original sample size, the number of cases considered to be outliers using a conservative $p \leq .01$, and the number of cases following removal of all outliers using pair wise deletion for each OSCSS scale.

Table 23
*Forms A and B Sample Size, Number of Outliers, and Sample Size with Removal*

| Scale | Sample Size | Outliers | Sample with Removal of Outliers |
|---|---|---|---|
| Respect for Students’ | 327 | 68 | 259 |
| Fairness of Rules’ | 325 | 66 | 259 |
| Clarity of Rules* | 651 | 73 | 578 |
| Student Influence’ | 318 | 59 | 259 |
| Planning and Action’ | 315 | 56 | 259 |
| Drug Availability* | 657 | 81 | 576 |
| Rebellious Behavior’ | 324 | 70 | 254 |
| Student Victimization* | 663 | 100 | 563 |
| Positive Peer Association* | 641 | 63 | 578 |
| Attachment to School’ | 297 | 38 | 259 |
| Belief in Rules’ | 322 | 63 | 259 |
| School Effort* | 658 | 80 | 578 |
| Social Integration* | 647 | 69 | 578 |
| School Rewards’ | 318 | 59 | 259 |
| Student Safety* | 638 | 60 | 578 |

**’Form A Mahalanobis significance threshold $\chi2$ (7, N = 327- 297) = 18.48, $p \leq .01$**

*Form B Mahalanobis significance threshold $\chi2$ (6, N = 657 - 638 ) = 16.81, $p \leq .01$

As may be seen in table 23, the number of cases determined to be outliers across the different scales ranges from 38 – 100, which results in a reduction in the total number of cases available for regression to between 254 and 578. Tabachnick and Fidell (1996, pp.132) provide as rules of thumb that one have $N \geq 50 + 8m$ (m is the number of IVs) cases for testing multiple correlations and $N \geq 104 + m$ cases when testing individual predictors. For regression equations using Form B data and five independent variables this would mean a minimum sample size of 90 – 109, and for regression equations using Form A data and six independent variables this would mean a minimum sample size of 98 – 110 cases, depending on which formula is used. In both instances removal of all cases determined to be outliers would result in sample sizes in excess of the minimum number of cases required for multiple regression. Due to the potential for increased Type I and II error rates and reduced generalizability of the results, and the fact that removal of all outliers does not violate the recommended case to independent variable ratio all outliers were removed from both samples for further analysis.

*Predicting Student Victimization*

Standard multiple regression procedures were used to predict the OSCSS Student Victimization (SV) scale using other OSCSS scales. The SV scale was entered into the regression equation as the dependent variable with the remaining OSCSS form B scales entered in simultaneously as independent variables. Table 24 below summarizes the results.

Table 24
*Model 1: Simultaneous Multiple Regression Predicting Student Victimization (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Clarity of Rules | .57 | .16 | .00 | -.91 | .05 | .09 |
| Positive Peer Assoc. | -.28 | -.17 | .00 | -.94 | -.05 | .09 |
| School Effort | -.06 | -.02 | .37 | -.85 | -.01 | .20 |
| Soc. Integration | -.01 | -.00 | .94 | -.93 | -.01 | .08 |
| Student Safety | -.70 | -.94 | .00 | -.97 | -.18 | .04 |

*** Unstandardized**

The multiple regression model resulted in an R = .97 and significant ANOVA F (5, 570) = 1992.86, p < .01 indicating that the model was able to account for 94% of the variance in the DV and that these results cannot be accounted for by chance. Examination of the results in table 24 indicates that, with the exception of Social Integration and School Effort, all independent variables significantly predict SV. However, further inspection indicates potential multicollinearity. Multicollinearity occurs when multiple independent variables are highly intercorrelated with each other. Not only does multicollinearity violate one of the assumptions of multiple regression but also destabilizes the solution and results in significant fluctuation in the inverted matrix with only minor changes in correlations. The consequence is reduced generalizability and instability of the beta coefficients (Tabachnick & Fidell, 1996, p.86). Tolerance provides a measure of multicollinearity. Tolerance represents 1 - R2 for the regression of that independent variable on all the other independents, ignoring the dependent. The higher the intercorrelation of the independents, the more the tolerance will approach zero. As a rule of thumb, if tolerance is less than .20, a problem with multicollinearity is indicated. With the possible exception of School Effort all IVs entered into the above regression equation demonstrate tolerance values substantially less than .20, which indicates multicollinearity (Tabachnick & Fidell, 1996). Multicollinearity is further substantiated by a review of the correlation coefficients. A zero-order correlation indicates the size and direction of the correlation between two variables, in this case the specific IV and the DV, regardless of each IV's correlation with other IV's. Semiparital correlations represent the correlation between the DV and an IV when the linear effects of the other independent variables in the model have been removed from the independent variable. Thus semipartial correlation represents the unique correlation between an IV and DV that is not shared by any other IV-DV combination (Tabachnick & Fidell, 1996). When zero-order correlations are high and semipartial correlations are low, as in the present case, this indicates that all IVs are highly correlated with the DV and highly correlated with each other given the low unique contribution of each IV. In other words there is a high degree of overlap among the IVs in their relationship with the DV.

Due to the significant problems with multicollinearity in the first model a second model was computed using standard multiple regression with SV entered as the DV. The second model was computed in an attempt to reduce multicollinearity and eliminate unnecessary variables. All IVs were entered with the exception of Social Integration, which did not demonstrate statistical significance in the first model and did demonstrate a low tolerance statistic. Although School Effort did not predict the DV at a statistically significant level it was retained in the second regression model due to the fact that it was the only IV with an acceptable tolerance statistic. Table 25 below summarizes the results of the second model.

**Table 25**

*Model 2: Simultaneous Multiple Regression Predicting Student Victimization (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Clarity of Rules | -1.26 | -.35 | .00 | -.91 | -.15 | .19 |
| Positive Peer Assoc. | -.22 | -.07 | .01 | -.85 | -.03 | .21 |
| School Effort | -.92 | -.56 | .00 | -.94 | -.20 | .13 |

*** Unstandardized**

The second multiple regression model resulted in an R = .95 and significant ANOVA $F(3, 572) = 1729.15$, $p < .01$ indicating that the model was able to account for 90% of the variance in the DV and that these results cannot be accounted for by chance. Examination of the results in table 25 indicates that multicollinearity continues to be present in the second model although to a lesser degree. Only one tolerance statistic is substantially below .20. It is interesting to note that two tolerance statistics increased and one (School Effort), which was previously at an acceptable level, decreased. The second model accounted for approximately 4% less of the variance in the DV. All IVs in the second model significantly contributed to the prediction of variance in the DV.

Consideration was given to running a third model by dropping the IV School Effort in an attempt to reduce multicollinearity. However, this was decided against based on the relatively high semipartial correlation (-.20) between School Effort and the DV. This indicates that School Effort makes a substantial unique contribution to the prediction of the DV and that its elimination would likely unnecessarily reduce the percentage of variance accounted for in the DV. The second regression model appears to provide the best fit to the data in that it predicts a substantial portion of the variance in the DV with an acceptable level of multicollinearity and does so more efficiently, with fewer IVs than does the first regression model.

*Predicting Student Rebellious Behavior*

Standard multiple regression was used to predict the OSCSS Student Rebellious Behavior (RB) scale using other OSCSS scales. The RB scale was entered into the regression equation as the dependent variable with the remaining OSCSS Form A scales entered simultaneously as independent variables. Table 26 below summarizes the results.

