School Psychology Review,
2015, Volume 44, No. 2, pp. 224–240

Relationship Between School Suspension and Student Outcomes: A Meta-Analysis

Amity L. Noltemeyer and Rose Marie Ward
Miami University

Caven Mcloughlin
Kent State University

Abstract. Although the association between school suspension and deleterious outcomes is widely acknowledged, policy and practice need to be informed by an evidence base derived from multiple studies revealing consistent trends. This meta-analysis aims to address this void by examining the degree to which different types of school suspensions (in-school versus out-of-school) are associated with both academic achievement and school dropout, while concurrently examining study or participant characteristics that moderate these relationships. Data sources included peer-reviewed and non-peer-reviewed studies from 1986–2012 obtained via bibliographic databases. A meta-analysis was conducted on 53 cases from 34 studies. The results revealed a significant inverse relationship between suspensions and achievement, along with a significant positive relationship between suspensions and dropout. Furthermore, study or participant characteristics and type of suspension significantly affected the relationship between suspensions and the outcome variables. Implications for policy, practice, and research are emphasized.

Exclusion from school is by no means a new educational practice. Although it has been used in various forms throughout the history of U.S. schooling, the use of suspensions as a deterrent for misbehavior increased widely beginning in the 1970s (Losen & Skiba, 2010). Whereas the term *expulsion* refers to the more permanent removal of a student from the school by the superintendent, the term *suspension* generally refers to the denial of school attendance for a specific amount of time that may be 10 days or less. Although *suspension* most often refers to out-of-school suspension (OSS), there also has been an increasing trend toward in-school suspension (ISS). ISS often involves the student being removed to a separate classroom for at least a full day, where he or she must complete work and cannot participate in mainstream activities alongside peers (Hyman, 1997).

The rationale undergirding both OSS and ISS use is that these practices will serve as punishment, decreasing the likelihood of future negative behaviors. However, some students may actually find the conditions of school more punishing than removal (Hyman, 1997). Although there is evidence of negative outcomes associated with suspension, it is un-

Please address correspondence regarding this article to Amity L. Noltemeyer, Miami University, 210 E Spring St, Oxford, OH 45056; e-mail: anoltemeyer@miamioh.edu

known how much they affect academic achievement and school dropout. In addition, the relative impact of study design and participant characteristics on the relationship between school suspension and the outcomes is unclear. This study aims to use meta-analytic techniques to fill these gaps in the extant literature. Consistent with the call for more research on the connections between school discipline data and important student outcomes (e.g., Losen, 2011), the relationship between school suspensions and achievement and dropout and the moderators of those relationships are important areas for research that could lead to a more accurate and comprehensive understanding of the issues and reveal implications for policy and practice.

SCHOOL SUSPENSIONS

In the 2011–2012 school year, 3.5 million U.S. students received ISS and 3.45 million received OSS (see U.S. Department of Education Office of Civil Rights, 2014). This widespread use of suspension has been disproportionately applied to some groups of students. For instance, Black students are significantly overrepresented as recipients of school suspension (e.g., Shirley & Cornell, 2011; Skiba et al., 2011; Sullivan, Klingbeil, & Van Norman, 2013). Male students (e.g., Raffaele Mendez & Knoff, 2003; Sullivan et al., 2013) and economically disadvantaged students (e.g., Sullivan et al., 2013; Wu, Pink, Crain, & Moles, 1982) are at a similarly heightened risk. Some research suggests that school characteristics also predict higher rates of suspension. Urban schools (e.g., Noltemeyer & Mcloughlin, 2010) and schools with a high percentage of low-income students (e.g., Christle, Nelson, & Jolivette, 2004; Fowler & Walberg, 1991) have higher suspension rates.

