

Validity of Office Discipline Referral Measures as Indices of School-Wide Behavioral Status and Effects of School-Wide Behavioral Interventions



Larry K. Irvin
Tary J. Tobin
Jeffrey R. Sprague
George Sugai
Claudia G. Vincent
University of Oregon

Abstract: Office discipline referrals (ODRs) are widely used by school personnel to evaluate student behavior and the behavioral climate of schools. In this article, the authors report the results of a review of the relevant literature to evaluate the validity of ODR data as indices of school-wide behavioral climate, the effects of school-wide behavioral interventions, and differing behavior support needs across schools. They used Messick's unified approach to validity by focusing on examples of evidence for empirical and ethical foundations of interpretations, uses, and social consequences of ODR measures at the school-wide level. The authors also discuss ongoing issues, study limitations, and related recommendations for interpretations and uses of ODR measures as school-wide indices, based on the existing literature.

In 1708, Cotton Mather implored his fellow colonists to send their children to school to “qualify them for future Serviceableness and have their Manners therewithal well-formed under a Laudable Discipline” to prevent “barbarous ignorance” leading to “outrageous wickedness” that could threaten the very survival of the colony (Mather, 1828/1975, p. 21). Much has happened in education in the United States since Mather made his impassioned plea; yet, some of the same issues continue to fuel debates regarding the effectiveness of U.S. schools in fulfilling the responsibility of educating children to be not only literate and knowledgeable about the arts and sciences but also well-behaved citizens.

Recently, the antisocial, and even violent, behavior of some children in schools has become a most pressing concern (Sprague et al., 2002; Walker & Shinn, 2002). Within this context, schools staff are being held responsible for ensuring safe environments where all children can learn appropriate academic and social skills. In accomplishing this responsibility, they face a number of challenges, foremost of which are decisions regarding instructional and disciplinary practices in the context of multiple—and sometimes conflicting—cultural values concerning the role of schools in our society. Mather spoke from the vantage point of a privileged citizen in a rather homogeneous society, but contemporary U.S. society typically does not grant a similarly largely unquestioned authority to individual persons.

U.S. communities and schools have become pluralistic networks of citizen patrons and policymakers, educational administrators, diverse family and student populations, teachers, and curricular strategies. Within this increasingly diverse context, communities and their school systems are experiencing a variety of antisocial behaviors that challenge their abilities to effectively educate children (see Paine & Paine, 2002; Snell, Mackenzie, & Frey, 2002).

In response to antisocial and disciplinary incidents, many school administrators and licensed staff have turned to school-wide behavioral support programs (Dwyer, 2002; Sugai, Horner, & Gresham, 2002). These programs are premised on the assumption that when all school staff members in all school settings actively teach and consistently reinforce appropriate behavior, the number of students with serious behavior problems will be reduced and the school climate will improve. More specifically, the conceptual foundation for school-wide behavior support programs is the following logic or “interpretive argument” (Kane, 1992): Behavioral climate of schools should vary as a function of the local building and district-level interplay of several factors:

- student behaviors and attitudes
- school and classroom characteristics
- school-wide behavioral support planning, including data-based decision-making

- effectiveness of school-wide behavioral support interventions
- educator and student values and related perceptions regarding considerations of school safety, school effectiveness, and the like
- school accountability contingencies at the district, community and state/federal levels (Colvin, Kame'enui, & Sugai, 1993; Lewis & Sugai, 1999; Sugai, Horner, et al., 2000)

The foundations for the interpretive argument regarding school-wide positive behavior support (PBS) are not new (Sugai, Horner, et al., 2000). They are logical extensions of the basic principles of behavior articulated by Skinner (1953) and adapted for practical use through applied behavior analysis and organizational behavior management. Foremost among these principles is the appropriate design of environments (e.g., in schools). Appropriate environments (a) present clear expectations for positive social behavior, (b) offer active instruction to operationally define appropriate behavior, (c) deliver consistent acknowledgement for appropriate behavior, and (d) include systematic intervention to prevent problem behaviors from compromising the effectiveness of ongoing delivery of instruction. Recent extensions of these principles to whole-school interventions (Mayer, 1995) have emphasized (a) the need for investing in the establishment of appropriate social behavior rather than simply punishing problem behavior and (b) the importance of the policies, personnel allocations, and basic organizational systems that govern staff behavior in schools (Zins & Ponti, 1990).

Given a commitment to effective behavior support, schools need valid indices of the school-wide behavioral climate, behavior support needs, and the effectiveness of school-wide behavioral intervention programs in improving the behavioral climate in the school through effective behavior support practices. In addition to being interpretable in these ways, to have a reasonable likelihood of use, such indices must be easily documented and maintained within the daily operations of the school.

Office discipline referrals (ODRs) are commonly available and frequently used school-wide measures of student behavior that, on their face, could serve as such an index. ODR data are collected and stored routinely by schools in a variety of formats (e.g., electronic or paper files) along with attendance records, grades, and academic test scores. Sugai, Sprague, Horner, and Walker (2000) defined an ODR as “an event in which (a) a student engaged in a behavior that violated a rule/social norm in the school, (b) a problem behavior was observed by a member of the school staff, and (c) the event resulted in a consequence delivered by administrative staff who produced a permanent (written) product defining the whole event” (p. 96).

Current common uses of ODRs are in school-based behavior support planning and decision-making about

student discipline and related parent–teacher conferences focused on individual students. ODRs are potentially useful for documenting a school’s behavioral climate and the effects of intervention programs for improving the climate. Given the ready availability of and recommended uses for ODRs (see U.S. Department of Education, 2002), their validity as indices of school behavioral climates and intervention effectiveness need to be evaluated. Such an evaluation will serve to document both the empirical and ethical foundations of such interpretation and use. In their recent article on convergent validity of ODRs, Nelson, Benner, Reid, Epstein, and Currin (2002) specifically noted the need for investigating the validity of using ODRs as indices of effectiveness of school-wide behavioral intervention programs.

In this article, we report the results of our review of examples of empirical and ethical justifications for interpreting and using ODR measures as school-wide indices of various features of schools’ behavioral climates. Our review was guided by the interpretive argument underlying school-wide behavior support programs, because it provided a credible theoretical/conceptual foundation for examining validity of ODRs as indices of school-wide behavioral climate. We used a protocol described by Gersten, Keating, and Irvin (1995) for validating interpretations and uses of measures in education. Gersten et al. noted that the burden of proof rests jointly on educators and researchers, who must work together continually over time, as evidence becomes available, to document the validity of interpretations and uses of measures. The foundation for their protocol was derived from the “unified approach to construct validity” of Messick (1988).

Messick’s Unified Approach to Construct Validity

Over several decades, Messick (1988) developed a framework that has proven broadly applicable for assessing the validity of measures used in educational and psychological assessment. Examples include (a) performance measurements now common in statewide academic assessments in public education (Messick, 1994), (b) assessments of academic achievement in mathematics (Gersten et al., 1995), and (c) questionnaires about incidents of harassment and aggression at school that are commonly known as bullying (Tobin & Irvin, 1996).

Messick has been persuasive in arguing that the more common statistical/methodological approaches to validity (e.g., criterion-related, predictive, concurrent), taken independently of one another, are necessary but insufficient foundations for interpretations and uses of educational measures. He emphasized that we can justify interpretations and uses only by applying a more unified approach to validity, that is, by considering information from all of

the common approaches to validity in a larger, more integrated conceptual frame. In Messick's approach, the validity of any measure can be understood more comprehensively by answering questions of four types:

1. What is the evidential basis for interpretations of the measure?
2. What is the evidential basis for the relevance, utility, and uses of the measure?
3. What is the consequential basis of the values implied in the interpretations?
4. What are the social/educational consequences that result from uses of the measure? (see Table 1).

Table 1 illustrates the comprehensiveness of the unified approach to construct validity. Messick's *evidential basis for interpretation* includes the common statistical ap-

proaches to construct validity (criterion-related, predictive, concurrent). Messick contended that validity evidence supporting interpretation of data from administration of a measure to persons, classes, or schools, although essential, does not necessarily justify using that measure. In addition, that evidence (traditional content and criterion-related validity) does not provide professionals with any information on the ethical, moral, or intellectual consequences of using data to evaluate a student or a school (value implications) or the intended or unintended consequences of its uses (social/educational consequences). Taken together, these latter two aspects of validity constitute Messick's *consequential bases of validity*.