Table 26

*Model 1: Simultaneous Multiple Regression Predicting Student Rebellious Behavior (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Respect for Students | -.42 | -.13 | .04 | -.46 | -.10 | .56 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Fairness of Rules | -.49 | -.09 | .13 | -.45 | -.07 | .58 |
| Student Influence | .16 | .04 | .41 | -.09 | .04 | .84 |
| Planning and Action | .68 | .13 | .02 | -.19 | .11 | .68 |
| Attachment to School | -.42 | -21 | .00 | -.53 | -.14 | .44 |
| Belief in the Rules | -1.38 | -.45 | .00 | -.62 | -.38 | .71 |
| School Rewards | -.39 | -.09 | .08 | -.30 | -.08 | .78 |

The multiple regression model resulted in an R = .70 and significant ANOVA F (7, 246) = 33.32, p < .01 indicating that the model was able to account for 49% of the variance in the DV and that these results cannot be accounted for by chance. Examination of the results in table 26 indicates that, with the exception of Fairness of Rules, Student Influence, and School Rewards all independent variables significantly predicted Student Victimization. Multicollinearity is not a concern in this model as all Tolerance values exceed .20 and most are well above this.

Although the seven independent variables included in the above regression equation accounted for a substantial portion of variance in RB and do not demonstrate multicollinearity not all of the IVs were statistically significant. In attempt to improve the efficiency of the regression model a second regression model was run eliminating those IVs that were not statistically significant in the first model. Standard multiple regression was again used with RB entered in as the IV and all previous DVs, with the exception of those noted above. Table 27 summarizes the results of the second model.

Table 27
*Model 2: Simultaneous Multiple Regression Predicting Student Rebellious Behavior (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Respect for Students | -.46 | -.14 | .02 | -.46 | -.11 | .57 |
| Planning and Action | .61 | .12 | .02 | -.19 | .11 | .80 |
| Attachment to School | -.55 | -.27 | .00 | -.53 | -.20 | .54 |
| Belief in the Rules | -.1.43 | -.47 | .00 | -.62 | -.41 | .75 |

The second multiple regression model resulted in an R = .69 and significant ANOVA F (4, 249) = 56.18, p < .01 indicating that the model was able to account for 48% of the variance in the DV and that these results cannot be accounted for by chance. As with the first model all Tolerance statistics were well above the recommended minimum value of .20 and all of them increased slightly. This finding indicates that multicollinearity is not a concern with the second model. It is significant to note that the removal of three IVs only resulted in a 1% drop in the percentage of variance accounted for in SB and slightly increased all Tolerance values. The second model with only four IVs is able to predict SB as well as the first model with seven IVs. This finding indicates that the second model is a more efficient and parsimonious model for predicting Student Rebellious Behavior.

*Predicting Drug Availability*

Standard multiple regression was used to predict the OSCSS Drug Availability (DA) scale using other OSCSS scales. The DA scale was entered into the regression equation as the dependent variable with the remaining OSCSS Form B scales entered simultaneously as independent variables. Table 28 below summarizes the results.

**Table 28**

*Model 1: Simultaneous Multiple Regression Predicting Drug Availability (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Clarity of Rules | -1.14 | -.42 | .00 | -.77 | -.13 | .09 |
| Positive Peer Assoc. | -.22 | -.18 | .04 | -.77 | -.05 | .09 |
| School Effort | -.73 | -.31 | .00 | -.76 | -.14 | .21 |
| Social Integration | -.18 | -.10 | .25 | -.76 | -.03 | .08 |
| Student Safety | .08 | .15 | .25 | -.77 | .03 | .04 |

The multiple regression model resulted in an R = .81 and significant ANOVA F (5, 570) = 213.74, p < .01 indicating that the model was able to account for 66% of the variance in the DV and that these results cannot be accounted for by chance. Examination of the results in table 28 indicates that, with the exception of Social Integration and Student Safety all independent variables significantly predicted DA. With the exception of one IV all other IVs demonstrate Tolerance statistics below .20 and thus indicate a problem with multicollinearity.

In an attempt to reduce the problem of multicollinearity evidenced in the first model and improve the efficiency of the model be eliminating IVs that did not significantly predict the DV a second model was computed using standard multiple regression with Drug Availability entered as the DV. All IVs were entered with the exception of Social Integration and Student Safety both of which did not demonstrate statistical significance in the first model. Table 29 below summarizes the results of the second model.

**Table 29**

*Model 2: Simultaneous Multiple Regression Predicting Drug Availability (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Clarity of Rules | -1.04 | -.38 | .00 | -.77 | -.17 | .19 |
| Positive Peer Assoc. | -.19 | -.16 | .03 | -.77 | -.06 | .126 |
| School Effort | -.75 | -.31 | .00 | -.76 | -.15 | .22 |

The multiple regression model resulted in an R = .81 and significant ANOVA F (3, 572) = 355.82, p < .01 indicating that the model was able to account for 66% of the variance in the DV and that these results cannot be accounted for by chance. All IVs included in the second model significantly predicted DA. Although two out of three IVs' Tolerance statistics fall below .20, indicating multicollinearity, the Tolerance statistics for each of the included IVs increased relative to their values in the first model. This indicates that multicollinearity continues to pose a problem in the second model but this problem has been decreased with the elimination of several IVs. It is significant to note that the percentage of variance accounted for in the DV with the removal of two IVs did not change indicating that the second model was able to predict the DV as well as the first with fewer IVs.

Because the second model continued to demonstrate multicollinearity but increased the tolerance statistics with the deletion of two IVs, a third model was computed to see if multicollinearity might be further reduced without a substantial reduction in the percentage of explained variance. The third model was conducted using standard multiple regression with Drug Availability entered as the DV and two variables, Clarity of Rules and School Effort, as the DVs. Positive Peer Association was eliminated from the model due to it having the lowest Tolerance statistic of the three. Table 30 below summarizes the results of the third model.

**Table 30**

*Model 3: Simultaneous Multiple Regression Predicting Drug Availability (DV)*

| Variable | Beta | | *p* | Correlations | | Tolerance |
|---|---|---|---|---|---|---|
| | B* | β | | r | sr | |
| Clarity of Rules | -1.27 | -.46 | .00 | -.77 | -.27 | .34 |
| School Effort | -.92 | -.39 | .00 | -.76 | -.22 | .34 |

The third multiple regression model resulted in an R = .81 and significant ANOVA F (2, 573) = 527.54, p < .01 indicating that the model was able to account for 66% of the variance in the DV and that these results cannot be accounted for by chance. Examination of the results in table 30 indicates that elimination of the third IV (Positive Peer Association) increased all tolerance statistics to .34 thus reducing multicollinearity to acceptable levels while predicting the same percentage of variance in the DV. The third regression model is an improvement over the second in that it accounts for the same percentage of variance in the DV with fewer IVs and while reducing the problem of multicollinearity to within acceptable levels. Based on these results the third regression model, which includes Clarity of Rules and School Effort, appears to provide the best linear prediction of variance in Drug Availability. The first regression model predicts a greater percentage of variance in the DV but does so at the expense of substantial multicollinearity and five IVs. Due to the problem of multicollinearity in the first model the beta coefficients are likely unstable, the model's prediction over inflated, and the results poor in generalizabiltiy. Thus the third regression model with two IVs provides the most stability and generalizability of results and does so with the greatest parsimony of IVs.

# Discussion

This study, which was primarily exploratory in nature, was designed to answer several questions related to the measurement of school safety and disorder. The questions as originally proposed included: (1) What are the factor structures and internal consistency coefficients of the SET, EBS 1.5, OSSS, and OSCSS? (2) What are the zero-order correlation coefficients between the SET, EBS 1.5, OSSS, and OSCSS at the full-scale and subscale levels? (3) What are the zero-order correlation coefficients between the community and school archival data, SET, EBS 1.5, OSSS, and OSCSS full-scale and subscale scores and the school safety measure? (4) What combination of community and school archival variables, measures, and subscales best predicts the construct of school safety/disorder? The first two questions were analyzed as proposed however, due to limitations encountered in the data set as previously noted, the third question was not possible to answer and was eliminated, whereas the fourth question was modified to include only data from the OSCSS. Due to the complexity and number of analyses undertaken I will begin with a brief review of the results of the study followed by a more in-depth interpretation.