The extant literature is replete with commentary on the ineffective nature of suspensions. Sharkey and Fenning (2012) established a rationale for prevention-oriented school discipline contexts, citing research showing that suspension not only is ineffective at producing positive behavioral change but also is linked with subsequent academic failure and school dropout. Moreover, Chin, Dowdy, Jimerson, and Rime (2012) called for alternatives to suspension because suspension may result in counterproductive outcomes rather than students learning from their actions. However, despite repeated calls for more proactive interventions to address misbehavior within a prevention-science framework, the actual amount of quantitative research investigating the link between suspensions and detrimental outcomes, especially academic outcomes, has been relatively sparse.

SUSPENSION AND ACADEMIC ACHIEVEMENT

Academic achievement has often been discussed within the suspension dialogue because school suspensions are inversely related to academic achievement for individual students and for broader systems. Schools with high suspension use have shown lower mean scores on state achievement tests than schools with lower suspension use (Rausch & Skiba, 2004). Furthermore, even when poverty and Black student enrollment were controlled for, a school's suspension rate remained a significant predictor of its passage rate on a state achievement test in both elementary and secondary schools (Skiba & Rausch, 2004). There were negative correlations between suspension in Grade 6 and math or reading achievement in Grades 7 and 8 for both White and Black students (Raffaele Mendez, 2003), and students made significantly less achievement gains the more they were suspended across a 3-year period (Arcia, 2006).

There are several possible explanatory mechanisms for the relationship between suspension and academic achievement. The links may be explained by missed instructional time, school disengagement resulting from the suspension, preexisting student academic or behavioral difficulties that resulted in the suspension and concurrently influenced achievement (Arcia, 2006), or additional opportunities to interact with delinquent peers when out of school (Gifford-Smith, Dodge, Dishion, & McCord, 2005). Establishing the degree to which the association between suspension and

academic achievement exists across the literature is an important first step toward better understanding these explanations.

SUSPENSION AND SCHOOL DROPOUT

The high school dropout rate is another variable commonly examined in relation to suspension. The dropout rate is the percentage of students who leave high school without a successful outcome such as passing the General Educational Development test or receiving a high school diploma (Chapman, Laird, & KewalRamani, 2010), which was the case for 8.0% of individuals in the United States aged 16–24 years in 2008, and these figures were even higher for male students and Hispanic students (Chapman et al., 2010). Students in special education and students with chronic attendance or achievement problems also experience a heightened risk for dropout (Mac Iver, 2011). Failure to complete high school is problematic, considering that early school departure has been associated with an increased risk of such negative outcomes as fewer job opportunities and lower earning potential (Northeastern University–Center for Labor Market Studies, 2009), incarceration (Lochner & Moretti, 2002), reliance on public economic assistance (Waldfogel, Garfinkel, & Kelly, 2007), and substance use (Townsend, Flisher, & King, 2007).

Given the negative consequences, it is important to identify factors that might contribute to the decision to drop out of school prematurely. Several risk factors have been speculated, including retention in a grade (e.g., Carpenter & Ramirez, 2007), low socioeconomic status (e.g., Suh & Suh, 2007), and academic failure (e.g., Suh & Suh, 2007). In this investigation, we are concerned with the role of school suspensions as a contributor to premature departure from school. Some studies have found that suspensions are positively associated with high school dropout rates. For example, suspension rates in 196 Kentucky schools were significantly correlated with dropout rates (Christle, Jolivette, & Nelson, 2007). At the individual-student level, a study using data from 30,000 high school sophomores found that students who dropped out of school were more likely to have a history of suspension (Eckstrom, Goertz, Pollack, & Rock, 1986). Similar relationships have been documented for diverse student groups, including English language learners (Kim, 2011).

It is plausible that the relationships between these variables could be explained by demographic factors (e.g., rather than suspensions causing students to drop out, it could be that economically disadvantaged students, for example, may be both more likely to be suspended and more likely to drop out). Consequently, some studies have controlled for such factors. For example, Lee, Cornell, Gregory, and Fan (2011) found that even when controlling for student demographic variables and attitudes, the relationship between suspensions and dropout remained significant. These findings suggest that the act of suspension itself, rather than only the attitudinal or demographic variables predicting or related to it, may directly contribute to the decision to drop out. Once again, however, little is known about the extent of this relationship across the existing literature.

