Handbook of School Violence and School Safety
International Research and Practice
Shane R. Jimerson, Amanda B. Nickerson, Matthew J. Mayer, Michael J. Furlong

Using Office Discipline Referrals and School Exclusion Data to Assess School Discipline

Publication details
Kent McIntosh, Emily S. Fisher, Kelly S. Kennedy, Calli B. Craft, Gale M. Morrison
Published online on: 05 Dec 2011

Using Office Discipline Referrals and School Exclusion Data to Assess School Discipline

Kent McIntosh
THE UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER

Emily S. Fisher
LOYOLA MARYMOUNT UNIVERSITY, LOS ANGELES

Kelly S. Kennedy
CHAPMAN UNIVERSITY, ORANGE, CA

Calli B. Craft
THE UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER

Gale M. Morrison
THE UNIVERSITY OF CALIFORNIA, SANTA BARBARA

Abstract

This chapter provides an overview of approaches to assessing problem behavior and school climate with existing discipline data, such as office discipline referrals and school exclusion records. Discipline data are potentially valuable in assessing school violence and school safety, including use for screening, identifying needs for change, and evaluating school-wide intervention outcomes. However, limitations to their reliability and validity must be considered, including disproportionate use with students from ethnic minority backgrounds. To address these concerns, many schools have taken steps to standardize office disciplinary procedures and enhance the cultural responsiveness of schoolwide discipline systems. This chapter will also provide recommendations for enhancing discipline data for decision making and implementing culturally responsive practices.

Given the broad range of discipline challenges facing today’s schools, coupled with budget-ary restrictions requiring school personnel to do more with less, discipline data can become
exceedingly valuable in assessing school violence and school safety. With the right measures, school teams can identify challenges, generate potential solutions, and evaluate effectiveness of the support provided. However, to be valuable, these measures must meet certain psychometric standards. For example, the data used need to be valid, reliable, efficient to collect, and useful for decision making (Horner, Sugai, & Todd, 2001).

School discipline indices, such as office discipline referrals, suspensions, and expulsions, provide data by which school personnel can maximize effectiveness and efficiency of decision making and the support provided to students. These data are appealing because of their ease of collection and ready availability—in most schools, collecting and summarizing discipline data is a district mandate. In addition, these data are sensitive to low-frequency, high intensity behaviors that are difficult to capture with direct observation (Sprague & Horner, 1999). As such, they often meet the efficiency and utility standards of measurement (McIntosh, Reinke, & Herman, 2009). Thus, assessing their adequacy for decision making, in comparison to resource intensive supplemental measures (e.g., direct observation, screening systems), is worthwhile. Unfortunately, school discipline data can suffer from poor reliability and, most concerning, disproportionate use with students from ethnic and racial minority backgrounds. Yet the steps to enhancing technical adequacy for making useful decisions with these data are known. This chapter addresses approaches to assessing problem behavior and school climate with existing discipline data, the benefits and limitations of using such data, and the particular challenge of disproportionate use. Finally, recommendations for enhancing the approaches for using these data to inform decision making will be provided.

**Office Discipline Referrals**

Office discipline referrals (ODRs) are forms used to document student problem behavior on school grounds. ODRs are typically completed by school personnel as a means of communicating the event to school administrators when a student is sent to the office and requires administrative involvement (i.e., additional support or a punitive consequence). Sugai, Sprague, Horner, and Walker (2000) have defined an ODR as:

> an event in which (a) a student engaged in a behavior that violated a rule or social norm in the school, (b) the problem behavior was observed or identified 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)