## *Review of Results*

### Internal Consistency of Measures

Alpha coefficients were computed on each of the scales for the OSSS, EBS 1.5 (Current Status and Priority for Improvement), SET, and OSCSS. The OSSS included two scales, Risk and Protect, both of which demonstrated excellent and essentially identical alpha coefficients.

The EBS 1.5 included four scales designed to assess the Current Status and Priority for Improvement of selected positive behavior support systems in a school. Current Status alpha coefficients ranged from .82 - .95; Priority for Improvement alpha coefficients ranged from .96 - .91. Both Priority for Improvement and Current Status scales demonstrated highly similar and acceptable alpha coefficients thus indicating that the EBS 1.5 scales have strong internal consistency both in the rating of Current Status and Priority for Improvement.

The SET includes seven scales. Alpha coefficients across the seven scales ranged from excellent to unacceptably low (.91 - .29). Four scales demonstrated acceptable alpha coefficients indicating good internal consistency. Three scales demonstrated poor to unacceptable alpha coefficients (.56 - .29). This finding indicated that these three scales have poor internal consistency. Item level analysis of the SET scales demonstrating poor internal consistency indicated several possible reasons for these low scores including low number of items in the scale and poor item-total correlations.

The OSCSS is comprised of two forms with a total of fifteen scales. Ten of the scales demonstrated acceptable alpha coefficients indicating strong internal consistency. Three scales evidenced questionable alpha coefficients (.63 - .58) thus indicating less than desirable internal consistency. Finally two scales demonstrated poor to unacceptably low alpha coefficients indicating poor internal consistency of these scales.

## Factor Structure of Measures

The SET, EBS 1.5 and the OSSS were factor analyzed in order to establish appropriate factor structures. The OSSS data indicated a clear two-factor orthogonal solution accounting for approximately 40% of the variance.

The EBS 1.5, as previously mentioned, asks respondents to rate each item in terms of Current Status and Priority for Improvement. As a result, two separate factor analyses were conducted on the EBS 1.5, one for each response category. The EBS-Current Status (EBS-CS) factor analyzed into a four-factor oblique solution accounting for approximately 49% of the variance. The EBS-CS demonstrated a relatively clear factor solution.

The EBS-Priority for Improvement (EBS-PI) also factor analyzed into a four-factor oblique solution accounting for approximately 57% percent of the variance. The EBS-PI demonstrated a relatively clear factor solution. Although the EBS-CS and EBS-PI both demonstrated oblique four-factor solutions the factors did not exactly overlap in terms of the item content of the factors.

The SET factor analyzed into a four-factor oblique solution accounting for approximately 56% of the variance. However, the factor analysis of the SET did not demonstrate a clear factor structure with multiple items loading substantially on more than one factor.

## Intercorrelations among Factors and Measures

The correlations between the above derived factors of the EBS 1.5, SET, and OSSS were next examined. Within the EBS 1.5 the majority of EBS-CS and EBS-PI factors were significantly and negatively correlated. Only the first and fourth factors of the SET were significantly positively correlated with the EBS 1.5 and then only with the EBS-CS and not with EBS-PI with the exception of the SET fourth factor and EBS-PI first factor. The SET data demonstrated no significant correlations with the OSSS. In contrast the first three factors of EBS-CS demonstrated significant positive correlations and the first three factors of the EBS-PI evidenced a significant negative correlations with the OSSS Protect factor. Factors one through four of the EBS-PI were significantly and positively correlated with the OSSS Risk factor.

## Multiple Regression: Predicting School Safety/Disorder

Several scales from the OSCSS Form A and Form B were included in standard multiple regression equations predicting the school safety/disorder OSCSS scales of Student Victimization, Rebellious Behavior, and Drug Availability. The scales Clarity of Rules, Positive Peer Association, and School Effort all significantly and negatively predicted Student Victimization. The model was able to account for approximately 90% of variance in Student Victimization.

The scales Respect for Students, Planning and Action, Attachment to School, and Belief in the Rules significantly negatively predicted Student Rebellious Behavior. The model was able to account for approximately 48% of the variance in Student Rebellious Behavior.

The scales Clarity of Rules and School Effort significantly negatively predicted Drug Availability. The model was able to account for approximately 66% of the variance in Drug Availability.

***Discussion of Factor Analytic Results***

Having briefly reviewed the results, these findings will be discussed and interpreted beginning with the factor analytic results. The OSSS was designed to assess a set of known risk and protective factors in the school and neighborhood. Items 1 - 17 assess risk factors while items 18 - 33 assess protective factors. Exploratory factor analysis confirmed both the two-factor structure and the item content within the two factors. All items intended to assess risk factors loaded on the first factor with all items demonstrated strong factor-loading coefficients. With the exception of item 26, all items intended to assess risk factors loaded on the second factor, also with strong factor loading coefficients. Item 26 did not demonstrate a meaningful factor-loading coefficient on either factor. Item 26 assessed the extent to which "Suicide prevention/response plans" exist in the school and neighborhood. This finding suggests that suicide prevention/response plans may not serve as a risk or protective factor. However, the fact that item 26 did not load substantially on either factor may also be an artifact of the samples used or of a relatively low base-rate of suicidal behavior in the schools examined. Further research should be conducted to confirm or disconfirm this finding. It is significant to note that the OSSS factor solution is orthogonal, meaning the subscales are considered to be independent. This fact indicates that the constructs of risk and protect are relatively independent of each other and thus it is possible to encounter schools and communities with either high risk and high protective factors or low risk and low protective factors or any combination in-between.

Exploratory factor analysis of the EBS 1.5 indicated a four-factor structure for both Current Status and Priority for Improvement response sets. This finding is important given that the EBS 1.5 was designed to include four scales. Consequently the results of the factor analysis appear to confirm a four-factor structure for the EBS 1.5. However, the items loading within each factor and the items in the original scales proposed by the instrument's authors do not exactly overlap and differ between the Current Status and Priority for Improvement factor solutions. The original scales of the EBS with their corresponding item numbers are: School Wide Systems (101 - 115), Nonclassroom Setting System (201 - 209), Classroom Systems (301 – 311), and Individual Student Systems (401 - 408).

The first factor of the EBS-CS includes items 101 - 209 and thus encompasses all items in the original scales School Wide Systems and Nonclassroom Setting Systems. This suggests that nonclassroom and school wide positive behavior support systems may not be two distinct systems but rather are one large system of positive behavior support.

The second factor of the EBS-CS conforms perfectly in item content to the Classroom Systems scale. The third factor of the EBS-CS also conforms perfectly in item content to the Individual

Student Systems scale. These results suggest that Classroom Systems and Individual Student Systems constitute two distinct systems of positive behavior support and serves to confirm the structure of the EBS 1.5 proposed by its authors, at least with regard to these two factors.

The fourth EBS-CS factor only includes three items all three of which are from the School-Wide Systems scale. The three items are: (116) “All staff are involved directly and/or indirectly in school-wide interventions”, (117) “The school team has access to on-going training and support from district personnel”, and (118) “The school is required by the district to report on the social climate, discipline level or student behavior at least.” All three items demonstrated substantial factor loading coefficients suggesting strong internal consistency. However, the fourth EBS-CS factor is poorly defined with only three items. It is also unclear how conceptually these three items differ form the others and are similar to each other within the original scale. Two of the items address staff involvement and school team access to continued training. The third item assesses district requirements regarding school reporting. Further research may be needed to confirm the stability of this factor and to examine the nature of this factor and how it differs from the other factors.

The EBS-PI factor structure differed slightly from the EBS-CS factor structure. With the exception of two items, the first EBS-PI factor conformed in item content to the School Wide Systems scale. Similarly, the EBS-PI second factor conformed perfectly in item content with the Classroom Systems scale. The third EBS-PI factor conformed perfectly in content with the Individual Student Systems scale. Finally, the fourth EBS-PI factor, with the exception of one item that loaded substantially on two factors, conformed in item content with the Nonclassroom Setting Systems scale.