RATIONALE AND PURPOSE

Many individual studies have examined the relationship between suspension and achievement, as well as between suspension and dropout. However, research is needed to reconcile and summarize the existing literature in an effort to elucidate the magnitude and characteristics of the relationship between suspension and these outcome variables. Specifically, we know that suspension is related to the outcomes, but we cannot yet confidently assert the strength of this relationship across all studies that have been conducted. In addition, little is known about the moderating effects of race or ethnicity, gender, type of suspension (OSS or ISS), socioeconomic status, and publication type. Finally, a comprehensive study examining the characteristics of empirical research on this topic, which could reveal areas in need of further research, is missing

from the literature on this topic. This meta-analysis aims to address each of these gaps in the research on school suspensions. Specifically, the research questions are as follows: (a) What is the quantitative relationship between school suspension and achievement, as well as between school suspension and dropout, across existing studies? (b) Are there moderating effects of race or ethnicity, gender, type of suspension, socioeconomic status, and publication type? (c) What are the characteristics of the existing empirical research on this topic?

METHOD

This study used meta-analytic methods. Meta-analysis is a technique that integrates findings across multiple studies to reveal patterns that exist across the research literature base (Hunter & Schmidt, 2004). Meta-analysis has multiple advantages. For example, it has the potential to correct for sampling error, measurement error, distorting effects, and other artifacts that result in contradictory findings across studies (Hunter & Schmidt, 2004). Although meta-analysis was originally used to compare the effect of intervention and control conditions, it also can be used to examine the strength of relationships between variables.

Search Procedures

Potential publications were researched using keywords in the following electronic databases: ERIC, Education Full Text, Educational Research Complete, PsycINFO, Psychological and Behavioral Sciences Collection, and Social Science Citation Index. For each keyword search, the term "suspen*" was combined with one of the following terms: "outcome*," "graduat*," "dropout," "achievement," "school completion," "school persistence," "test," and "grades." Using asterisks in the search criteria allowed for any variation of the term to also emerge from the search (e.g., "graduat*" would produce "graduate," "graduated," and "graduation"). After identification of an initial pool of articles using these keyword searches, the reference section of each article was consulted to locate additional articles. Finally, after the complete list of articles was assembled, two experts in the field were consulted to ensure that no relevant articles were missed.

We devised specific a priori inclusion criteria to determine which studies would be included in the meta-analysis (see Table 1). In doing so, we drew from a framework adapted from Lipsey and Wilson (2001) to define which studies would and would not be included, in an effort to delimit the included studies to those that best aligned with the goals of our research and measurement best practices. Each study could provide multiple cases or independent samples that were eligible for inclusion.

Each article found through any of these three methods (i.e., keyword search, reference list search, and expert review) was first screened based on the title, the abstract, and quick perusal of the article. Any articles that showed potential for meeting the inclusion criteria were then read closely to determine if they did in fact meet the criteria. Articles that were not accessible through any of our library resources were requested through other libraries or sources or by contacting the authors.

Coding Scheme

A three-step process was used to develop and apply the coding system. First, we constructed an initial coding scheme and manual based on the guidelines of Robey and Dalebout (1998). As recommended by Quintana and Minami (2006), coding variables were defined precisely and objectively to avoid coding bias. These variables captured salient aspects of the research related to the variables of interest, such as author names; year of publication; type of publication; page numbers; journal or publication name; design; number of participants; dependent variables of interest; information related to the actual effect size (ES) definitions and calculations; race, gender, and socioeconomic makeup of participants; mean grade and age of participants; level of analysis (e.g., student, school, district, state, national, or international); level at which the dependent variables were measured (e.g., school or individual-student level); and how each article was found. Operational

Table 1. Eligibility Criteria Across Seven Dimensions

Dimension	Criteria for Eligibility
Distinguishing features	Must document the relationship between school suspension (in-school or out-of-school) and at least one of the outcome variables.
Research respondents	Must involve participants in preschool through Grade 12 at the time data were collected.
Key variables	Must report at least one quantitative measure of suspension rate (in-school or out-of-school) and at least one quantitative measure of one of the outcome variables; must also be possible to obtain a correlation between suspensions and at least one of the outcome variables.
Research methods	Research designs must allow for a measure of association between variables and must not have egregious methodologic flaws.
Cultural and linguistic range	Must be written in English.
Time frame	Must be conducted since 1970 because the use of suspensions increased widely beginning in the 1970s (Losen & Skiba, 2010).
Publication type	Could be published or unpublished material. This decision was made to avoid the potential of upward bias because published studies tend to report more significance (Lipsey & Wilson, 2001).