Although we considered other conceptual frameworks for examining the validity of ODRs in school-wide contexts (e.g., Cone, 1997), we chose the Messick framework specifically because of the nature of its emphasis on vali-

Table 1. Messick's (1988) Unified Approach to Construct Validity Applied to School-Wide Interpretation and Use of ODR Measures: Validity Questions and Evidence

Relevant validity question (Messick, 1998)	Evidential basis: Necessary evidence	Consequential basis: Necessary evidence
Interpretation of ODR measures		
What is the empirical evidence justifying interpretations of ODR measures in school-wide contexts?	Traditional psychometric indices of content-related, criterion-related, & construct validity: <ul style="list-style-type: none"> • Correlations with other measures • Within-scale subscale correlations • Factor analyses within & across measures • Differences within & across population groups 	
What are the moral/ethical & intellectual consequences of the implied values when ODR measures are interpreted for school-wide contexts?		Analysis of value connotations: <ul style="list-style-type: none"> • Ethical/moral analysis of implications of labeling based on interpretations of ODRs in school-wide contexts • Scholarly analysis of contribution to theory development
Use of ODR measures		
What is the empirical evidence justifying actual uses, usefulness, & social validity of ODR measures in school-wide contexts?	Empirical data on: <ul style="list-style-type: none"> • Actual uses • Users' perceptions of usefulness • Social validity • Sensitivity to program interventions 	
What are the educational & social consequences of using ODR measures in school-wide contexts?		Analysis of consequences of use: <ul style="list-style-type: none"> • Cost/benefit analyses • Educational/social consequence analysis (e.g., improvement of school climate) • Analysis of unintended side effects of use of ODRs in school-wide contexts

Note. ODR = office discipline referral.

dation of the *uses* of measures in education. The particular contributions of Messick's approach are twofold: First, as do other approaches to assessing validity, Messick directs us to seek evidence justifying our uses of educational measures. Second, in a unique way, Messick urges us to attend to the *value implications* and *educational/social consequences* of our uses of measures. These *consequential implications* are specifically focused on ethical and intellectual considerations related to potential positive and negative value connotations and on actual effects and social consequences that may result from our uses of educational measures.

In our review, we used Messick's (1988) framework as a lens for helping clarify "the appropriateness, meaningfulness, and usefulness" (p. 35) of interpreting and using ODR measures on a school-wide basis. In this article, we first offer examples of empirical evidence we found for interpreting the meanings of ODR measures as school-wide indicators of (a) the behavioral climate of schools, (b) the effectiveness of school-wide behavioral interventions, and (c) the behavior support needs within schools. Second, we report examples of empirical evidence documenting and justifying the actual use of school-wide ODR measures for various purposes, including as indices of the effectiveness of behavioral interventions in schools. Third, we describe examples of evidence regarding foundations for values underlying school/community commitments to effective school behavioral climate. We conclude our review with examples of empirical evidence regarding consequences (costs and benefits) of the validated interpretations and uses of ODRs in school-wide contexts. Finally, we discuss ongoing issues, limitations, and related recommendations for interpretations and uses of ODR measures as school-wide indices based on our review results. Table 2 presents a summary of our research questions organized within the Messick framework.

Strategy for Identifying Relevant Literature

We did not attempt an exhaustive review of the literature on interpreting meanings and making use of ODRs in school-wide contexts; rather, we adapted Kane's (1992) "argument-based approach to validity" by stipulating our interpretive argument and using Messick's framework to complete an initial review to establish whether sufficient evidence existed as preliminary support for our interpretive argument. Such evidence, if identified, would then warrant further, more comprehensive analyses of ODRs as indices of school-wide behavioral climate and the effectiveness of school-wide behavioral support programs. We applied the following criteria in sampling the range of available evidence:

1. The study must be relevant to understanding the meaning and/or uses of measures of problem

behaviors at school that become disciplinary events, including being sent to the principal's office; being "written up" for misbehavior or violating school rules and being assigned a punitive consequence by a school administrator, such as losing a recess; receiving an in-school suspension; being required to attend "Saturday School" or being kept after school; or being sent home, suspended, or, pending the outcome of an expulsion hearing, expelled from school. Suspensions and expulsions from school are two of the most serious consequences delivered by administrative staff that are recorded on the ODR. Because suspensions and expulsions are typically documented as discipline referrals in schools, we included suspensions and expulsions as ODRs, as did Sugai, Sprague, et al. (2000).

2. The study must provide information on one or more facets of Messick's (1988) approach to validity: interpretation of the meaning and/or uses of ODR measures in school-wide contexts or consequences of such interpretations or uses.

We took the following steps in our search:

1. reviewed articles on file from previous literature reviews on school discipline, effective behavior support, and prevention programs;
2. conducted manual searches of tables of contents in current periodicals that report school-based prevention programs and research on interventions to reduce behavior problems (e.g., *Behavioral Disorders, Education and Treatment of Children, Educational Evaluation and Policy Analysis, Exceptional Children, Journal of Emotional and Behavioral Disorders, Journal of Positive Behavior Interventions, Journal of School Psychology, Prevention Science, School Psychology Review*);
3. searched reference lists from Steps 1 and 2; and
4. searched computer databases (e.g., ERIC, PsycINFO, ArticleFirst) using keywords such as *office discipline referral, suspension, expulsion, and school discipline*. The results of our review, organized into the four facets of the Messick framework, are presented next.

Evidence for Interpreting the Meanings of ODRs

ODRS AS INDICATORS OF SCHOOL-WIDE BEHAVIORAL CLIMATE

We found a variety of empirical evidence justifying interpretations of school-wide ODR measures as indicators of school-wide behavioral climate.

Table 2. Messick's (1988) Framework Applied to Defining Research Questions and Relevant Evidence for Interpretation and Use of ODRs as Indices of School Behavioral Climate, Intervention Effectiveness, and School Behavior Support Needs

Relevant research question	Evidential basis: Necessary evidence	Consequential basis: Necessary evidence
Interpretation of ODR measures		
Is there empirical evidence justifying interpretations of the meaning of ODR measures as an index of school-wide behavioral climate?	Predictive & concurrent correlational relationships between ODRs & the following: <ul style="list-style-type: none"> • Student behaviors: aggression, drug use, defiance, behavior disorders, juvenile delinquency • Student attitudes toward: rules, commitment to education • School/classroom features: orderliness, victimization, safety, crime at school 	
Is there evidence regarding the educational/social values associated with interpreting ODRs as indices of school-wide behavioral climate?		Correlational relationships between ODRs &: <ul style="list-style-type: none"> • School safety • School effectiveness • Social control needs • Equity in education
Use of ODR measures		
Is there empirical evidence of actual uses, usefulness, sensitivity to intervention, & social validity of ODRs as indices of effectiveness of school-wide interventions?	Descriptive quantitative/qualitative data on ODRs: <ul style="list-style-type: none"> • Actual uses • Users' perceptions of usefulness • Sensitivity to program interventions • Correlations with users' perceptions of effects of program interventions (social validity) 	
Is there empirical evidence of the educational & social costs/benefits of using ODRs in school-wide contexts?		Program evaluation data on costs/benefits of using ODRs for: <ul style="list-style-type: none"> • School-wide behavior support planning • Data-based decision making • Accountability at district, state, & federal levels

Note. ODR = office discipline referral.

General Misbehavior at School

Bryant, Schulenberg, Bachman, O'Malley, and Johnston (2000, p. 81, Fig. 2) have shown that eighth-grade school misbehavior (as evidenced by being sent to the office, being kept after school, skipping class/school, or being suspended/expelled) had the following correlations with other student behaviors ($p < .001$): Positive: .46 with cigarette smoking; .70 with continued school misbehavior (in 10th grade); Negative: $-.44$ with academic achievement; $-.38$ with school bonding (interest in school work and liking/disliking school).

According to Najaka, Gottfredson, and Wilson (2002), who reviewed the literature on prevention programs and problem behavior, "Crime and delinquency, alcohol and

other drug use, [school] attendance problems, and conduct problems . . . repeatedly demonstrated positive intercorrelations" (p. 260). In the category of conduct problems, they placed misbehavior at school, including "rebellious behavior, non-criminal antisocial behavior, aggressive behavior, defiance of authority, disrespect for others, school suspension, and school expulsion" (p. 259).

Using interview results and responses to an Internet questionnaire from a national sample of 213 girls ages 9 to 19 years (mean age = 15.6, $SD = 2.15$), Smith and Thomas (2000) found that the most common reason given for being expelled was fighting and the second most common was verbal aggression. Girls who had been (a) suspended or expelled for fighting or for bringing a weapon to school or (b) charged with a violent offense by the juvenile justice

system were also “more likely to dislike school and to perceive school discipline as unfair” (p. 547).

Student Perceptions

G. D. Gottfredson and Gottfredson (1999) found that when the self-reports of middle and high school students regarding misbehavior and attitudes were compared to school records, ODRs were most highly correlated with rebellious behavior (.32; e.g., steal, fight, threaten, show off), commitment to education (−.31); peer drug models (.22), drug availability (.21), and attitudes favoring drug use (.17). Other negative correlations included belief in rules (−.23) and parental supervision (−.19). All of these correlations were significant at $p < .01$.

Teacher Perceptions

Tobin, Lewis-Palmer, and Sugai (2002) compared changes in school-wide rates of ODRs to teacher perceptions of changes in student behaviors for a 2-year period. School-wide rates of ODRs were determined every month, using rate data that were standardized across schools of different sizes into a measure of rate per day per 100 students per month. Teachers from five elementary schools who had been teaching at the same school for at least 2 years were asked if they thought the students' behavior, in general, was getting better, staying the same, or getting worse. Teachers also were invited to provide explanations. For four of the five schools, the teachers' perceptions were consistent with schoolwide ODR changes from 1 year to the next.

Classroom Orderliness

Nelson and Roberts (2000) reported direct observation data about the sequences of interactions between teachers and students leading to the child being sent out of the classroom, either to the principal's office or to the hallway. They found that the conditional probability of a direction to leave the room was highest when the disruptive behavior involved a chain of more than four reciprocal teacher–student interactions leading up to the ODR. ODRs thus were associated with high degrees of disorder (on average, four negative teacher–student interactions preceded the ODR).