It is important to note that the EBS-CS factor structure included the items for the original scales School-Wide Systems and Nonclassroom Setting Systems in a single factor while the EBS-PI factor structure differentiated these items into two distinct factors that conformed to a high degree with the original scales of the instrument. It is unclear what this difference suggests. One possibility is that school-wide and nonclassroom systems constitute a single factor when their current status is considered but when considering the less tangible priority for improvement they constitute two distinct factors. Further research is needed to confirm the stability of these factor structures and to explore possible reasons for the observed difference. The results of the factor analysis of the EBS-CS and EBS-PI largely confirm the four-factor structure and item content of the factors of the EBS 1.5 however with a few exceptions. It should also be noted that the factor analytic results clearly indicate the factors of the EBS 1.5 are highly interrelated with each other. Thus although school wide positive behavior support as assessed by the EBS 1.5 does appear to consist of separate constructs these constructs are highly interrelated.

Exploratory factor analysis of the SET produced an oblique four-factor structure. However, the factor structure was plagued by a large number of items that demonstrated substantial factor loadings across multiple factors, which makes interpretation of the factor structure difficult. It is also important to note that the data set used included large numbers of missing items across multiple items. A comparison of the cases with and without missing items indicated that they differed to a statistically significant degree suggesting that the missing data points were not random. Because of the large number of missing data points missing data was imputed through an estimation procedure. Caution is urged in interpreting the results of the SET factor analysis as these results very likely lack both validity and generalizability.

The SET includes seven scales designed to assess and evaluate the critical features of school-wide effective behavior support. The seven scales are labeled using letters and include: Expectations Defined (A), Behavioral Expectations Taught (B), On-going System for Rewarding Behavioral Expectations (C), System for Responding to Behavioral Violations (D), Monitoring & Decision-Making (E), Management (F), and District-Level Support (G). The first factor was defined by 15 items and included items from each of the seven scales. Because of the diversity of items included in the first factor it is nearly impossible to interpret the factor as a unitary whole. Six items defined the second factor. Similar to the first factor, the second factor included items from multiple original scales. A review of the items in the second factor does not reveal a distinct pattern and thus it is not possible to determine the distinct construct that this factor measures.

The third factor includes only three items and thus is poorly defined. Two of the three items assess administrator report on team function and constitution. The third item asks, "Can the administrator identify an out-of-school liaison in the district or state?" It may be that the third factor relates to the consistency of administrator reporting independent of what is reported. Alternatively it may be that the third factor relates to the constitution and functioning of the team.

The fourth factor was defined by five items. Two items assess the accuracy of staff and student reporting of school rules, two items assess for the presence of a crisis plan and agreement between staff and administration on how to handle emergencies. The fifth item is unrelated to any of the others in the factor. Based on these items the fourth factor may represent the level of agreement among staff, students and administration on a common set of rules and procedures. Several issues have very likely contributed to the conceptually unclear factor structure of the SET. The first is the large number of missing items in the data set. A second has to do with the number and nature of scales in the SET. The SET includes seven scales some of which are assessed by only two items. Three of the seven original scales have been shown to have poor to unacceptably low alpha coefficients. This finding suggests poor internal consistency of a number of items intended to measure a similar construct. Given that three out of seven scales demonstrate poor internal consistency and several are poorly defined by only a few items it appears reasonable to expect that only four factors would emerge with the items from the other three scales spread throughout these four. To a degree this appears to be the case. However, even the items in scales demonstrating strong internal consistency did not consistently cluster together in distinct factors. In sum, although exploratory factor analysis of the SET produced a four factor oblique solution this solution does not demonstrate simple structure and consequently is very difficult if not impossible to interpret using the present data set. Further research is needed to confirm or disconfirm the factor structure of the SET. These results also suggest that item deletion and/or addition may be needed to improve the factor structure of the SET.

### *Discussion of Correlations between Factors and Scales*

The factors of the SET, EBS 1.5, and OSSS were correlated to determine the nature and direction of the relationships between each. With only one exception, all EBS-CS and EBS-PI factors were significantly correlated. This suggests that both measures are assessing similar constructs, however the strength of the correlations do not suggest that both the EBS-CS and EBS-PI are measuring identical constructs.

It is significant to note that, with the exception of the fourth factor of the EBS-CS and EBS-PI, all factors between the two measures showed a significant inverse relationship. This finding indicates that when school staff rate the current level of implementation of positive behavior support systems in a school as high they also rate the priority for improving those systems as low. Conversely, when staff rate the current level of implementation of positive behavior support systems as low they rate the priority for improvement of those systems as high. This finding has face validity given that one would not expect staff to indicate a need for improvement of something that is already in place nor would one expect staff to rate the priority for improvement of something as low that is not in place.

Only the first and fourth factors of the SET showed any significant correlations with any of the EBS-CS or EBS-PI factors. The first factor of the SET was not interpretable therefore it is not clear what this correlation means. The fourth factor of the SET was interpreted to represent the level of agreement among staff, administration, and students on a common set of rules and procedures. This finding suggests a degree of overlap between the SET and EBS 1.5 in that when the SET factor four indicates a high level of agreement among staff, students, and administration the EBS 1.5 also indicates a high level of current implementation. The SET fourth factor demonstrates two significant negative correlations with two EBS-PI factors. The SET measures the current status of implementation, like the EBS-CS. A negative correlation between the SET and EBS-PI indicates, as was the relationship between the EBS-CS and EBS-PI factors, that when the current level of implementation of positive behavior support is rated as low on the SET the Priority for Improvement on the EBS-PI is rated as high, at least across several factors.

These results indicate minimal but statistically significant correlations between the SET and EBS 1.5. However, these results should be interpreted with caution due to the problems previously noted with the SET results. No statistically significant relationships were found between the OSSS and the SET. This may indicate a true lack of relationship between the two measures or may be a product of the problems noted with the SET and the data sample. Most likely it is a product of some combination of the two. Provided that the SET and EBS 1.5 do correlate significantly across several factors in the present study it is significant to note that the SET and OSSS do not correlate significantly across any factors despite the fact that the EBS 1.5 and OSSS do. This finding warrants further investigation.

The OSSS Protect factor was significantly and positively correlated with all four factors of the EBS-PI and significantly and negatively correlated with the majority of EBS-CS factors. The OSSS Risk factor was significantly and negatively correlated with all factors of the EBS-PI but not significantly related to any EBS-CS factors. This finding is important, indicating a clear relationship between the implementation (or lack) of positive behavior support in a school and school and community risk and protective factors. More specifically, these results suggest that positive behavior support systems in a school are associated with protective factors and that the lack of such systems in a school is associated with increased risk factors in a school. Due to the fact that these results are only correlational and not the result of controlled experimental manipulation causation may not be inferred. However, these results do suggest the possibility of a causal relationship and suggest an area for further study.

*Discussion of predicting school safety/disorder variables using student report*

Using student report, the Student Victimization scale of the OSCSS was significantly predicted by three other OSCSS scales including Clarity of Rules, Positive Peer Association, and School Effort. All three independent variables were negatively related to Student Victimization as evidenced by negative beta coefficients. This finding indicates that as students report greater clarity of the rules in their school, more positive peer associations, and greater effort in doing their schoolwork they also report fewer instances of being victimized. The strongest predictor of Student Victimization was School Effort followed by Clarity of rules.

Self-reported Student Rebellious Behavior was significantly predicted by the OSCSS scales Respect for Students, Planning and Action, Attachment to School, and Belief in the Rules. With the exception of the scale Planning and Action all independent variables showed an inverse relationship with Student Rebellious Behavior. This indicates that as students report feeling more respected by staff, feel a greater bond to their school and espouse greater belief in the validity of the rules in the school they are less likely to engage in a variety of inappropriate behaviors at school including threatening other students, cheating, and stealing. The strongest predictor of Student Rebellious Behavior was Belief in the Rules followed by Attachment to School.