Note. The dimensions were proposed by Lipsey and Wilson (2001).

definitions and response options were written for each of these variables and are available from the first author on request.

Next, we applied the coding scheme to several articles to reveal any weaknesses or flaws and allow an opportunity for revision of the coding scheme (Robey & Dalebout, 1998). Finally, after the coding scheme was adjusted accordingly, two graduate students in school psychology and the first author coded the studies. The graduate students were trained until 100% interrater agreement was obtained. These two students coded only aspects of the study unrelated to the actual ES because Quintana and Minami (2006) recommended that coding be done blind to the ESs of the articles. If a single study included multiple analyses that were conducted on different populations, these were coded as two separate cases (Lipsey & Wilson, 2001). For example, if one analysis used a sample of sixth-grade White students and a second analysis in the same study used a sample of eighth-grade Latino students, these were coded as separate entries. It should also be noted that the total number of participants was estimated in eight cases using approximations calculated from data provided within the study. For example, in one study the exact total number of students attending the participating schools was not provided; however, we computed it by multiplying the mean enrollment of the participating schools by the number of participating schools.

Approximately 25% of the cases were double coded by the first author, and discrepancies were resolved via consensus. Interrater reliability for categorical variables was calculated using Cohen's κ, and all values were statistically significant at $p < .01$. The percent agreement for these variables ranged from 93% to 100% ($M = 99.1\%$, $Mdn = 100\%$). Interrater reliability for continuous variables was calculated using Pearson's *r,* and these values ranged from $r = 0.97$ to $r = 1.0$. For these variables, the percent agreement ranged from 75% to 100% ($M = 88.9\%$, $Mdn = 91.7\%$), with the 75% agreement occurring because of the age variable, for which there were few data points and minor inconsistencies (e.g., one rater said the mean age was 14.25 years and the other said 14.00 years). Together, these interrater agreement

results suggest a high degree of agreement between raters.

ES Computation

We decided to convert all findings into *r* values, considering that this is a common ES measure used to assess the relationship between two variables (Wolf, 1986) and most of the studies in our meta-analysis examined the relationships between variables (Quintana & Minami, 2006). The metric *r* was selected over r^2 because the latter can be misleading and obscure the direction of the effect (Hunter & Schmidt, 2004). In some cases, data contained in the publications were reported directly as *r* values. In other cases, data were reported using other methods and were converted to *r* values using an ES calculator provided by Lipsey and Wilson (2001). For each publication, we coded how the *r* value was obtained (i.e., directly from the source or through conversion). In instances in which multiple *r* values were obtained for the same case and the same outcome variable (e.g., math and reading achievement scores were analyzed using the same population), these *r* values were averaged together into a single *r* value for that outcome variable. Regarding the dropout variable, it should be noted that for some studies, the ES represented the dropout rate, whereas for others, the ES represented the graduation rate. In the latter circumstances, the inverse ES value was used to approximate the dropout rate.

Analyses

Several types of analyses were conducted on the data. First, descriptive analyses (e.g., frequency counts, means) were calculated on the coding categories to better understand the topography of the extant literature focusing on the relationship between suspensions and outcomes. Next, ES values were aggregated and analyzed to quantify the relationship between suspensions and each of the outcome variables. Finally, analyses were conducted to detect moderator effects, which involved examining the linear relationship between the ES and the moderator (Shadish & Sweeney, 1991). The *Q* statistic was used to examine the heterogeneity of the true ESs. A significant *Q* statistic indicated that there was variability in the ESs or there might have been subgroups present in the pool of ESs. A weighted analysis of variance was used to test these moderating effects (see Quintana & Minami, 2006) using Comprehensive Meta-Analysis software (Version 2; Borenstein, Hedges, Higgins, & Rothstein, 2007).