ODRs provide immediate appeal to school personnel because of their ready availability and potential use for a wide variety of decision making tasks. ODR data can be used to assess overall school climate, evaluate the effectiveness of schoolwide practices, analyze areas for additional intervention (e.g., locations, student competencies, groups of students), assess the need for professional development, and screen and monitor progress for individual students (Irvin et al., 2006; Irvin, Tobin, Sprague, Sugai, & Vincent, 2004; Luiselli, Putnam, Handler, & Feinberg, 2005). As an illustrative example, a school team may use ODR data to monitor the overall effectiveness of their schoolwide behavior support systems. These data can be used to note an increase in ODRs above rates in previous months, the same month of different years, and district averages. Team members may examine the ODR patterns further to identify that the increase in overall school incidents is due specifically to increases in physical aggression in the hallways. The team can use this specific information to provide additional support to students in a particular location (e.g., reteach expectations in the hallway) and particular social competency (e.g., conflict resolu-
Assessing School Discipline

Assessing School Discipline

307

tion) and identify professional development needs for school personnel (e.g., active supervision). Finally, ODR data can then be used to assess the effectiveness of these strategies on reducing aggression in the hallways.

Regardless of their utility, there are some potentially major threats to the reliability and validity of ODRs in assessing school discipline, making school-to-school and even within-school comparisons (across time) difficult. ODR data can be compromised by variations in use by school personnel (Kaufman et al., 2010; Morrison, Peterson, O’Farrell, & Redding, 2004; Rusby, Taylor, & Foster, 2007). Variability in ODR rates across classrooms and schools may not necessarily be due to differences in student misbehavior, but rather variations in instructional effectiveness (Scott, Nelson, & Liaupsin, 2001), classroom management skills (Blankemeyer, Flannery, & Vazsonyi, 2002; Reinke & Herman, 2002), and tolerance levels for student behavior (Gerber, 1988; Kaufman et al., 2010; Wright & Dusek, 1998). For example, ODRs may be more likely to be issued in classrooms where expectations are not explicit, behavioral management systems are inconsistent, and instruction time is less structured (Lo & Cartledge, 2007). In addition, the schoolwide discipline policy may be ambiguous, so teachers may not be informed as to what behaviors should result in ODRs.

Another limitation of using ODRs to assess school discipline is variation in recording procedures (Morrison et al., 2004). Many schools either do not use a formal office referral form or use such a form only for the most serious behaviors. Additionally, variation occurs in what is recorded on the form. For example, some teachers may indicate multiple types of misbehavior on the same form, and others may simply select the most pressing or severe behavior, resulting in varying data across ODRs.

Administrative response to ODRs may also affect how often teachers refer. For example, teachers may issue fewer ODRs if they feel that no additional support will be provided or anticipate that the ODR will reflect poorly on their teaching or management skills. Also, an overemphasis from administrators on reducing ODRs, rather than providing quality support, can lead to suppression, limiting their validity and utility for decision making (Kern & Manz, 2004).

To address these concerns and enhance the reliability and validity of ODR data, many schools have taken steps to standardize ODR forms and procedures. McIntosh and colleagues (2009) have offered five critical criteria for schools to employ:

(a) a common form that details important information about the incident (e.g., location, time of day, others involved), (b) clear definitions of what behaviors warrant a referral, (c) clear definitions of what behaviors are expected to be handled without a referral, (d) regular training on use and discrimination between reportable and nonreportable behaviors, and (e) a system for compiling and analyzing ODR data.

The School-Wide Information System (SWIS; May et al., 2008) is one example of an ODR system, used by over 7,500 schools, that includes a predefined set of ODR behaviors and definitions, plus requirements for staff training, data entry, and ODR forms (Educational and Community Supports, 2010). Data from ODRs can also be made more reliable when forms include check boxes (e.g., list of defined behavior offenses) instead of blank fields, so that forms are easier to use and data become more consistent among teachers (Wright & Dusek, 1998).

There is clear evidence that without such steps to enhance standardization, ODR data suffer from lack of validity (Nelson, Benner, Reid, Epstein, & Currin, 2002), but there is a growing research base demonstrating enhanced technical adequacy with standardized ODRs. The number of ODRs issued at a school is significantly related to teacher and student ratings of school climate, school engagement, classroom disruptiveness, substance use, and overall levels of
conflict at school and home (Irvin et al., 2004; Spaulding et al., 2010). Moreover, ODRs are sensitive to implementation of schoolwide interventions to reduce problem behavior and improve school climate (Bradshaw, Mitchell, & Leaf, 2010; Putnam, Luiselli, Handler, & Jefferson, 2003; Sprague et al., 2001).