ACADEMIC FAILURE AND SOCIAL MALADJUSTMENT

Tobin and Sugai (1999b) found that three or more suspensions in ninth grade predicted school failure in high school. Boys referred for fighting more than twice and girls referred even once for harassing, as sixth graders, were not likely to be on track for graduation when in high school. In an analysis of school records, Tobin and Sugai (1999b) also found the following correlations between grade point averages (GPAs) and specific types of ODR behaviors for

boys in sixth grade: fighting (−.26), harassing and threats of violence (−.46), and nonviolent misbehavior (−.44). All correlations were statistically significant at $p < .001$.

In a study of an experimental in-school suspension program, Morrison, Anthony, Storino, and Dillion (2001) examined the relationships among records of ODRs, student self-report surveys, GPAs for the previous and current years, and two reports from assistant principals: (a) a form completed “at the time of the offense . . . included student grade, history of office referrals, . . . previous suspensions” (p. 281) and (b) a rating after the suspension “(1 = *did not improve behavior* to 5 = *improved behavior*)” (p. 281). Students with no previous referrals had higher GPAs, higher levels of personal optimism and social responsibility, more likelihood of endorsing peer norms favoring academic achievement, and lower levels of family conflict than did students with previous referrals ($p < .05$).

Using a “discipline” scale of 1 (*never*) to 5 (*four or more times this year*), with an internal consistency reliability of $\alpha = .81$, Murdock, Anderman, and Hodge (2000) measured the frequency of four student self-reported events: (a) sent to the assistant principal, (b) detention, (c) in-school suspension, and (d) out-of-school suspension. These discipline scores for ninth graders demonstrated a correlation of −.50 with grades ($p < .01$).

Juvenile Delinquency

Huisinga and Jakob-Chien (1998) used data from the *Denver Youth Survey* (Huisinga, Esbensen, & Weiher, 1991) to measure associations between school suspensions and delinquency. Of the students between the ages of 11 and 17 years old who had been suspended, 67% were juvenile offenders (boys = 73%, girls = 53%). When youth were grouped according to the seriousness and violence of their behavior, rates of suspension from school were as follows: 55% suspension rate for youth who committed serious and violent offenses; 42% suspension rate for the youth classified as having “serious but not violent” offenses; 30% suspension rate for youth with minor offenses (e.g., underage drinking, disorderly in public); and 15% suspension rate for youth who were not delinquent.

Behavior Disorders

Interview data from youth with behavior disorders who received special education services in two high schools revealed that 94% had been suspended from school (Kortering, Braziel, & Tompkins, 2002). In research at the middle school and high school levels, Morgan-D'Atrio, Northrup, LaFleur, and Spera (1996) found relationships between suspensions and the following types of measures: (a) the three types of behavior report forms (the *Child Behavior Checklist* [CBCL], the *Youth Self-Report*, and the *Teacher Report Form* [TRF]) standardized by Achenbach

and Edlebrock (1991); (b) the *Social Skills Rating System* (Gresham & Elliott, 1990); and (c) reading and academic achievement tests. A higher percentage of suspended high school students (43%) than middle school students (38%) had clinically elevated scores (2 *SD* above the mean) on one or more of the subscales of the CBCL. A higher percentage of suspended high school students (42%) than middle school students (23%) were ranked high on the Aggression subscale of the TRF. About 52% of high school students and 67% of middle school students had social skills deficits according to one or more measures. Half of the suspended students in both the middle and high schools had academic deficits, and 30% had both academic and social skills deficits.

DURABILITY OF ODR PATTERNS

In an analysis of school records, Tobin and Sugai (1999b) found that ODRs for fighting in sixth grade were positively correlated ($p < .001$) with ODRs for fighting in eighth grade (girls, $r = .61$; boys, $r = .50$). They also found that the frequency of discipline referrals in sixth grade was a statistically significant ($p < .001$) predictor of chronic discipline problems in later middle school, which in turn was a statistically significant ($p < .001$) predictor of frequency of suspensions in ninth grade. Similarly, Murdock et al. (2000) found that the best statistically significant discriminators between students with self-reported high and low levels of discipline referrals in ninth grade were self-reported discipline referrals in seventh grade.

Other research findings have shown that the behavior problems that result in ODRs in school are likely to persist into adulthood. Discipline problems for boys at 8 to 10 years of age have been shown to predict (a) self-reported violence at 16 to 18 years of age and at 32 years of age (odds ratio = 1.9, or almost twice as likely as for boys without discipline problems in elementary school) and (b) violent crime convictions between ages 10 and 32 (odds ratio = 2.4; Farrington, 1989, cited in Hawkins et al., 1998, pp. 119, 133). "The odds ratio indicates the increased likelihood of violence associated with presence of a particular predictor" (p. 108).

In an analysis of 5 years of ODR data for a middle school, Tobin, Sugai, and Colvin (1996) examined correlations between referrals in the first term of sixth grade and referral outcomes over the next four terms of middle school, as measured by (a) frequency of ODRs per term, with a focus on the term when the student had the highest frequency, called the *single term high*, and (b) the percentage of terms with at least one ODR, a measure of *duration* and an indicator of chronic disciplinary problems. Tobin et al. found the following significant correlations for sixth grade: first-term ODR frequency and single term high = .74 ($p < .001$), first-term ODR for harassment and single

term high = .52 ($p < .05$), first-term ODR frequency and later duration = .59 ($p < .05$), and first-term ODR for harassment and later duration = .56 ($p < .05$).

In a series of studies following the maturation of two groups of boys, one identified as "antisocial" and the other as "at-risk controls," discipline referrals in the elementary and middle schools were statistically significant predictors ($p < .01$) of later antisocial behaviors (Walker, Colvin, & Ramsey, 1995; Walker, Shinn, O'Neill, & Ramsey, 1987; Walker, Stieber, & O'Neill, 1990; Walker, Stieber, Ramsey, & O'Neill, 1990, 1991, 1993). Another series of reports using discipline referrals and other school records to describe events and outcomes found that early in their school careers, both girls and boys with more discipline referrals than their peers were more likely to continue to have discipline problems later, were at elevated risk for identification as "emotionally disturbed" and restrictive placements, and were not likely to be on track for graduation when in high school (Tobin & Sugai, 1999a, 1999b; Tobin et al., 1996; Tobin, Sugai, & Colvin, 2000; Walker, Block-Pedego, Todis, & Severson, 1991).

Wright and Dusek (1998) also found evidence of the durability of ODR patterns when they examined records from two urban schools. They studied recidivism, which they defined as "multiple discipline referrals for the same student in the same school year" (p. 142). In one school, the mean probability that any student with at least one ODR would be referred again in the same school year was 81% in Year 1 ($SD = 7.06\%$), 80% in Year 2 ($SD = 7.22\%$), and 77% in Year 3 ($SD = 6.5\%$). For the second school, the mean probability of another referral for any student referred once was roughly similar but a bit lower: 77% in Year 1 ($SD = 10.4\%$) and 72% in Year 2 ($SD = 7.9\%$). The researchers concluded that in spite of differences between the two schools, within-school rates of ODR recidivism were highly stable.

SUMMARY

Based on the empirical studies we reviewed (see Table 3), we found that elements of school-wide behavioral climate include general student misbehavior, school attendance, students' and teachers' perceptions of safety and victimization, classroom orderliness, students' experiences of academic success (or failure) and social adjustment (or maladjustment), the prevalence of juvenile delinquency and behavior disorders, and the durability of patterns of disciplinary incidents within a school over time. In the examples we cited previously, higher levels of school-wide ODRs were associated with higher levels of problematic behavioral climates in schools. When a school experiences increases in ODRs, these increases probably have occurred in the form of one or more student misbehavior, victimization, academic failure, social maladjustment, juvenile

Table 3. Messick's (1988) Framework Applied to Summarizing Sources of Evidence for Interpretation and Use of ODRs as Indices of School Behavioral Climate, Intervention Effectiveness, and School Behavior Support Needs