It is interesting to note that when students reported greater levels of Planning and Action in their school they also reported engaging in higher levels of Student Rebellious Behavior. These findings contradict those of a previous study in which a measure of planning and action in schools was found to be negatively associated with measures of school disorder (Welsh, Green & Jenkins, 1999). One possible explanation for this finding might be that when students report high levels of organization and action in their school they may perceive this as too much control and so rebel against it. These results replicate a number of previous studies in schools that have found an inverse relationship between students reporting high levels of respect, perceived fairness of the school rules, and high levels of student influence in the school and several measures of school disorder. The finding that high levels of student belief in the school rules predicts lower levels of problem behaviors was also a replication of previous findings. (Welsh, Green & Jenks, 1999; Welsh, 2000, Welsh, 2001).

Student reported Drug Availability was significantly predicted by the scales Clarity of Rules and School Effort. Both independent variables were inversely related to Drug Availability indicating that as students reported greater clarity of the rules and greater effort on their schoolwork they also reported less availability of drugs. Sullivan and Farrell (1999) found that a measure of commitment to school served as a protective factor for later drug use. These appear to be similar findings to the extent that effort at schoolwork constitutes a form of commitment to school.

The scale Drug Availability asked students to indicate how easy it would be for someone like them to obtain several different drugs and did not directly assess the presence of drugs in the school. Because of this it is unclear whether schools with students reporting high Clarity of the Rules and high School Effort have less drugs available to students or whether students who indicate that the school rules are clear to them and who report putting forth effort on their school work are less exposed to the drugs that may be in their school.

### *Practical Implications of this Study*

There are several practical implications of the results of this study. The first, originally a product of practical necessity, relates to improved efficiency of measurement. The OSCSS was constructed from selected items of two published surveys, the What About You? student survey (Gottfredson & Gottfredson, 1989) and the Effective Schools Battery (Gottfredson, 1984). Not all of the scales from these measures were considered relevant for the present study. However, neither of these two measures contained all of the desired scales. The desired items from each of these two measures were divided between two OSCSS forms, Form A and Form B. The reason for this was to decrease the amount of time needed to complete the surveys. This approach did in fact decrease the amount of time each student needed to complete his or her survey, based on my informal observation. The OSCSS scales, with a few exceptions, demonstrated acceptable alpha coefficients many of which exceeded previously published Effective Schools Battery alpha coefficients. These findings indicate that it is possible to obtain useful school-wide information using student self report in less time by dividing the items between two measures and only administering a portion of the items to any single student. This finding has practical implications in that decreased data collection time presents less of a burden to students and schools and yet provides the same useful information.

A second related practical implication was that it was possible to substantially and significantly predict measures of student victimization, rebellious behavior, and drug availability, all highly relevant to school safety, using a relatively small set of IVs. The majority of these IVs represent school climate factors amenable to change. For example, one significant predictor of both student victimization and drug availability was the student perceived clarity of rules in the school. As students reported greater perceived clarity of the rules they reported less availability of drugs and victimization. The practical implications of these findings are that schools may reduce both student victimization and drug availability by making the rules more clear.

Student reported rebellious behavior was also significantly and substantially predicted by several IVs that would appear to be constituents of a school's climate and thus changeable. Students reported less rebellious behavior when they felt more respected at school and greater attachment to their school. Both of these IVs are changeable within a school. Schools that focus efforts on increasing respect for students and foster student attachment to school are likely to reap the benefits in reduced student behavior problems.

A final practical implication relates to the use of the EBS 1.5. Not only did the EBS 1.5 demonstrate good internal consistency and clear factor structures, which serve to at least partially validate the measure, but it also demonstrated a significant and meaningful pattern of correlations with the OSSS. When staff rate the school's Current Implementation of positive behavior support systems in their school as high this would appear to constitute a protective factor and thus increase a school's safety, based on the significant positive correlations between the EBS-CS factors and the OSSS Protective factor. When staff rate the Improvement Priority of positive behavior support systems in their school as high, they rate their Current Status of implementation as low and this appears to serve as a risk factor and thus may decrease the safety of their school. This is based on the significant positive correlation between the EBS-PI factors and the OSSS Risk factor. Thus the EBS 1.5 appears to provide a valid measure of the level of implementation of positive behavior supports in a school and also provides a measure of a school's safety to the extent that risk and protective factors determine school safety.

### *Limitations of this Study*

Although a number of significant and relevant results were found in this study there were also a number of limitations. The initial design of the study involved correlating multiple sources of information including the SET, EBS 1.5, OSSS, OSCSS, and various school and community archival data in order to determine the relationships between these various measures and sources of information. The study also originally intended to explore the role of all of these measures and sources of data in predicting the measures of school safety/disorder. However, because these analyses necessarily involved the school as the unit of analysis the sample size of at most four schools did not permit these analyses. It would have been very useful information to have completed these analyses.

A second limitation of the study, as previously discussed, was the large number of missing data points in the SET data. Although the missing data points were estimated, analyses indicated that these missing data were not random. Nonrandom missing data systematically distorts the results and thus affects their generalizability. Consequently, the results of the SET, particularly the factor analytic results, must be interpreted with caution and very likely lack generalizability. This finding also limits our understanding of the relationships between the EBS 1.5 and OSSS based on the results of this study and ultimately limits our understanding of the validity of these measures. Thus the problems encountered in analyzing the SET substantially limited the advancement of our understanding of these measures. However, not to paint too bleak a picture, useful knowledge was gained.

A third limitation of the study, not realized until after data analysis began, deals with the OSCSS. The OSCSS includes fifteen scales divided between two forms, Form A and Form B. Only one of the two forms was administered to any one student. Consequently, it was only possible to conduct correlational analyses between scales included on the same form. This potentially limited the ability to predict the three measures of school safety/disorder. Had all OSCSS scales been available for all analyses the results of the multiple regression analyses may have been different due to the inclusion of different independent variables.

### *Directions for Future Research*

Future research in this area should attempt to answer the questions that were originally propose in this study but were not possible to answer due to limitations in sample size. More specifically, a large enough sample size should be developed in order to examine the relationships between all of the measures and data originally proposed. Such an analysis will very likely yield important findings and contribute to our further understanding of the construct of school safety and disorder and our understanding of how to improve schools.

The analyses involving the SET, EBS 1.5, and OSSS were all exploratory and should be confirmed with further study. This is especially the case with the SET. Future research might also examine ways to improve these measures by either eliminating items or adding new items. It may also be possible to combine these measures into a smaller number of more comprehensive and efficient measures. The results of this suggest study suggest both areas that need improvement and the degree of overlap between measures.

Although the results of this study contribute to our understanding of the internal structures of the SET, EBS 1.5 and OSSS and improve our understanding of external validity by demonstrating the correlations between measures more research is needed. All three of these measures have not as yet been well validated. It is important to validate measures by comparing them with established and well-validated measures. This may be somewhat difficult at present due to the emerging nature of this area of study and the relative dearth of good measurement tools. However, it is important to conduct further validation studies of these measures. A variety of data and measures should be used in doing so. This will serve to address some of the current limitations of school safety measurement previously discussed.

Finally, future research might examine the constructs of school safety and disorder using hierarchical linear modeling. Such statistical methods are more suited to examining nested data as occurs when exploring the relationships between organizations, individuals, and communities. Such statistical models would very likely shed new light on these constructs and would not be as limited by the small sample size that results for more traditional statistics when schools are the unit of analysis. Hierarchical linear modeling would also allow for greater incorporation of community data due to its ability to accommodate nested data. Research indicates that community factors significantly affect children and families. Children bring their communities with them to school and the problems associated with their communities (Plybon & Kliewer, 2002; Smith & Jarjoura, 1988; Schwartz & Gorman, 2003). Consequently, it is important to understand the impact that community variables have on school safety and disorder.