There is also evidence for the reliability and validity of standardized ODRs for assessing individual student behavior. For example, there is evidence of the stability of ODR growth throughout the school year (McIntosh, Frank, & Spaulding, 2010), throughout elementary school (McIntosh, Horner, Chard, Boland, & Good, 2006), middle school (Tobin, Sugai, & Colvin, 1996), and from middle to high school (McIntosh, Flannery, Sugai, Braun, & Cochrane, 2008; Tobin & Sugai, 1999). The number of ODRs received is related to a range of negative individual outcomes, including school exclusion, academic underachievement, and family conflict (Irvin et al., 2004; Morrison, Anthony, Storino, & Dillon, 2001; Tobin & Sugai, 1999). ODRs have also been shown to be stronger predictors of teacher and parent rated problem behavior than family risk factors, such as poverty (Rusby et al., 2007).

Recent research has also examined the concurrent validity of ODRs with commonly used behavior rating scales. Walker, Cheney, Stage, and Blum (2005) found significantly higher ratings on the Problem Behavior Scale of the Social Skills Rating System (Gresham & Elliott, 1990) for students with multiple ODRs. In another study, McIntosh, Campbell and colleagues (2009) found significantly different ratings on the Externalizing Composite of the Behavior Assessment Scale for Children 2 (Reynolds & Kamphaus, 2004) for students with 0 to 1, 2 to 5, and 6 or more ODRs. Pas (2010) found moderate correlations between ODRs and scores on the Teacher Observation of Classroom Adaptation (Koth, Bradshaw, & Leaf, 2009). However, concurrent validity scores of unstandardized office referrals with the Child Behavior Checklist (Achenbach & Rescorla, 2001) were low (Nelson et al., 2002). With teacher training and standardized forms to increase reliability, ODRs can provide more accurate information to monitor the school climate, examine patterns of problem behavior, provide targets for school reform, and monitor progress toward school goals related to behavior (Irvin et al., 2006). Nevertheless, it is critical to note that without standardization and ongoing training, ODRs have questionable reliability and validity for decision making (Nelson, Gonzalez, Epstein, & Benner, 2003), and most schools in the United States are unlikely to be using standardized ODRs.

**School Exclusion**

School exclusion is an inclusive term to describe school discipline practices that remove students from the school environment, such as out of school suspension and expulsion. School exclusion often begins with an ODR, but then continues with an additional form that is completed to document the offense warranting exclusion and type and intensity of the consequence (e.g., number of days suspended, indication of a zero tolerance offense). Offenses leading to school exclusion are generally defined by state or provincial education codes and most often include offenses such as possession of a weapon, possession or selling of illegal drugs, physical harm to others, willful disobedience, destruction of property, disruption of school activities, and threat of harm (Kingery & Coggeshall, 2001).

According to the most recent U.S. statistics, 7% of all students were suspended from school at least once per school year, and 0.2% were expelled (Planty et al., 2009). Such data, when considering the offenses listed above, might be construed to indicate that schools are more dangerous than has been suggested. However, these statistics should be considered in light of the fact that it is becoming more common to use school exclusion (particularly suspension) for less severe incidents, such as interpersonal difficulties with peers or adults at school (Evenson, Justinger, Pelischek, & Schulz, 2009; Theriot, Craun, & Dupper, 2010). For example, 43% of
school exclusions in the United States were provided for insubordination, whereas less than 1% were due to possession of a firearm or explosive device (Dinkes, Kemp, & Baum, 2009).