RESULTS

In all, 53 cases from 34 studies met all the inclusion criteria and were used in the analyses. Of these, 42 cases from 24 studies examined the relationship between suspension and achievement outcomes and 11 cases from 10 studies examined the relationship between suspension and school completion outcomes. See Table 2 for descriptive statistics on the characteristics of these cases. On average, the cases included over 7,000 participants and over 100 schools, with the mean student age being 13 years and the mean student grade being the eighth grade. The majority of cases (62.3%) focused on OSS, measured suspensions at the student level (58.5%), and were conducted with U.S. populations (96.2%; see Table 2). Many of the included cases were in the form of dissertations (47.2%) and peer-reviewed articles (34.0%), with other publication types being less frequently observed (see Table 2). Finally, the disciplines of the primary authors varied, with half represented by the field of educational administration (50.9%; see Table 2). The studies were published between 1986 and 2012.

Publication Bias

To account for the file-drawer effect or publication bias, we included unpublished and published studies. Publication bias was assessed using Egger's regression test (Egger, Smith, Schneider, & Minder, 1997) of the intercept. A nonsignificant test indicated that the studies in the dropout analysis did not seem to have a publication bias, $t(10) = 1.82$, $p = .10$. In addition, the studies in the achievement analysis did not seem to have a publication bias, $t(38) = 0.95$, $p = .35$.

Table 2. Characteristics of 53 Cases Included in Meta-Analysis

	Sample Characteristics						
	n	*M*	Minimum	Maximum	*SD*	Frequency	%
Number of participants	45	7,579.13	15	71,342	1,140.84		
Number of participating schools	24	134.46	1	1,336	295.59		
Grade of sample	34	8.81	2.5	12	1.90		
Age of sample	11	13.75	9.38	16	1.74		
Suspension types represented							
Out of school						33	62.3
In school						12	22.6
Out of school and in school combined						1	1.9
Not specified						7	13.2
Publication forms							
Peer-reviewed journal						18	34.0
Professional newsletter						2	3.8
Technical report						4	7.5
Dissertation						25	47.2
Conference publication						4	7.5
Level of analysis							
National						4	7.5
State						25	47.2
District						16	30.2
School						8	15.1
Level at which suspension was measured							
Student level						31	58.5
School level or greater						22	41.5
Nationalities represented							
United States						51	96.2
Australia						2	3.8

Estimates of ES for ISS and OSS

In addition to the achievement and dropout rate bifurcation of studies, ESs were calculated based on the type of suspension examined (see Table 3). Twelve studies (achievement, $n = 11$; dropout, $n = 1$) examined ISS. The ISS–achievement studies had a significant Q statistic, $Q(10) = 78.89$, $p < .001$, indicating that the included ESs (or r values) were heterogeneous. The estimated ES from the random-effects model was −0.10, 95% CI [−0.13, −0.07], $SE = 0.001$, and was statistically significant, $z = -6.14$, $p < .001$. A significant negative correlation indicates that there is an inverse relationship between achievement scores and rates of ISS. Schools that tend to report higher levels on achievement variables also tend to report lower levels of ISS. There was only one study that examined the relationship between dropout rate and ISS, and it is included in the overall estimate of ES between suspension and dropout rate.