## References

Bay, B. (1975). (Chairman). Our nation's schools- a report card: "A" in school violence and vandalism. Preliminary report to the subcommittee to investigate juvenile delinquency based on investigations, 1971-1975. Committee on the Judiciary, United States Senate, Washington, D.C: U.S. Government Printing Office.

Bowen, N. K., & Bowen, G. L. (1999). Effects of crime and violence in neighborhoods and schools on the school behavior and performance of adolescents. *Journal of Adolescent Research*, 14(3), 319-342.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

CBSNews.com (2005, March 22). Shooter drove police car to school. CBSNews.com. Retrieved April 16, 2005, from http://cbsnews.com/stories/2005/03/21/national/printable68075.shtml

Cernkovich, S. A., & Giordano, P. C. (1992). School bonding, race and delinquency. *Criminilogy*, 30(2), 261-291.

Cohen, J. (1988). Statistical power analysis for the behavioral science (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cohen, J, & Cohen, P. (1983). Applied multiple regression/correlation analysis for the behavioral science (2nd Ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., Ramey, S. L., Shure, M. B., & Long, B. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist*, 48, 1013-1022.

Colder, C. R., Mott, J., Levy, S., & Flay, B. (2000). The relation of perceived neighborhood danger to childhood aggression: A test of mediating mechanisms. *American Journal of Community Psychology*, 28(1), 83-103.

Comrey, A. L., & Lee, H. B. (1992). A first course in factor analysis. (2nd Ed.) Hillsdale, NJ: Erlbaum.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation Design & Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company.

Corvo. K. N. (1997). Community-based youth violence and prevention: A framework for planners and funders. *Youth and Society*, 28, 291-316.

Crosnoe, R., Erickson, K. G., & Dornbusch, S. M. (2002). Protective functions of family relationships and school factors on the deviant behavior of adolescent boys and girls: Reducing the impact of risky friendships. *Youth & Society*, 33(4), 515-544.

Derzon, J. H. (2001). Antisocial behavior and the prediction of violence: A meta-analysis. *Pscyhology in the Schools*, 38(2), 93-106,

Dishion, T.J., Patterson, G.R., Stoolmiller, M., & Skinner, M.L. (1991). Family, school, and behavioral antecedents to early adolescent involvement with antisocial peers. *Developmental Psychology*, 27 (1), 172-180.

Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy*,27(3), 373-390.

Dodge, K. A., Lansford, J. E., Burks, V. S., Bates, J. E., Pettit, G. S., Fontaine, R., & Price, J. M. (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development*, 74(2), 374-393.

Dryfoos, J. G. (1990). *Adolescents at risk: Prevalence and prevention*. New York: Oxford University Press.

Elliott, D. S., Hamburg, B., & Williams, K. R. (1998). Violence in American schools: An overview. In D. S. Elliott, B. A. Hamburg, & K. R. Williams (Eds.), *Violence in American Schools* (pp. 3-28). New York, NY: Cambridge University Press.

Farrington, D.P. (1995). The twelth Jack Tizard memorial lecture the development of offending and antisocial behaviour from childhood: key findings from the Cambridge study in deliquent development. *Journal of Child Psychology and Psychiatry*, 360 (6), 929-964.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment, 7, 286-299.

Furlong, M., & Morrison, G. (2000). The school in school violence: Definitions and facts. *Journal of emotional and behavioral disorders*,8(2), 71-82.

Gallup Poll Analysis. (2001, March). Majority of parents think a school shooting could occur in their community. The Gallup Organization. Retrieved November 20, 2001. Available: http://www.gallup.com/poll/releases/pr010305c.asp

Garmezy, C. J. (1983). Stressors of childhood. In N. Garmezy and N. Rutter (Eds.), *Stress, coping and development in children* (pp. 43 – 84). New York: McGraw-Hill

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 updated (4th ed.). Boston: Allyn & Bacon.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Chronbach alpha reliability coefficient for Likert-type scales. Paper presented at the Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, The Ohio State University, Columbus, OH, October 8-10, 2003. Retrieved August 24, 2004, from http://www.alumni-osu.org/midwest/midwest%20papers/Gliem%20&%Gliem--Done.pdf.

Gorsuch, R. L. (1983). Factor Analysis second edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1983.

Gottfredson, D. C. (2001). *Schools and delinquency*. New York, NY: Cambridge University Press.

Gottfredson, G. D. (1984). The Effective Schools Battery. Odessa, Fla: Psychological Assessment Resources.

Gottfredson, G. D., & Gottfredson, D. C. (1989). What About You? Ellicot City, MD: Gottfredson Associates, Inc.

Gottfredson, G. D., & Gottfredson, D. C. (1992, November). Development and applications of theoretical measures for evaluating drug and delinquency prevention programs. Paper presented at the annual meeting of the American Society of Criminology, New Orleans

Governor's Juvenile Justice Advisory Committee. (1999). Washington's juvenile justice state advisory group. State of Washington, Dept. of Social and Health Services, Office of Juvenile Justice. Retrieved February 18th, 2003. Available: http://wa.gov/juvenilejustice/abouthist.html

Guadagnoli, E. & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. Psychological Bulletin, 103(2), 265-275.

Hatcher, L. (1994). A Step-by-step approach to using the SAS system for factor analysis and structural equation modeling. Cary, NC: SAS Institute Inc., 1994.

Hawkins, J. D., Catalano, R. F., & Miller, J. L. (1992). Risk and protective factors for alcohol and other drug problems in early adulthood: Implications for substance abuse prevention. *Psychological Bulletin*, 112, 64-105.

Hellman, D. A., & Beaton, S. (1986). The pattern of violence in urban public schools: The influence of school and community. Journal of Research in Crime and Delinquency, 23(2), 102-127.

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

Hsia, H. M., & Bownes, D. (1998). Title V: Community prevention grants program. *OJJDP Fact Sheet*, December, #89. Retrieved February 18th, 2003. Available: http://ncjrs.org/pdffiles1/fs-9889.pdf

Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2002). The school-wide evaluation tool (SET): A research instrument for assessing school-wide positive behavior support. Submitted for publication.

Iowa Department of Education (2003). Safe and drug-free schools and communities overview. Iowa Department of Education. Retrieved: February 18th, 2003. Available: http://state.ia.us/educate/ecese/cfcs/sdfsc/

**Kaiser, H. F. (1974). An index of factorial simplicity. Psychometrika, 39: 31-36**

Kaufman, P., Chen, X., Choy, S. P., Peter, K., Ruddy, S. A., Miller, A. K., Fleury, J. K., Chandler, K. A., Planty, M. G., & Rand, M. R. (2001). Indicators of school crime and safety: 2001. *Education Statistics Quarterly*,3(4), 47-66. Retrieved November 6th, 2002. Available: http://nces.ed.gov/pubs2002/2002607_1.pdf

Laird, R. D., Jordan, K. Y., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2001). Peer rejection in childhood, involvement with antisocial peers in early adolescence, and the development of externalizing behavior problems. *Development & Psychopathology*,13(2), 337-354.

Larson, J. (1994). Violence prevention in the schools: A review of selected programs and procedures. *School Psychology Review*, 23, 151-164.

Lee, V. E., & Croninger, R. B. (1996). The social organization of safe high schools. In Borman, K. M., Cookson, jr., P. W., Sadovnik, A. R., & Spade, J. Z. (Eds.), *Implementing Educational Reform: Sociological Perspectives on Educational Policy* (pp. 359-388). Norwood, NJ: Ablex Publishing Corporation.

Lipsey, M. W. (1990). Design sensitivity statistical power for experimental research. Newbury Park, CA: Sage.

Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical Psychology Review*, 10, 1-41.

Loeber, R., & Dishion, T. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin*, 94(1), 68-99.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment a review of evidence and issues. *American Psychologist*, 56(2), 128-165.