There are two possible explanations for what many consider an overreliance on school exclusion. First, community pressure related to increased publicity for serious acts of violence has led to the adoption of “zero-tolerance” policies, which mandate school exclusion for a broad range of discipline offenses (American Psychological Association, 2008; Evenson et al., 2009). Though these policies were designed to both ensure student safety and promote equitable treatment by taking administrator discretion out of the decision making process, neither has occurred as a result of such policies (Skiba & Peterson, 2000).

The second explanation is that school administrators, lacking positive strategies for supporting students with challenging behavior, overuse school exclusion in an attempt to deter future incidents (Fenning & Rose, 2007). Osher, Bear, Sprague, and Doyle (2010) have described school exclusion as “a short-term fix to what often is a chronic and long-term problem” (p. 48). Removing students from school may reinforce problem behavior through escaping an undesirable situation and therefore may lead to increased problem behavior when returning to school (Lo & Cartledge, 2007). Unfortunately, a range of studies has shown that even when controlling for initial levels of behavior, the use of repeated school exclusion with individual students greatly enhances their risk for future mental health problems, antisocial behavior, academic failure, unemployment, homelessness, crime, and incarceration (Hemphill, Toumbourou, Herrenkohl, McMorris, & Catalano, 2006; Skiba & Peterson, 2000; Theriot et al., 2010; Wald & Losen, 2003).

Like ODRs, school exclusion data can potentially provide useful information for administrators. School exclusion data represent a measure of more severe problem behavior than ODRs because they often represent more serious, or at least more chronic, offenses (McIntosh, Campbell, et al., 2009). As a result, some view school exclusion data as a more valid index of school violence than ODRs because low-intensity problem behavior is not included. In addition, school exclusions may be viewed as less susceptible to underreporting. Because suspension provides immediate relief from having to manage a student’s behavior (Maag, 2001), school personnel may be more reluctant to suppress school exclusions than ODRs.

Comparing the rates of ODRs and school exclusions at individual schools can be helpful to assess school policies and examining needs for professional development (Morrison et al., 2004). For example, high relative rates of school exclusions may indicate high incidences of serious events, under reporting of ODRs, or a lack of proactive strategies to prevent problem behavior. Just as with ODRs, lack of standardization (e.g., inconsistency in counting the number of incidents or students involved in each incident) can hamper accurate assessment with school exclusion data (Leone, Mayer, Malmgren, & Meisel, 2000). To counter these challenges, it is worthwhile to examine both overall rates of suspension and percent of students suspended at least once during the year (Raffaele Mendez, Knoff, & Ferron, 2002).

**Disproportionality Issues in School Discipline**

There has been considerable research documenting overrepresentation of certain groups in school discipline data. Studies have consistently reported that students of color, students from low socioeconomic backgrounds, students in special education, and males are disproportionally represented in ODRs and school exclusions (American Psychological Association, 2008; Evenson et al., 2009; Kaufman et al., 2010; Raffaele Mendez & Knoff, 2003; Raffaele Mendez et al., 2002). In addition, particular attention has been paid to the exceedingly high number of African American students excluded from school (Fenning & Rose, 2007; Lo & Cartledge, 2007; Skiba, Michael, Nardo, & Peterson, 2002). This trend begins in elementary school and continues
through middle and high school (Bradshaw, Mitchell, O’Brennan, & Leaf, 2010; Raffaele Men-
dez & Knoff, 2003). In addition, recent research has found that though girls are excluded signifi-
cantly less than boys, African American girls are excluded at disproportionately higher rates than
Caucasian (Lo & Cartledge, 2007) and Hispanic girls (Losen & Skiba, 2010).

Given the connection between race and socioeconomic status in the United States, some
researchers have suggested that racial disproportionality in school exclusion is primarily due to
racial disproportionality in socioeconomic status, with socioeconomic status being more import-
ant to consider (Theriot et al., 2010). However, other researchers have convincingly demon-
strated that student ethnicity affects disproportionality above and beyond the contribution of
socioeconomic status (Skiba, Poloni-Staudinger, Simmons, Feggins, & Chung, 2005; Wallace,
Goodkind, Wallace, & Bachman, 2008). In these studies, ethnicity is a significant predictor of
ODRs and school exclusions, even after accounting for socioeconomic status and neighborhood
characteristics.