Example	Evidential basis: Sources of evidence	Consequential basis: Sources of evidence
Interpretation of ODR measures		
Evidence justifying interpretations of the meaning of ODRs as an index of school-wide behavioral climate	<p><i>General student misbehavior & classroom orderliness:</i> Bryant et al. (2000), G. D. Gottfredson & Gottfredson (1999), Najaka et al. (2002), Nelson & Roberts (2000), Smith & Thomas (2000), Tobin et al. (2002)</p> <p><i>Academic failure & social maladjustment:</i> Morrison et al. (2001), Tobin & Sugai (1999b)</p> <p><i>Juvenile delinquency:</i> Huisinga et al. (1991)</p> <p><i>Behavior disorders:</i> Kortering et al. (2002), Morgan-D'Atrio et al. (1996)</p> <p><i>Durability of ODR patterns:</i> Hawkins et al. (1998), Murdock et al. (2000), Tobin & Sugai (1999b), Tobin et al. (1996), Walker et al. (1995), Wright & Dusek (1998)</p>	
Evidence of the educational/social values foundation for interpreting ODRs as indices of school-wide behavioral climate		<p><i>School safety:</i> G. D. Gottfredson et al. (2000)</p> <p><i>School effectiveness:</i> Bryant et al. (2000)</p> <p><i>Social control needs:</i> Jessor & Jessor (1977)</p> <p><i>Equity in education:</i> Abbeville County (SC) School District (1998), Scott (2001)</p>
Use of ODR measures		
Evidence of actual uses, usefulness, sensitivity to intervention, & social validity of ODRs	<p><i>Assessment of school policies & procedures:</i> Bingham et al. (1993), Morgan-D'Atrio et al. (1996)</p> <p><i>Assessment of risks from student threats:</i> Ryan-Arredondo et al. (2001)</p> <p><i>Sensitivity to program interventions (including social validity):</i> Character Counts! (2002), Cheney et al. (2004), Colvin et al. (1993), Cowen et al. (1996), Dwyer & Osher (2000), Flay et al. (2001), D. C. Gottfredson (2001), Hagan-Burke et al. (2004), Metzler et al. (2001), Nelson et al. (1996), Sprague et al. (1999), Sugai et al. (2000), Tobin et al. (2002)</p>	
Evidence of the educational & social costs/benefits of using ODRs in school-wide contexts		<p><i>School-wide behavior support planning:</i> Wright & Dusek (1998)</p> <p><i>Data-based decision making:</i> Cordori (1987), Guzmán (1996), Nakasato (2000), Sadler (2000), Todd et al. (2002)</p> <p><i>Accountability at district, state, & federal levels:</i> Scott (2001), U.S. Office for Civil Rights (1998)</p>

Note. ODR = office discipline referral.

delinquency, or behavior disorders. Without school-wide behavior support interventions, high ODR levels and problematic school behavioral climate are likely to persist. The evidence supports the interpretation of ODRs as school-wide behavioral climate indicators.

Evidence for Relevance, Use, and Utility of ODRs

Office discipline referral measures have often been used as outcome measures in both formative and summative eval-

uations of interventions, as well as in research to determine the effectiveness of interventions intended to improve the behavioral climate of a school environment. Wright and Dusek (1998) listed four specific uses of ODR measures:

1. to measure aggression,
2. to compare schools and different interventions or approaches to management in schools,
3. to compare and make predictions about behaviors within student subgroups, and
4. to identify early in the school year students who have special behavioral support needs.

According to Guskey (2000), school records concerning being sent to the office for disciplinary action, student behavior problems, vandalism, and suspensions are valuable for assessing staff development efforts, making comparisons between classrooms and schools, and making comparisons from one time to another. Guskey noted that they are “vitaly important to many stakeholders” (p. 234). Examples of the use of ODR measures in assessing the concerns of local staff members and students include assessment of school policies and procedures and of the risks posed by threats by students.

ASSESSMENT OF SCHOOL POLICIES AND PROCEDURES

Office discipline referral data also may be used to evaluate the extent to which school staff members are following written school discipline policies or are in agreement on key issues. Morgan-D’Atrio et al. (1996) used a computer printout of ODRs to determine that about 45% of the disciplinary actions in a school did not correspond to its written policies and 20% of the suspensions violated those policies. When administrators do not work with teachers to monitor and solve school-wide discipline problems, they may not be aware of the lack of consensus among staff members. For example, Bingham, Haubrich, and White (1993) surveyed teachers and principals and analyzed their agreements and disagreements on a variety of issues, including the levels of staff consensus regarding how to handle school discipline problems and the overall quality of the school. A correlation of $r = .36$ ($p < .001$) was found between level of disagreement on how to handle discipline problems and level of disagreement about school quality. Principals tended to be more likely than teachers to think that staff consensus on discipline was high and that the quality of the school was high.

ASSESSMENT OF RISKS POSED WHEN STUDENTS MAKE THREATS

Office discipline referral data are part of the *Dallas Violence Risk Assessment*, a systematic assessment of the level of risk of violence when students make threats at school and a decision must be made quickly about which actions

to take (Ryan-Arredondo et al., 2001). The authors noted that “all threats are taken very seriously” (p. 187) and advised that the student’s parents be called. The extent of additional action (e.g., calling the police, placement changes) depends on the level of risk as indicated by ODR data. The student’s discipline record is interpreted as indicating relatively low risk if no history of previous ODR is indicated; medium risk if previous referrals were for fighting, harassing, or being verbally abusive; and high risk if the student has a history of “discipline problems, criminal offenses, has been removed or expelled” (p. 187). Use of ODR data in this way was based on information from the National School Safety Center (1998).

ODRS AS INDICES OF SCHOOL-WIDE AND CLASSROOM INTERVENTION EFFECTIVENESS

Office discipline referrals appear to be sensitive measures of the effects of interventions designed to change student behavior and to improve school and classroom climate. We found a variety of examples of empirical evidence supporting this type of use.

Findings From Reviews of Intervention Research

In a review of research on school-wide discipline programs, Hagan-Burke et al. (2004) found that the most common measures (used in about 75% of the intervention studies) were from school archival data, including ODR measures and achievement test scores. In a literature review and comprehensive analysis of 110 studies of school-based interventions for preventing delinquency, which was limited to group designs, D. C. Gottfredson (2001) reported a similar finding. Using a category such as the one described earlier in Najaka et al. (2002), where ODR measures were included along with other indicators of conduct problems at school (e.g., survey data where students self-reported misbehavior, teacher ratings), Gottfredson found that it was the most commonly used outcome measure (79% of the group-design intervention studies). The next most commonly used measure was academic performance (54%).

Sensitivity of School-Wide ODRs to Intervention Effects

In a study using an experimental/control group design, Colvin et al. (1993) found

a 12% increase in office referrals in the control middle school and a 50% decrease in the experimental school. In both schools, harassment, defiance, disruption, and fighting were the most common problem behaviors referred to the office. However, after implementation of the TOT [Teacher-of-Teachers, a staff development project related to a proactive, instructional approach to school-wide discipline] program, decreases were observed in each of these

behavioral categories at the experimental school. . . . Slight increases were noted in the control school.” (p. 375)

Nelson, Colvin, and Smith (1996) reported direct observation data on positive and negative interpersonal interactions (both student to student and student to adult) in the cafeteria and in a “before-school” setting (i.e., a common area). They also reported ODR measures for the same times and places. The direct observation data and the ODR measures were both dependent variables designed to measure changes in students’ social behavior as a result of an instructional intervention to establish clear standards for expected, appropriate behavior. Changes in the ODR measures were consistent with changes in the data from the direct observations. Both types of data led the researchers to conclude that the intervention was successful. The average number of ODRs per day dropped from 1.2 to 0.3 in the cafeteria and from 1.8 to 0.2 in the before-school setting. The mean percentage of intervals of disruptive behavior (i.e., negative social interactions) dropped from 37% to 18% in the cafeteria and from 44% to 23% in the before-school setting. Changes in ODR measures reflected changes in directly observed behavior in this study, which used trained observers who were blind to the purposes of the research.

Office discipline referral data for 30 classroom cohorts from five elementary schools were examined over a 3-year period and compared to the number of school staff members who participated in staff development classes on functional behavioral assessment (Tobin et al., 2002). If two or fewer staff members participated, 50% of the classroom cohorts in the school had low/reduced rates of ODRs. If six or more staff members participated, 66.67% to 83.33% of the classroom cohorts had low/reduced ODR rates.

ODRs as Outcome Measures in Program Evaluation

We also examined reports of program evaluations in which ODR measures were used as a dependent or outcome variable. Evidence regarding the utility of ODR measures as indicators of intervention effectiveness in schools was found.

Project ACHIEVE (Knoff & Batsche, 1995) is a school reformation process that emphasizes school-wide, systematic instruction in social skills; a collaborative, team-based approach to problem solving to find ways to provide individualized support for students with challenging behaviors; and decision making based on data and with assistance from consultants, when needed. According to Dwyer and Osher (2000), a formal evaluation of the impact of Project ACHIEVE indicated that the number of disciplinary referrals to the principal’s office and out-of-school suspensions decreased, grade-level retention of students significantly decreased, and the number of students who scored above the 50th percentile on end-of-year

achievement tests significantly increased (see also Knoff, 2000).

For 13 out of 14 schools, the implementation of a character education program called Character Counts! was associated with decreases in the number of suspensions and discipline referrals. For the 14th school, although the same changes did not occur, improvement in school discipline measures was reported. “Discipline referrals and other school records of disruption are useful evidence, since they show actual behavior [as opposed to reports based on surveys]” (Character Counts!, 2002, p. 3 of Does CC! Work?).

Flay, Allred, and Ordway (2001) reported the results of an evaluation of Positive Action, a comprehensive elementary (K–Grade 6) school reform model that included (a) character education, (b) efforts to prevent problem behaviors and enhance academic instruction, (c) daily 15- to 20-minute classroom lessons related to the school-wide reform goals, and (d) family and community outreach. Flay et al. used a matched-control design to evaluate the program in two school districts and found a reduction of 78% to 85% in the number of discipline referrals ($p < .05$) and improvements of 16% to 52% in academic achievement ($p < .0001$).