Minogue, N., Kingery, P., & Murphy, L. (1999). Approaches to assessing violence among youth. The Hamilton Fish National Institute on School and Community Violence. Retrieved May 23, 2003. Available: http://www.hamfish.org

Moffitt, T.E., & Silva, P.A. (1988). Self-reported delinquency, neuropsychological deficit, and history of attention deficit disorder. *Journal of Abnormal Child Psychology*, 16 (5), 553-569.

Najaka, S. S., Gottfredson, D. C., & Wilson, D. B. (2001). A meta-analytic inquiry into the relationship between selected risk factors and problem behavior. *Prevention Science*, 2(4), 257-271.

Nansel, T., Overpeck, M., Pilla, R., Ruan, W., Simons-Morton, B., Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjusment. *Journal of the American Medical Association*, April 25.

National Crime Prevention Council. (2001). The 2001 National Crime Prevention Survey: Are we safe? Focus on Teens. Washington, D.C: Author.

National Institute of Education (1978). Violent schools – safe schools: the safe school study report to the congress. Washington, DC: U.S. Department of Education.

National School Safety Center (2001). NSSC review of school safety research. National School Safety Center. Retrieved 2/16/03. Available: http://www.nssc1.org/studies/statistics%20resourcespdf.pdf.

Norusis, M. J. (1994). SPSS Professional Statistics 6.1. SPSS Inc., Chicago, IL.

North Central Regional Education Laboratory. (2003). Summary of Goals 2000: Educate America Act. Retrieved February 18th, 2003. Available: http://ncrel.org/sdrs/areas/issues/envrnmnt/stw/sw0goals.htm

Orphinas, P., Kelder, S., Murray, N., Fourney, A., Conroy, J., McReynolds, L., & Peters, R., Jr. (1996). Critical issues in implementing a comprehensive violence prevention program for middle schools: Translating theory into practice. *Education and Urban Society*, 28, 456-472.

Pearson, F. S., & Toby, J. (1991). Fear of school-related predatory crime. *Sociology and Social Research*, 75, 457-479.

Plybon, L. E., & Kliewer, W. (2002). Neighborhood types and externalizing behavior in urban school-aged children: Tests of direct, mediated and moderated effects. *Journal of child and family studies*, 10(4), 419-437.

Ringwalt, C. L., Greene, J. M., Ennett, S. T., Iachan, R., Clayton, R. R., & Leukefeld, C. G. (1994). Past and future directions of the D.A.R.E program: An evaluation review. Research Triangle Institute. Retrieved: February 19, 2003. Available: http:\\ncjrs.org/txfiles/darerev.txt.

Rose, L. C., & Gallup, A. M. (2000). The 32nd annual Phi Delta Kappa/Gallup Poll of the public's attitudes toward the public schools. The Gallup Organization.

Schwartz, D., & Gorman, A. H. (2003). Community violence exposure and children's academic functioning. *Journal of Educational Psychology*, 95(1), 163-173.

Shoffner, M. F., & Vacc, N. A. (1999). Psychometric analysis of the Inviting School Safety Survey. *Measurement and Evaluation in Counseling and Development*, 32, 67-74.

Simons-Morton, B. G., Crump, A. D., Hayne, D. L., & Saylor, K. E. (1999). Student-school bonding and adolescent problem behavior. *Health Education Research*, 14(1), 99-107.

Smith, D. A., & Jarjoura, G., R. (1988). Social structure and criminal victimization. *Journal of Research in Crime and Delinquency*, 25(1), 27-52.

Snyder, H.N., & Sickmund, M. (1995). Juvenile offenders and victims: A focus on violence. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.

Snyder, H.N., & Sickmund, M. (1999). Chapter 2: Juvenile victims. Juvenile *Offenders and Victims: 1999 National Report*. National Center for Juvenile Justice.

Spector, P. E. (1992). Summated rating scale construction an introduction (Sage University Paper series on Quantitative Applications in the Social Sciences, No. 07-082). Newbury Park, CA: Sage.

Sprague, J., Colvin, G., & Irvin, L. (1995). The Oregon School Safety Survey. Eugene: University of Oregon.

Sprague, J., Colvin, G., Irvin, L., & Stieber, S. (1998). Assessing school safety in Oregon: How do school principals respond? *Effective School Practices*, 17(2), 36-44.

Sullivan, T. N., & Farrell, A. D. (1999). Identification and impact of risk and protective factors for drug use among urban African American adolescents. *Journal of Clinical Child Psychology*, 28(2), 122-136.

Tabachnick, B. G., & Fidell, L. S. (1996). Using Multivariate Statistics third edition. New York, NY: HarpersCollins College Publishers, 1996.

Thornberry, T. P., Lizotte, A. J., Krohn, M. D., Farnworth, M., and Jang, S. J. (1991). Testing interactional theory: An examination of reciprocal causal relationships among family, school, and delinquency. *Journal of Criminal Law and Criminology*, 82, 3-35.

U.S. Department of Education & U.S. Department of Justice. (1999). Chapter 1: The nature and scope of school violence. In 1999 Annual Report on School Safety. Author.

U.S. Department of Health and Human Services (2001). *Youth violence: A report of the Surgeon General*. Rockville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; Substance Abuse and Mental Health Services Administration, Center for Mental Health Services; and National Institutes of Health, National Institute of Mental Health.

Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., & Kaufman, M. J. (1996). *Journal of Emotional and Behavioral Disorders*, 4(4), 194-209.

Walker, H. M., Steiber, S., & Bullis, M. (1997). Longitudinal correlates of arrest status among at-risk males. *Journal of Child and Family Studies*, 6(3), 289-309.

Walker, H. M., & Sprague, J. R. (1999). The path to school failure, delinquency, and violence: Causal factors and some potential solutions. *Intervention in School and Clinic*, 35(2), 67-73.

Welsh, W. N. (2000). The effects of school climate on school disorder. *Annals of the American Academy of Political and Social Science*, 567, 88-107.

Welsh, W. N. (2001). Effects of student and school factors on five measures of school disorder. *Justice Quarterly*, 18(4), 911-947.

Welsh, W. N., Greene, J. R., & Jenkins, P. H. (1999). School disorder: The influence of individual, institutional, and community factors. *Criminology*, 37(1), 73-116.

White, J.L., Moffit, T.E., Earls, F., Robins, L., & Silva, P.A. (1990). How early can we tell?: Predictors of childhood conduct disorder and adolescent delinquency. *Criminology*, 28 (4), pp. 507-527.

Yu, C.H. (2004). An introduction to computing and interpreting Chronbach Alpha in SAS, Paper 246-26. Retrieved August 24, 2004, from http://www2.sas.com/proceedings/sugi26/p246-26.pdf