The reasons for this overrepresentation are multifaceted, but research highlights the need
to examine student-teacher interactions and strategies used by school personnel. For example,
African American males are more likely to receive ODRs and school exclusions for less severe
and subjective behaviors (e.g., disrespect, disruption), and these behaviors are more likely to
be handled through zero tolerance policies when exhibited by African Americans (American
Psychological Association, 2008; Theriot et al., 2010). In addition, when receiving ODRs for
the same offenses as students from other ethnicities, African American males are more likely to
receive more severe punitive consequences (Skiba et al., 2002). There is also evidence that school
exclusion rates are higher for teachers who lack adequate training in classroom management and
culturally competent practices (American Psychological Association, 2008; Raffaele Mendez et
al., 2002). In a review of ethnographic research, Fenning and Rose (2007) noted that teachers
and administrators were more likely to use ODRs and school exclusions preventively (i.e., to
avoid a perceived loss of classroom control) with students of color.

Recommendations for Enhancing the Use of Extant Data for Decision Making

Standardize ODRs to Improve Reliability and Validity

Given the information provided earlier in this chapter, relying on unstandardized ODRs for
decision making is ill advised. Completing a process of standardization can improve the technical
adequacy of ODRs, thus enhancing the quality of the decisions made from their analysis. How-
ever, standardization and ongoing training should not be expected to mitigate the challenges of
ODRs completely. ODRs should always be used as part of multi-method assessment, especially
when making high stakes decisions regarding individual students (McIntosh et al., 2010).

Use Discipline Data to Enhance Schoolwide Decision Making

Discipline data can be used effectively for schoolwide decision making, helping administrative
teams allocate resources and identify areas for reform (Irvin et al., 2006). The first step in this
process is to graph ODRs in useful ways (e.g., by rate, by problem behavior, location, by time of
day, by student) and share these data with staff (Clonan, McDougal, Clark, & Davison, 2007).
However, using discipline data for decision making is not always a simple task, and many teams
require specific training in the process of data based decision making. Newton and colleagues
(2009) have developed a training program called Team Initiated Problem Solving that guides
teams through the steps for effective decision making. An experimental study showed that teams
receiving the training used school discipline data significantly more effectively (Todd et al.,
2011).
One key point regarding discipline data is not to use it to punish or single out school personnel. Although data may identify specific teachers who are potentially struggling with classroom management, school administrators should not use ODRs alone to identify teachers in need of additional training. Otherwise, teachers may choose not to issue ODRs. Instead, a more comprehensive approach to assessing classroom management can be used, including direct observation and self-assessments (Simonsen, Fairbanks, Briesch, & Sugai, 2006). Using a range of tools to provide feedback to teachers about their use of specific classroom management practices has been shown to increase both teacher use of effective strategies and student academic engagement (Jeffrey, McCurdy, Ewing, & Polis, 2009). If all teachers receive regular feedback regarding classroom management, such assessments can become less threatening.

**Use Discipline Data to Identify Students for Additional Support**

Once effective schoolwide practices are in place, ODRs can help school personnel identify individual students who require additional behavior support to be successful (Rusby et al., 2007). Horner, Sugai, Todd, and Lewis-Palmer (2005) identified ODR cut points that signal the need for additional behavioral support as follows: 0 to 1 ODRs per year indicates that the student is adequately supported by schoolwide support, 2 to 5 ODRs indicate the need for efficient, targeted support, and 6 or more ODRs indicate the need for intensive support. These cut points have been validated by concurrent validity with behavior rating scales (McIntosh, Campbell, et al., 2009; Walker et al., 2005). Yet given the need to identify students and intervene proactively, additional research has found that any elementary student with 2 or more ODRs for any behavior by the end of October is highly likely to receive at least 6 ODRs over the course of the school year (McIntosh et al., 2010). Research in middle school suggests that receiving either 2 ODRs or a single ODR for harassment in the fall of Grade 6 predicts chronic problem behavior through middle school (Tobin et al., 1996). As a result, school teams can use ODRs at the end of October to screen students for additional intervention. However, it is important to note that additional sources of data should be used for screening, as ODRs are unlikely to identify a substantial portion of students with internalizing behavior problems (McIntosh, Campbell, et al., 2009).