In an evaluation report of the Primary Mental Health Project (PMHP) for the Department of Education in the state of Connecticut, Cowen et al. (1996) compared children who had been referred to the PMHP with non-referred children and noted that the former had more “formal disciplinary actions, and [incidents of] being sent to the principal” (p. 152). Two evaluations of PMHP in Connecticut found that after intervention, students had (a) “fewer discipline referrals . . . [and improved] social skills, interpersonal relationships, and educational performance” and (b) improved on “adjustment ratings made by teachers, school mental health professionals, and were less likely than comparison children [who did not receive the intervention] to have been . . . referred to the principal for disciplinary action” (p. 186).

In a school that implemented the Effective Behavior Support (EBS) approach to school-wide discipline, a statistically significant ($p < .05$) decrease was documented in the frequency of discipline referrals for seventh graders from baseline (Year 1) to the 2nd year of implementation. This improvement was maintained during the 3rd year (Metzler, Biglan, Rusby, & Sprague, 2001). For boys at all grade levels combined, referrals for harassment declined. In addition, the number of students with high levels of discipline referrals (e.g., more than 10 per year) declined. Both teachers’ and students’ perceptions of safety improved. Most of the school staff members (86%) agreed that behavior had improved, and 100% agreed that “providing recognition to students for positive behavior had a positive impact on students’ behavior” (p. 472).

Cheney et al. (2004) evaluated the effectiveness of a school-wide positive behavior support program across four elementary schools over the course of 2 school years. Using a contingency table analysis of changes in the number of ODRs for three categories (high [> 5 yearly] to low [< 5 yearly], low to high, and no change) across the 2 years for 48 typically developing students and 37 students with externalizing behaviors or emotional and behavioral disorders, Cheney et al. found a significant association ($\chi^2 = 22.52, p = .007$) between ODR change and group membership. Almost all of the typical students showed no change (43 out of the 48 students were low in the number of ODRs at Year 1 and remained low at Year 2), whereas 10 out of the 37 students in the disabled group were high on the number of ODRs at Year 1 and low at Year 2. The Cheney et al. data, along with those of the other evaluation studies cited previously, provide evidence of the sensitivity of ODRs as a measure of school-wide intervention effectiveness.

In another set of evaluation studies, Sprague, Sugai, Horner, and Walker (1999) and Sugai, Sprague, et al. (2000) reported the frequency of ODRs for each of 4 years (1996–2000) of a school-wide positive behavior intervention in 14 elementary and 8 middle schools. As anticipated, most of the schools reported reductions in the number of ODRs across the intervention years of 1996 to 1998. In the 1999–2000 school year, some of the schools adopted a new Web-based ODR tracking system—the School Wide Information System, or SWIS—and reported increases in the number of ODRs during that year. Further detailed analyses of the 1999–2000 data, along with interviews of building administrators, indicated that the increases were not due to dramatic changes in student behavior but instead represented a change in the manner of reporting. The decreases in ODRs from 1996 to 1998 and the increases in 1999–2000 illustrate how sensitive school-wide ODR measures are to behavioral interventions and other contextual characteristics, such as school-wide adoption of an electronic ODR tracking and information system.

SUMMARY OF EVIDENCE

We found evidence regarding the use and utility of ODR measures in (a) reviews of research on school-wide discipline and school-based efforts to prevent juvenile delinquency, (b) assessments of the effects of interventions to improve student behavior at different levels (e.g., on a school-wide basis, in specific settings, by class cohorts, for individual students), and (c) in program evaluations (see Table 3). In many of these examples of evidence for the treatment validity and use of ODRs in program evaluation, other variables (e.g., academic achievement, observed social behaviors) were also measured. These studies thus also offer examples of evidence related to interpretation of

ODRs. In the construct validity literature, this is known as establishing evidence for a hypothesized “nomological network” of constructs that supports the interpretation and use of measures (see Patterson & Bank, 1986).

Foundations for Value Interpretations of ODR Measures

In his unified approach to construct validity, Messick (1988) stated that investigators must be explicit about the value connotations of interpretations and the uses of educational measures. All interpretations and uses of ODR measures connote values about the kinds of behaviors that are considered to be problematic and about what should be done about them. Often these values are implicit and not clearly stipulated. Office discipline referrals are measures of problem behavior, which has been defined as “behavior that is socially defined as a problem, a source of concern, or as undesirable by the norms of conventional society and the institutions of adult authority, and its occurrence usually elicits some kind of social control response” (Jessor & Jessor, 1977, p. 33, cited in Najaka et al., 2002). When we interpret school-wide ODR measures as indices of the behavioral climates of schools or the effectiveness of schools’ efforts to create environments conducive to behavioral success, we therefore need to be explicit about the fact that within such interpretations we are valuing order, safety, and control in the service of school effectiveness.

In our review, we found examples of evidence that appears to support interpreting ODRs as indices of how well such values foundations are being realized in schools (see Table 3). For example, Bryant et al. (2000) noted that “43% of secondary teachers reported that student misbehavior limits to a great or moderate extent their ability to maintain order . . . 25% felt their schools were not effective in preventing school misbehavior” (p. 85). Similarly, with respect to general values for school safety, Gottfredson et al. (2000, Table 1 in Appendix G) found that classroom orderliness, as measured by teachers’ self-reports concerning the use of a variety of behavior management and disciplinary methods, had a high negative correlation ($-.77$) with teachers’ reports of teacher victimization by students and a positive correlation ($.68$) with students’ reports of perceived safety at school. With regard to values for equity in education, Scott (2001) stated that analyses have been conducted of school- and district-wide ODR data to examine the overrepresentation of students from minority ethnic groups in exclusionary practices (e.g., referrals to alternative schools, suspensions, expulsion).

Messick’s contribution regarding the social value connotations of interpretations of educational measures also means that researchers and educators must consider the potential socially devaluing consequences of some inter-

pretations and uses of ODRs. For example, ODR-based labels on school "behavior report cards" could have socially devaluing consequences by characterizing students or schools as behaviorally problematic. In this regard, it seems advisable to reframe the potentially devalued connotations of *behavior problems* as positive behavior supports needed for *taking action toward socially valued outcomes*.

Consequences of Using School-Wide ODR Measures

PLANNING BEHAVIORAL SUPPORT

Individual student histories of behavioral and disciplinary problems may be useful in planning behavioral supports for children in schools. Also particularly useful are aggregated data that provide information about the extent to which systems supporting socially valued student behavior are in place in the school. Office discipline referrals also can be useful as part of teacher consultation and support. Wright and Dusek (1998) pointed out that behavior specialists may arrive at a better understanding of teachers' behaviors by studying discipline referral records and initiating "consultation with the teachers of individual students who display high rates of referrals to assist the teacher in developing a behavior plan . . . also . . . to consult with teachers who frequently refer large numbers of students, to help those instructors use proactive classroom management strategies" (p. 146).

MAKING DATA-BASED DECISIONS

In some schools, a team of teachers, administrators, and other interested school staff members work together to use ODR measures to make data-based decisions (Nakasato, 2000):

An elementary school reviewed its ODR data for the previous 2 years and discovered little, if any, improvement in student behavior despite having . . . [classroom] incentive programs . . . 48% of the total behavioral problems were related to harassment incidents, of which 58% occurred on the playground, with peaks occurring before and after lunch recess. . . the leadership team decided to focus social skills training efforts specifically on the playground and on training lunch recess playground supervisors to provide acknowledgements and incentives for good behavior. (p. 248)

Cordori (1987) reported that a computerized ODR database was used to develop school action plans and was helpful because referrals to the assistant principal were reduced. Similarly, Sadler (2000) noted the following:

One middle school EBS team has learned that 50% to 80% of its discipline referrals were coming from classrooms. On

the basis of these data, the team arranged for increased classroom management support for their teachers. Another middle school team, along with their cafeteria staff, implemented a lunch activity program. As a result, they noticed that office discipline referrals during lunchtime decreased from 10% in the fall of 1998 to 4% in the fall of 1999. (p. 243)

By studying ODRs, the administrators of another school found that fighting at recess was the behavior problem most in need of attention. Using these data, school leaders decided to develop an intervention specifically for the problem, and results were monitored using both ODR measures and a staff survey. The intervention resulted in an 80% reduction in recess discipline problems the 1st year after implementation, and this level of improvement was maintained during the following year. The staff survey indicated that the intervention was perceived as effective and efficient (Todd, Haugen, Anderson, & Spriggs, 2002).

Building consensus and action plans using the processes described by Nakasato (2000), Sadler (2000), and Todd et al. (2002) can help school staff members increase their capacity to support students with behavioral challenges in general education settings and improve the overall quality of the school climate. This type of teamwork may also improve teacher morale. Guzmán (1996) interviewed principals of schools rated as among the most successful in their district in including students with disabilities who had discipline problems and found that key aspects of their success included (a) school-wide efforts, (b) staff teamwork, and (c) use of the advice of behavior specialists.

OFFERING ACCOUNTABILITY

Another consequence of the use of school-wide ODRs relates to satisfying school accountability needs. Increasingly, schools are being asked to demonstrate that they are effective, equitable, and efficient. Scott (2001) reported that a

school was under a great deal of scrutiny from the school district for excluding high numbers of students. In addition, the federal Office for Civil Rights was involved and asking that the school address the overrepresentation of minorities in its exclusionary practices. . . . During the positive behavior support year, . . . a 65% decrease in number of days suspended was realized. . . [and] a 75% decrease in both the total number of students suspended and total number of minority students suspended. (p. 91)

In a 1998 U.S. Office for Civil Rights (OCR) hearing, the Abbeville County (South Carolina) School District successfully used school records of suspensions, expulsions, and referrals to alternative schools in defending against multiple parent claims of racial discrimination. The OCR hearing officer used the school district records of suspen-

sion, expulsion, and referrals as the basis for findings of nondiscrimination in each claimed instance.