Thirty-four studies (achievement, $n = 25$; dropout, $n = 9$) examined OSS. The OSS–achievement studies had a significant Q statistic, $Q(24) = 16{,}079.26$, $p < .001$, indicating that the effects were heterogeneous or not from the same population. Therefore, we used a random-effects model. The estimated ES was −0.24, 95% CI [−0.29, −0.18], $SE = 0.02$, and was statistically significant, $z = -8.06$, $p < .001$. Similar to the ISS, there was a significant inverse relationship between OSS and the achievement variables. This re-

Table 3. Meta-Analytic Results for Achievement and Dropout by Suspension Type

					95% CI		
	k	*Q*	Effect Size (*r*)	*SE*	*LL*	*UL*	*z*
Achievement tests							
In-school suspension	11	78.89*	–.10	.001	–.13	–.07	6.14*
Out-of-school suspension	25	16,079.26*	–.24	.02	–.29	–.18	8.06*
Combined[a]	2	0.35	–.13	.01	–.15	–.11	–13.33*
Unknown[b]	5	21.30*	–.32	.01	–.40	–.22	–6.43*
All types of suspension	43	17,337.13*	–.21	.02	–.26	–.17	–9.47*
Dropout							
Out-of-school suspension	9	281.56*	.25	.01	.18	.32	6.90*
All types of suspension	12	316.03*	.28	.01	.22	.33	8.81*

Note. CI = confidence interval; *LL* = lower limit; *UL* = upper limit.
[a]Studies examined both in- and out-of-school suspension simultaneously.
[b]Studies did not specify the type of suspension examined.
*$p < .001$.

lationship indicates that as rates of OSS increase, lower scores on achievement variables are reported. The OSS–dropout studies had a significant Q statistic, $Q(8) = 281.56$, $p < .001$. By use of a random-effects model, the estimated ES was 0.25, 95% CI [0.18, 0.32], $SE = 0.01$, and was statistically significant, $z = 6.90$, $p < .001$. Studies with higher dropout rates also tend to report higher levels of OSS rates.

Estimates of ES for Combined Suspension Types

Three studies (achievement, $n = 2$; dropout, $n = 1$) explored OSS and ISS together. The combined suspension–achievement studies had a nonsignificant Q statistic, $Q(1) = 0.35$, $p = .55$, indicating a homogeneous sample. The estimated ES using a fixed-effects model was −0.13, 95% CI [−0.15, −0.11], $SE = 0.01$, and was statistically significant, $z = -13.33$, $p < .001$. In studies that examined both ISS and OSS combined and achievement variables, as suspension rates increased, achievement variables decreased. There was only one study examining both types of suspension and the dropout rate. This study is included in the overall estimate of ES between suspension (both ISS and OSS combined) and dropout rate.

Estimates of ES for Studies in Which Type of Suspension Was Not Reported

Six studies (achievement, $n = 5$; dropout, $n = 1$) did not report what type of suspension was examined. The five achievement studies had a significant Q statistic, $Q(4) = 21.30$, $p < .001$. Therefore, we used a random-effects model for the estimated ES and found −0.32, 95% CI [−0.40, −0.22], $SE = 0.01$, $z = -6.43$, $p < .001$. Similar to the findings for the ISS, OSS, and combined studies, there was a significant negative correlation between rates of suspension and achievement variables. Studies with higher levels of suspension tended to report lower levels of achievement variables. Because there was only one dropout study that did not specify the suspension type, this study was included in the overall estimate of effect (both ISS and OSS combined and dropout rate).

Examination of Moderators

Additional analyses examined the role of moderators with respect to the estimated