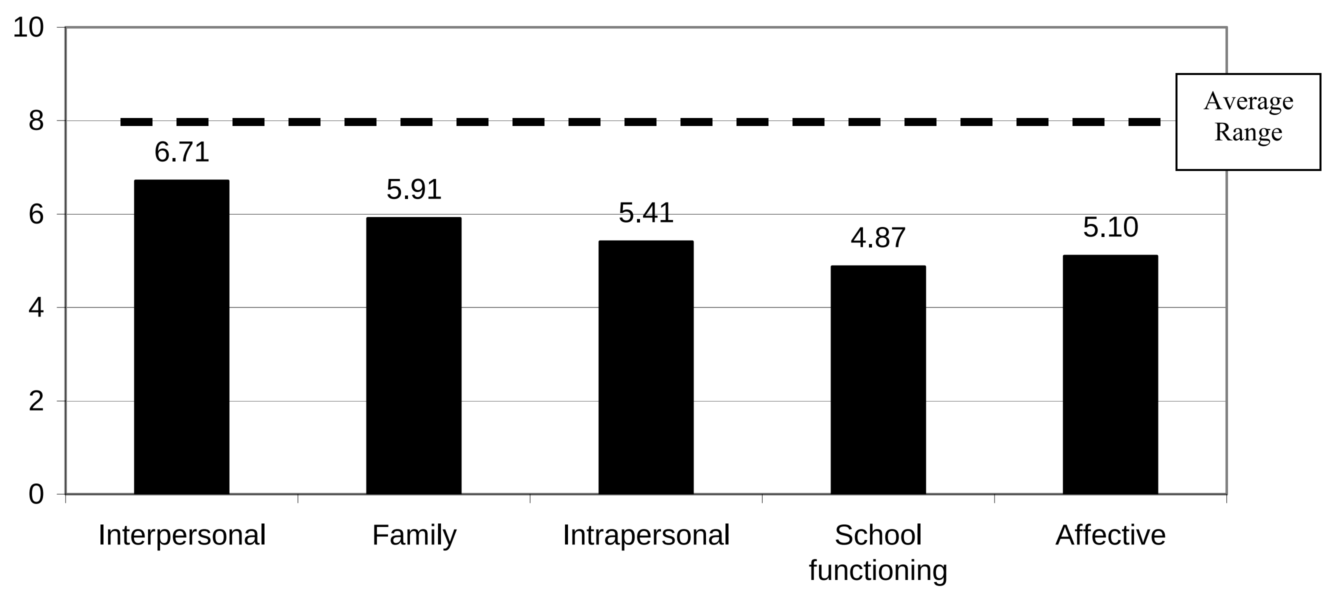
**Attachment B**

***Universal Screening Results***

**Universal Screening.** The SFS project continues to investigate the efficacy of Universal Screening procedures that assist middle and high schools in early identification of students at-risk for dropout, academic failure, and school violence. We currently have universal screening data for 5 high schools and 1 middle school that serve 7,405 students. Each participating high school administered Universal Screening procedures for their 9th grade students and the middle school administered these procedures for their 6th grade class. We purposely selected these grade levels because our research indicates students in 6th and 9th grade report higher levels of school adjustment problems, victimization, and overt aggression in schools. As a result, our school partners administered the universal screening procedures to 2,185 students in grades 6 and 9. Of these 2,185 students, teachers nominated 454 students or 21% of the total sample for further screening using the Behavioral and Emotional Rating Scale (BERS).

As articulated in previous reports, we used the BERS (Epstein & Sharma, 1998) as the screener for several reasons: (a) this measure meets the psychometric standards for individual assessment, (b) is a strengths-based measure that outlines skill and protective factors that would improve the student's success in school, and (c) is easily administered by classroom teachers for students aged 6-18 years old. The participating teachers completed the BERS rating scale for 454 students. Of these 454 completed surveys, 11 or 2% had missing data that prevented scoring. Figure 1 shows the mean scores for the five BERS subscales; a score of 8 or higher is in the average and above range. Overall, the total BERS mean score for students identified as at-risk by the universal screening procedures was <u>69.95</u> or in the <u>very poor</u> range of protective factors.

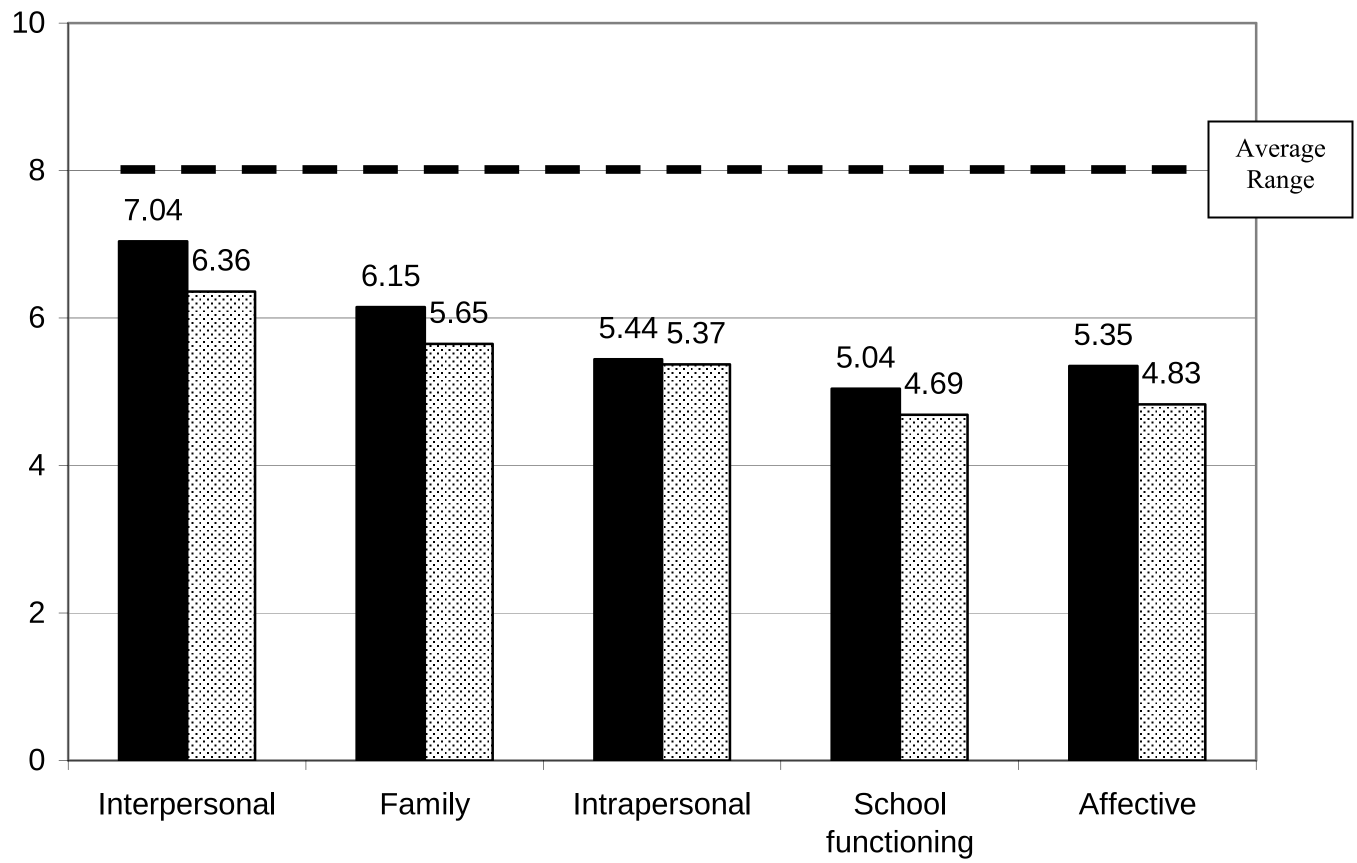
Figure 1 Group Mean Scores for BERS Subscales
(n = 370 students)



Overall, 370 of the nominated students or 84% scored in the very poor to below average range of protective factors. Consequently, the universal screening procedures identified **17%** of the total sample as at-risk or high risk for emotional and behavioral disorders. Teachers reported the universal procedure took approximately 60 minutes or one preparation period to complete.

An important value for the development of the Universal Screening procedures developed for the Skills for Success project was consideration of gender bias that overlooks identification of at-risk girls. To date, the ratio of boys to girls identified as emotionally disturbed is 4 to 1. The SFS Universal Screening procedures identified 194 boys (52%) and 176 girls (48%). A comparison of the differences between boys and girls were different across their total BERS scores and the following BERS subscales at the .05 level of statistical significance: (a) interpersonal skills, (b) family involvement, and (c) affective skills. We found no differences between boys and girls across the intrapersonal and school functioning subscales.

Figure 2 Group Mean Scores for BERS Subscales
(n = 370 students)




The Universal Screening procedure continues to have promise in assisting schools to organize services for at-risk and high risk students. The SFS Universal Screening procedures use minimal teacher time, appears to reduce the over-representation of boys identified as emotionally and behaviorally disorders, and provides schools with skill information for individual students.