Discussion

In this review, we used the unified approach to construct validity template developed by Messick (1988) to document exemplars of the empirical and ethical foundations for the validity of interpretations and uses of school-wide ODR measures. We focused on ODR validity for assessing (a) school-wide behavioral climate, (b) the effectiveness of school-wide behavioral intervention programs, and (c) differing needs across schools in developing positive behavioral environments. We found a substantial basis for interpreting and using ODR measures in these ways. Several important issues require ongoing attention, however, if school-wide interpretation and use of ODR measures is to improve.

VALIDITY OF INTERPRETATIONS AND USE OF ODRS FOR INDIVIDUALS VERSUS SCHOOLS

In a recent review of the literature on administrative discipline contacts (ADCs; ODRs by another name), Nelson, Gonzalez, Epstein, and Benner (2003) concluded that the evidence on the predictive validity of ADCs is “mixed” (p. 271) and that “the concurrent and predictive validity of administrative discipline contacts appears to be relatively limited” (p. 249). Their concerns appear to be based on their and others’ observations of (a) high false positive rates of ODRs (detecting a problem when one does not exist) relative to various standardized social skills, academic achievement, and self-concept scales (Morgan-D’Atrio et al., 1996) and (b) high false negative rates of ODRs (failing to detect a problem when one exists) relative to established borderline and clinical cutoff scale scores on the TRF, particularly with students who demonstrate internalizing problems (Nelson et al., 2002). These concerns relate to important educational decisions that are—or are not—made about individual students based on ODR data; they are not minor, and they clearly deserve more attention and study.

Our focus here, however, is on *school-wide* interpretations and uses of ODR data wherein concerns about false positives and false negatives must necessarily be at the school-wide level. Thus, there are two relevant questions. First, does evidence of false positive/false negative rates of ODRs relative to individual students’ status on clinical, social, and academic measures have a bearing on the interpretation of evidence regarding the validity of ODRs for understanding and making decisions about school-wide concerns such as school climate and school-wide behavioral interventions? Second, is there any evidence of false positive/false negative issues regarding ODR interpretation and use at the school-wide level (e.g., evidence of

“underidentification”/“overidentification” of problematic school-wide behavioral climate and a resulting need for or effects of school-wide behavioral interventions)? Based on our research, our answer to both questions is no. This in no way detracts from the need to continue to investigate, understand, and act on the issues and concerns raised by Nelson et al. (2003) in relation to interpretation and use of ODR measures with individual students. The results of our review demonstrate to us that the issues they have raised are related most directly to the validity of using ODRs with individual students and not directly to such uses school-wide.

UNDERSTANDING ODRS AS INTERACTIONS

Although the weight of the evidence clearly supports using ODR measures in assessing school-wide behavioral climate, the effectiveness of school-wide behavioral intervention programs, and the differing needs across schools, several important issues remain largely unaddressed in the available literature. One of the foremost challenges is understanding that the validity of ODRs is conceptually difficult because each individual ODR measure (i.e., each individual referral) represents a “stream,” or sequence, of events. In other words, each ODR represents the convergence of a student’s response to a given situation, a teacher/staff member’s response to the student’s behavior, and an administrator’s response to the student–teacher interaction. This serial stream of responses occurs within and in response to a value system defined by the members of the school community.

The number of players involved in the ODR process and the potential complexity of interactions among them can be problematic for ensuring consistent outcomes. It seems reasonable to assume that each referring teacher, referred student, and administrator responsible for contingencies related to ODRs brings some degree of idiosyncratic behavioral/cultural standards, management skills/lack of skills, prejudices, expectations, and motivations to any disciplinary event (Morrison & Skiba, 2001). If this assumption is accurate, questions about the reliability of ODR measures (i.e., their consistency within a given school setting) are relevant and legitimate.

Reliability of ODRs

G. D. Gottfredson and Gottfredson (1999) examined students’ ODR records over time and calculated ODR test–retest reliability to be $r = .54$ ($p < .01$) from one year to the next for students in middle and high school. This correlation was not as high as for GPAs ($r = .75$, $p < .01$) from one year to the next. That is, a student was more likely to have the same ranking from year to year in comparison to his or her classmates when GPAs were compared than to have the same ranking from year to year when the numbers of disciplinary referrals were compared.

The ODR measure was found, however, to be more reliable than teacher reports on the revised version of the *Behavior Problem Checklist* (Quay & Peterson, 1988), which had a 1-year test-retest reliability of .36 ($p < .01$) in the G. D. Gottfredson and Gottfredson study. The authors concluded that “archival records may serve as more dependable criterion measures than the structured teacher ratings” (p. 9). In this same vein, Wright and Dusek’s (1998) analysis (described previously) led them to the same conclusion: At least for students who are referred to the administrator because of their behavior more than once, within-school patterns tend to be quite stable.

In the Tobin et al. (2002) study, the agreement of a computerized database of ODRs with the original paper ODRs averaged 86%. Tobin et al. noted that consistency was affected by clerical errors and procedures (e.g., keying in the date the administrator signed the referral rather than the date the teacher wrote the referral, delays in filing paper referrals or entering electronic referrals) that could be corrected through changes in office procedures. Clearly, we need additional studies on the reliability of ODR measures used in school-wide applications.

Standardizing and Triangulating ODR Measures

We believe that the validity of using ODR measures to assess or index school-wide behavioral climate and intervention effectiveness ultimately will depend on schools’ efforts in establishing policies and procedures that minimize or at least take into account the variability of their application(s) of ODR measures. To facilitate such “standardization” in our school-wide work with schools using EBS, we encouraged school staff to (a) create operational definitions of student behaviors, (b) establish mutually exclusive and exhaustive categories for inappropriate behavior, (c) consistently implement clearly defined consequences, and (d) regularly schedule summarizations and use of behavioral data to evaluate and change the behavioral climate of the school (Horner, Sugai, Lewis-Palmer, & Todd, 2001; Lewis & Sugai, 1999; May et al., 2002; Sugai, Sprague, et al., 2000).

In addition, Sprague et al. (1999) noted that in their experience with school-wide behavioral interventions, some schools were not systematically recording or tracking ODR measures. A major aspect of these authors’ work with schools was assisting the school staff in setting up usable, reliable ODR reporting systems. They also found that a common school practice was to purge ODR data at the end of the year, which resulted in incomplete data at the beginning of the next school year.

Along with Sprague et al. (1999) and others, we also encourage school personnel and researchers who are interested in using ODR measures for formative or summative evaluations to “triangulate” their school-wide indicators of behavioral climate. This can be done by collecting results from the use of other types of measures, such as the *Ore-*

gon School Safety Measure (Sprague, Colvin, & Irvin, 1995); other survey or interview data on school climate; or other interview data on the perceptions of teachers, students, administrators, parents, and support staff regarding school climate. Members of school-based behavior support teams can do this on a regular and even informal basis because they “live” in the school. If they see an unexpected or interesting dip, peak, “blip,” or change in trends when they look at ODR data on a monthly basis, they may be able to explain it based on their knowledge of the circumstances.

ONGOING ISSUES

At the school-wide level, the important validity questions about ODRs are largely issues of utility for informing decision making about concurrent and future school-wide behavioral climates. In addition, collaborating educational and research communities must continue to address the criteria we use to judge the validity of ODRs, that is, the *outcomes*. In this regard, more research must be done that focuses on important validity questions such as the following: What is the “behavioral climate” of a school supposed to be? What does an effective school climate look like? Answers will likely depend on the type of school (elementary, middle, high), the values of stakeholders, and a host of other potentially relevant contextual features operating within and around individual schools and districts.

Additional validity questions about school-wide outcomes against which we judge the validities of interpretations and uses of ODR measures need to be addressed in future studies. First, what are the results of school-wide behavior support programs supposed to be—fewer ODRs generally, fewer ODRs for specific subgroups of students, or more accurate record-keeping of the discipline climate? Second, how should the “behavioral support needs” of a school or district be determined? What does a high frequency of student referrals indicate—an increase in problem behavior, inaccurate use of the discipline system, or the need for more behavioral support? How do cultural expectations of the school, families, and the community affect these considerations? The answers to these questions depend, in part, on the goals schools and school districts set and on that same host of other potentially relevant contextual features operating in and around individual schools and districts.