**Use Data to Assess and Address Cultural Responsiveness**

ODR and school exclusion data can be used to assess the extent to which ODRs are used disproportionately with students of color, signaling a lack of cultural responsiveness (Townsend, 2000). Chinn and Hughes (1987) proposed a method for determining disproportionality in special education placement that could be applied to discipline data. Data may be disproportional if the percent of ODRs received by the group of interest is 10% or greater than their percent of the school enrollment. The discipline data application SWIS includes an ethnicity report that calculates the percent of students with ODRs by ethnicity and total ODRs by ethnicity to assess disproportionality (May et al., 2008). In addition, assessing patterns in ODRs can provide valuable information. For example, if students of color are receiving a large proportion of ODRs for a particular behavior (e.g., disrespect), behavior expectations or staff assumptions about behavior may need to be examined for cultural bias (Kaufman et al., 2010; Townsend, 2000).

School discipline data can then be used to assess the effectiveness of approaches to reducing disproportionate use. Schoolwide positive behavior support (Sprague & Horner, this volume) has been offered as an approach for promoting a safe, supportive school culture with shared, consistent expectations for behavior and reducing the use of ODRs and school exclusions. However, reductions in ODRs and school exclusions may not be equal across all student ethnicities, as schoolwide positive behavior support without attention to cultural responsiveness
may not reduce disproportionality (Kaufman et al., 2010; Skiba et al., 2008; Vincent & Tobin, in press). Several strategies are recommended to address these challenges. First, it is important for school personnel to self-reflect on their attitudes and understanding of students from diverse backgrounds (Hershfeldt et al., 2010; Skiba et al., 2008). Second, school expectations can be examined to assess whether they reflect the cultural norms of the diverse population of students in the school (Townsend, 2000). Third, it is important to examine student–teacher and student–administrator interactions for cultural responsiveness (Gregory, Skiba, & Noguera, 2010). When these practices are applied within a framework of schoolwide positive behavior support, disproportionality can be reduced considerably (e.g., Cregor, Smith, & Leverson, 2010; Jones, Caravaca, Cizek, Horner, & Vincent, 2006).

Conclusion

As this chapter has highlighted, school discipline data can be useful for decision making, but they are not without limitations. Educators must use caution and careful judgment when making decisions about school discipline and related factors, given the threats to reliability and validity that are often present in unstandardized ODR and school exclusion data. These data can become significantly more valid and useful if standardized, with ongoing training in both accurate administration and use for decision making. With these changes, extant discipline data can provide an efficient means to identify students and classrooms in need of additional support, measure school needs and response to intervention, and assess disproportionate use.

References


Table 23.1 Implications for Practice: Use of Extant Data for Decision Making

1. Repeated use of ODRs and school exclusions increases student risk for delinquency and school dropout.
2. ODRs can be made more valid and reliable through standardization, including using a standardized form, clarification on use, regular entry and analysis, and ongoing training.
3. Standardized ODRs can be used for a range of decisions, including screening, problem analysis, and evaluation of practices.
4. School teams can benefit from training and support in the use of discipline data for decision making.
5. Students of color receive a disproportionate amount of both ODRs and school exclusions, even when socioeconomic status and community risk factors are considered.
6. Discipline data can be used to assess the cultural responsiveness of schoolwide behavior support systems.
7. The cultural responsiveness of school-wide behavior systems can be enhanced through reflecting on attitudes and biases and changing expectations and practices to match the backgrounds and needs of students.


---

Kent McIntosh et al.