To accomplish this kind of validation of the outcomes of ODR interpretations and uses, complex and difficult issues must be addressed. These include (a) overrepresentation of children from minority ethnic/cultural groups in special education and in suspensions/expulsions from schools (Skiba, Peterson, & Williams, 1997); (b) students with mental health or medical problems and histories of abuse and neglect (Eckenrode, Laird, & Doris, 1993); (c) effectiveness and appropriateness of curricula reflecting cultural pluralism; (d) recruitment/retention of teach-

ers from minority/ethnic groups; and (e) shifts in societal tolerance for dealing with antisocial behavior. The results of several studies (Murdock et al., 2000; Scott, 2001; Smith & Thomas, 2000) have indicated that differences in ODR uses and outcomes could be related to differences in the cultural values of the players involved in the referral process. The large variety of value systems present in our culture makes the progression from empirical observations to normative statement problematic and any generalization across cultural subgroups challenging. The fact that data collection and analysis are always embedded in multiple, overlapping cultural contexts only emphasizes the challenge of formulating interpretations of data or recommendations of data uses that are likely to be accepted as valid across most of—or all—the members of our communities. It is also no small issue that researchers investigating these concerns are likely to use various degrees of interpretive freedom in their evaluations of data (Hirsch, 1967). Messick's inclusion of consequential bases for validity—that is moral, ethical, and social consequences of data interpretation and use—is aimed at increasing our sensitivity to differing standards across different cultural subgroups.

ABOUT THE AUTHORS

Larry K. Irvin, PhD, is a professor of special education at the University of Oregon. His research interests include program development and evaluation, measurement, and research methodology. **Tary J. Tobin**, PhD, is a research associate in educational and community supports at the University of Oregon and focuses on prevention research, function-based interventions, and school-wide behavior support. **Jeffrey R. Sprague**, PhD, is a senior research associate/associate professor in the College of Education and co-director of the Institute on Violence and Destructive Behavior at the University of Oregon. His research centers on violence prevention, school-wide behavior support, and school safety. **George Sugai**, PhD, is a professor of special education in the College of Education at the University of Oregon. His primary areas of interest are positive behavior support, emotional and behavioral disorders, applied behavior analysis, functional assessment, and school-wide discipline. **Claudia G. Vincent**, PhD, is an administrative assistant in educational and community supports at the University of Oregon and is interested in the cultural dimensions of discipline practices and positive behavior support. Address: Larry K. Irvin, Educational and Community Supports, 1235 University of Oregon, Eugene, OR 97403-1235; e-mail: larryi@darkwing.uoregon.edu

AUTHORS' NOTES

1. This study was made possible by Grant H324X010015 from the Office of Special Education Programs, U.S. Department of Education. Opinions expressed herein do not

necessarily reflect the policy of the Department of Education, and no official endorsement by the Department should be inferred.

2. The authors wish to express sincere thanks to Robert Horner, Douglas Cheney, and Russell Gersten for their encouragement and suggestions during the preparation of this article.

REFERENCES

- Abbeville County (SC) School District, 28 IDELR 1208, 28 LRP 5013 (OCR, 1998).
- Achenbach, T. M., & Edlebrock, C. (1991). *Child behavior checklist*. Burlington: University of Vermont, Department of Psychiatry.
- Bingham, R. D., Haubrich, P. A., & White, S. B. (1993). Explaining teacher/principal differences in evaluating schools. *Journal of Educational Administration*, 31(1), 70–76.
- Bryant, A. L., Schulenberg, J., Bachman, J. G., O'Malley, P. M., & Johnston, L. D. (2000). Understanding the links among school misbehavior, academic achievement, and cigarette use: A national panel study of adolescents. *Prevention Science*, 1, 71–87.
- Carran, D. T., Nemerofsky, A., Rock, E. E., & Kerins, M. (1996). Risk of unsuccessful program completion for students with serious emotional/behavioral disorders: An epidemiologic risk analysis. *Behavioral Disorders*, 21, 172–189.
- Character Counts! (2002). *Does CC! work? Records of disruption*. Retrieved February 7, 2003, from <http://www.charactercounts.org/doing/survey-reports.htm>
- Cheney, D., Stage, S., Walker, B., & Blum, C. (2004). *A multi-year, multi-method analysis of school-wide positive behavioral support*. Unpublished manuscript.
- Colvin, G., Kame'enui, E., & Sugai, G. (1993). Reconceptualizing behavior management and school-wide discipline in general education. *Education and Treatment of Children*, 16, 361–381.
- Cone, J. D. (1997). Issues in functional analysis in behavioral assessment. *Behaviour Research & Therapy*, 35, 259–275.
- Cordori, M. (1987). A computer-based recording system for analyzing disciplinary referrals. *NASSP Bulletin*, 71(501), 42–43.
- Cowen, E. L., Hightower, A. D., Pedro-Carroll, J. L., Work, W. C., Wyman, P. A., & Haffey, W. G. (1996). *School-based prevention for children at risk: The Primary Mental Health Project*. Washington, DC: American Psychological Association.
- Davies, J., & Brember, I. (2001). The closing gap in attitudes between boys and girls: A 5-year longitudinal study. *Educational Psychology*, 21(1), 103–114.
- Dwyer, K. (2002). Tools for building safe, effective schools. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 167–211). Silver Spring, MD: National Association of School Psychologists.
- Dwyer, K., & Osher, D. (2000). *Safeguarding our children: An action guide*. Washington, DC: U.S. Departments of Education and Justice, American Institutes for Research. Retrieved April 14, 2002, from <http://www.ed.gov/offices/OSERS/OSEP/ActionGuide>
- Eckenrode, J., Laird, M., & Doris, J. (1993). School performance and disciplinary problems among abused and neglected children. *Developmental Psychology*, 29(1), 53–62.
- Farrington, D. P. (1989). Early predictors of adolescent aggression and adult violence. *Violence and Victims*, 4, 79–100.
- Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the Positive Action Program on achievement and discipline: Two matched-control comparisons. *Prevention Science*, 2, 71–89.
- Gersten, R., Keating, T., & Irvin, L. K. (1995). The burden of proof: Validity as improvement of instructional practice. *Exceptional Children*, 61, 510–519.

- Gottfredson, D. C. (2001). *Schools and delinquency*. New York: Cambridge University Press.
- Gottfredson, G. D., & Gottfredson, D. C. (1999). *Development and application of theoretical measures for evaluating drug and delinquency prevention programs: Technical manual for research editions of What About You (WAY)*. Ellicott City, MD: Gottfredson Associates.
- Gottfredson, G. D., Gottfredson, D. C., Czeh, E. R., Cantor, D., Crosse, S., & Hantman, I. (2000). *A national study of delinquency prevention in schools*. Ellicott City, MD: Gottfredson Associates.
- Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system*. Circle Pines, MN: American Guidance Service.
- Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.
- Guzmán, N. (1996). Leadership for successful inclusive schools: A study of principal behaviours. *Journal of Educational Administration*, 35, 439–450.
- Hagan-Burke, S., Sugai, G., Ingram, K., Thier, K., Garrett, U., Derickson, E., et al. (2004). *Descriptive synthesis of school-wide discipline research*. Manuscript submitted for publication.
- Hawkins, J. D., Herrenkohl, T., Farrington, D. P., Brewer, D., Catalano, R. F., & Harachi, T. W. (1998). A review of predictors of youth violence. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 106–146). Thousand Oaks, CA: Sage.
- Hirsch, E. D. (1967). *Validity in interpretation*. New Haven, CT: Yale University Press.
- Horner, R. H., Sugai, G., Lewis-Palmer, T., & Todd, A. W. (2001). Teaching school-wide behavioral expectations. *Report on Emotional and Behavioral Disorders in Youth*, 1, 77–79, 93–96.
- Huisinga, D., Esbensen, F.-A., & Weiher, A. W. (1991). Are there multiple paths to delinquency? *Journal of Criminal Law and Criminology*, 82, 83–118.
- Huisinga, D., & Jakob-Chien, C. (1998). The contemporaneous co-occurrence of serious and violent juvenile offending and other problem behaviors. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 47–67). Thousand Oaks, CA: Sage.
- Jessor, R., & Jessor, S. L. (1977). *Problem behavior and psycho-social development: A longitudinal study of youth*. New York: Cambridge University Press.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112, 527–535.
- Knoff, H. M. (2000). Organizational development and strategic planning for the millennium: A blueprint toward effective school discipline, safety, and crisis prevention. *Psychology in the Schools*, 37, 17–32.
- Knoff, H. M., & Batsche, G. M. (1995). Project ACHIEVE: Analyzing a school process for at-risk and underachieving students. *School Psychology Review*, 24, 579–603.
- Kortering, L., Braziel, P., & Tompkins, J. (2002). The challenge of school completion among youths with behavioral disorders: Another side of the story. *Behavioral Disorders*, 27, 142–154.
- Lewis, T. J., & Sugai, G. (1999). Effective behavior support: A systems approach to proactive school-wide management. *Focus on Exceptional Children*, 31(6), 1–24.
- Mather, C. (1975). *Corderius Americanus: A discourse on the good education of children* [Microfilm]. New Haven, CT: Research Publications. (Original work published 1828)
- May, S., Ard, W., III., Todd, A. W., Horner, R. H., Glasgow, A., Sugai, G., et al. (2002). *School-wide information system*. Eugene: University of Oregon, Educational and Community Supports.
- Mayer, G. R. (1995). Preventing antisocial behavior in the schools. *Journal of Applied Behavior Analysis*, 28, 467–478.
- Messick, S. (1988). The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 33–45). Hillsdale, NJ: Erlbaum.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessment. *Educational Researcher*, 23(2), 13–23.
- Metzler, C. W., Biglan, A., Rusby, J. C., & Sprague, J. R. (2001). Evaluation of a comprehensive behavior management program to improve school-wide positive behavior support. *Education and Treatment of Children*, 24, 448–479.
- Morgan-D'Atrio, C., Northup, J., LaFleur, L., & Spera, S. (1996). Toward prescriptive alternatives to suspensions: A preliminary evaluation. *Behavioral Disorders*, 21, 190–200.
- Morrison, G. M., Anthony, S., Storino, M., & Dillon, C. (2001). An examination of the disciplinary histories and the individual and educational characteristics of students who participate in an in-school suspension program. *Education and Treatment of Children*, 24, 276–293.
- Morrison, G. M., & Skiba, R. (2001). Predicting violence from school misbehavior: Promises and perils. *Psychology in the Schools*, 38, 173–184.
- Murdock, T. B., Anderman, L. H., & Hodge, S. A. (2000). Middle-grade predictors of students' motivation and behavior in high school. *Journal of Adolescent Research*, 15, 327–351.
- Najaka, S. S., Gottfredson, D. C., & Wilson, D. B. (2002). A meta-analytic inquiry into the relationship between selected risk factors and problem behavior. *Prevention Science*, 2, 257–271.
- Nakasato, J. (2000). Data-based decision making in Hawaii's behavior support effort. *Journal of Positive Behavior Interventions*, 2, 247–251.
- National School Safety Center. (1998). *Checklist of characteristics of youth who have caused school-associated violent deaths*. Westlake Village, CA: Author.
- Nelson, J. R., Benner, G. J., Reid, R. C., Epstein, M. H., & Currin, D. (2002). The convergent validity of office discipline referrals with the CBCL-TRF. *Journal of Emotional and Behavioral Disorders*, 10, 181–188.
- Nelson, J. R., Colvin, G., & Smith, D. J. (1996). The effects of setting clear standards on students' social behavior in common areas of the school. *The Journal of At-Risk Issues*, 3(1), 10–18.
- Nelson, J. R., Gonzales, J. E., Epstein, M. H., & Benner, G. J. (2003). Administrative discipline contacts: A review of the literature. *Behavioral Disorders*, 28, 249–281.
- Nelson, J. R., & Roberts, M. (2000). Ongoing reciprocal teacher–student interactions involving disruptive behavior in general education classrooms. *Journal of Emotional and Behavior Disorders*, 8, 27–37.
- Paine, S., & Paine, C. K. (2002). Promoting safety and success in school by developing students' strengths. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 89–108). Silver Spring, MD: National Association of School Psychologists.
- Patterson, G. R., & Bank, L. (1986). Bootstrapping your way in the nomological thicket. *Behavioral Assessment*, 8, 49–73.
- Quay, H. C., & Peterson, D. R. (1988). *Revised behavior problem checklist*. Miami, FL: University of Miami.
- Ryan-Arredondo, K., Renouf, K., Egyed, C., Doxey, M., Dobbins, M., Sanchez, S., et al. (2001). Threats of violence in schools: The Dallas Independent School District's response. *Psychology in the Schools*, 38, 185–196.
- Sadler, C. (2000). Effective behavior support implementation at the district level: Tigard-Tualatin School District. *Journal of Positive Behavior Interventions*, 2, 241–243.
- Scott, T. M. (2001). A schoolwide example of positive behavioral support. *Journal of Positive Behavior Interventions*, 3, 88–94.
- Skiba, R., Peterson, R., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and Treatment of Children*, 20, 295–315.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Smith, H., & Thomas, S. P. (2000). Violent and nonviolent girls: Contrasting perceptions of anger experiences, school, and relationships. *Issues in Mental Health Nursing*, 21, 547–575.
- Snell, J. L., Mackenzie, E. P., & Frey, K. S. (2002). Bullying prevention in elementary schools: The importance of adult leadership, peer group support, and student socio-emotional skills. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 351–372). Silver Spring, MD: National Association of School Psychologists.

- Sprague, J. R., Colvin, G., & Irvin, L. K. (1995). *The Oregon school safety survey*. Eugene: University of Oregon.
- Sprague, J. R., Sugai, G., Horner, R. H., & Walker, H. (1999). Using office discipline referral data to evaluate school-wide discipline and violence prevention interventions. *Oregon School Study Council*, 42(2), 1–18.
- Sprague, J. R., Walker, H. M., Sowards, S., Van Bloem, C., Eberhardt, P., & Marshall, B. (2002). Sources of vulnerability to school violence: Systems-level assessment and strategies to improve safety and climate. In G. Stoner, M. Shinn, & H. M. Walker (Eds.), *Interventions for achievement and behavior problems* (pp. 295–314). Silver Spring, MD: National Association of School Psychologists.
- Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child & Family Behavior Therapy* 24(1/2), 23–50.
- Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Nelson, C. M., Scott, T., et al. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions*, 2, 131–143.
- Sugai, G., Horner, R. H., & Gresham, F. (2002). Behaviorally effective school environments. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 315–350). Silver Spring, MD: National Association for School Psychologists.
- Sugai, G., Sprague, J. R., Horner, R. H., & Walker, H. M. (2000). Preventing school violence: The use of office discipline referrals to assess and monitor school-wide discipline interventions. *Journal of Emotional and Behavioral Disorders*, 8, 94–101.
- Tobin, T., & Irvin, L. K. (1996). The Olweus Bully/Victim Questionnaire: Evidence and consequences regarding use in the United States. *Reclaiming Children and Youth: Journal of Emotional and Behavioral Problems*, 5(1), 29–33.
- Tobin, T. J., Lewis-Palmer, T., & Sugai, G. (2002). School-wide and individualized effective behavior support: An explanation and an example. *Behavior Analysis Today*, 3(1), 51–75. Retrieved May 1, 2002, from <http://www.behavior-analyst-online.org/BAT/BAT31.pdf>
- Tobin, T. J., & Sugai, G. M. (1999a). Discipline problems, placements, and outcomes for students with serious emotional disturbance. *Behavioral Disorders*, 24, 109–121.
- Tobin, T. J., & Sugai, G. M. (1999b). Using sixth-grade school records to predict violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders*, 7, 40–53.
- Tobin, T., Sugai, G., & Colvin, G. (1996). Patterns in middle school discipline records. *Journal of Emotional and Behavioral Disorders*, 4, 82–94.
- Tobin, T., Sugai, G., & Colvin, G. (2000, May). Using discipline referrals to make decisions. *NASSP Bulletin*, 84(616), 106–117.
- Todd, A. W., Haugen, L., Anderson, K., & Spriggs, M. (2002). Teaching recess: Low-cost efforts producing effective results. *Journal of Positive Behavior Interventions*, 4, 46–52.
- U.S. Department of Education, National Center for Educational Statistics, National Forum on Educational Statistics. (2002). *Safety in numbers: Collecting and using incident data to make a difference in schools* (NCES 2002-312, prepared by I. Hantman, G. Bairu, A. Barwick, B. Smith, B. Mack, S. Meston, et al. of the Crime, Violence, and Discipline Task Force). Washington, DC: Authors. Retrieved February 5, 2003, from <http://nces.ed.gov/pubs2002/2002312.pdf>
- Walker, H. M., Block-Pedego, A., Todis, B., & Severson, H. (1991). *School archival records search (SARS): User's guide and technical manual*. Longmont, CO: Sopris West.
- Walker, H. M., Colvin, G., & Ramsey, E. (1995). *Antisocial behavior in school: Strategies and best practices*. Pacific Grove, CA: Brooks/Cole.
- Walker, H. M., & Shinn, M. R. (2002). Structuring school-based interventions to achieve integrated primary, secondary, and tertiary prevention goals for safe and effective schools. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 1–25). Silver Spring, MD: National Association of School Psychologists.
- Walker, H. M., Shinn, M. R., O'Neill, R. E., & Ramsey, E. (1987). A longitudinal assessment of the development of antisocial behavior in boys: Rationale, methodology, and first-year results. *Remedial and Special Education*, 8(4), 7–16.
- Walker, H. M., Stieber, S., & O'Neill, R. E. (1990). Middle school behavioral profiles of antisocial and at-risk control boys: Descriptive and predictive outcomes. *Exceptionality*, 1, 61–77.
- Walker, H. M., Stieber, S., Ramsey, E., & O'Neill, R. E. (1990). School behavioral profiles of arrested versus nonarrested adolescents. *Exceptionality*, 1, 249–265.
- Walker, H. M., Stieber, S., Ramsey, E., & O'Neill, R. E. (1991). Longitudinal prediction of the school achievement, adjustment, and delinquency of antisocial versus at-risk boys. *Remedial and Special Education*, 12(4), 43–51.
- Walker, H. M., Stieber, S., Ramsey, E., & O'Neill, R. E. (1993). Fifth grade school adjustment and later arrest rate: A longitudinal study of middle school antisocial boys. *Journal of Child and Family Studies*, 2, 295–315.
- Wright, J. A., & Dusek, J. B. (1998). Compiling base rates for disruptive behaviors from student discipline referral data. *School Psychology Review*, 27, 147–183.
- Zins, J. E., & Ponti, C. R. (1990). Best practices in school-based consultation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (Vol. 2, pp. 673–694). Washington, DC: National Association of School Psychologists.

Action Editor: Glen Dunlap

Copyright of Journal of Positive Behavior Interventions is the property of PRO-ED and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.