Table 4. Moderation Meta-Analytic Results for Achievement

	Q_{within}	Effect Size (r)	SE	95% CI		k	z
				LL	UL		
Achievement overall	—	–.21	.02	–.26	–.17	43	–9.47*
Gender	17,285.08*						
<5% male	0.00	–.26	.00	–.41	–.10	1	–3.16*
5%–50% male	18.05*	–.31	.01	–.35	–.27	4	–15.42*
50% male	8.54	–.25	.03	–.35	–.15	4	–4.74*
50%–95% male	4,886.80*	–.25	.03	–.35	–.15	23	–4.74*
>95% male	0.00	–.27	.00	–.39	–.14	1	–3.97*
Cannot determine	12,371.69*	–.22	.02	–.22	–.22	10	–275.95*
Race	15,908.52*						
>60% White	3,221.29*	–.28	.03	–.29	–.27	11	–85.69*
>60% Black	148.58*	–.12	.004	–.13	–.10	8	–19.55*
>60% Hispanic	82.94*	–.11	.002	–.11	–.10	4	–23.88*
>60% other minority	0.00	–.35	.00	–.51	–.17	1	–3.65*
Mixed, none >60%	108.89*	–.24	.01	–.25	–.22	11	–36.58*
Cannot determine	12,346.82*	–.22	.03	–.22	–.22	7	–275.13*
SES	15,351.97*						
76%–100% low SES	0.00	–.11	.00	–.19	–.03	1	–2.75*
51%–75% low SES	113.80*	–.24	.01	–.25	–.22	10	–37.09*
26%–50% low SES	1,413.04*	–.42	.13	–.43	–.41	2	–87.64*
0%–25% low SES	0.00	–.27	.00	–.32	–.22	1	–10.11*
Unspecified	13,825.13*	–.22	.03	–.22	–.21	29	–277.19*
Publication type	2,274.85*						
Peer-reviewed journal	197.82*	–.24	.01	–.25	–.23	14	–44.24*
Professional newsletter	0.00	–.16	.00	–.16	–.16	1	–160.52*
Technical report	1,720.26*	–.35	.03	–.35	–.35	4	–263.42*
Dissertation	331.28*	–.12	.001	–.12	–.13	21	–43.28*
Conference publication	25.50*	–.24	.16	–.24	–.29	3	–9.42*
Level of analysis	17,329.17*						
National	0.00	–.24	.00	–.26	–.22	1	–18.84*
State	17,033.18*	–.22	.02	–.22	–.22	21	–285.77*
District	265.72*	–.21	.01	–.22	–.20	14	–43.29*
School	30.27*	–.23	.01	–.27	–.19	6	–11.40*

Note. Q statistic tests were performed to assess homogeneity of variance. CI = confidence interval; LL = lower limit; UL = upper limit; SES = socioeconomic status.
*$p < .01$.

ESs. Because of limited numbers of studies for each suspension type and the ambiguity in some of the studies concerning the type of suspension examined, ES was estimated across all types of suspension for achievement and dropout rate overall for the moderation analyses. See Tables 4 and 5 for the moderation results for achievement and dropout, respectively. Moderator analyses were not conducted for age or grade level because of missing data (i.e., too many studies did not include this information).

The effect of suspension on academic outcomes was evaluated with respect to gender (i.e., percent male), socioeconomic status, and racial composition of the sample; publication type; and level of analysis. The overall ES varied based on the gender composition of

Table 5. Moderation Meta-Analytic Results for Dropout

	Q_{within}	Effect Size (r)	SE	95% CI		k	z
				LL	UL		
Dropout overall	316.03*	.28	.01	.22	.33	12	8.81*
Gender	240.25*						
5%–50% male	0.00	.04	.00	–.08	.16	1	0.67
50% male	0.59	.31	.004	.29	.34	2	21.71*
50%–95% male	129.13*	.20	.03	.18	.22	4	23.74*
Cannot determine	110.54*	.27	.01	.25	.28	5	37.99*
Race	154.13*						
>60% White	122.46*	.19	.01	.17	.20	5	25.47*
>60% Black	0.00	.25	.00	.22	.28	1	16.26*
>60% Hispanic	0.00	.29	.00	–.26	.70	1	1.03
Mixed, none >60%	20.23*	.32	.05	.30	.33	2	33.84*
Cannot determine	11.44*	.32	.06	.29	.34	2	22.06*
SES	283.53*						
51%–75% low SES	0.00	.31	.00	.28	.34	1	21.08*
26%–50% low SES	42.97*	.32	.09	.26	.39	3	9.19*
0%–25% low SES	0.00	.17	.00	.10	.24	1	4.34*
Unspecified	240.56*	.24	.01	.23	.25	7	43.45*
Publication type	293.20*						
Peer-reviewed journal	89.06*	.25	.01	.24	.27	5	29.20
Professional newsletter	0.00	.54	.00	.42	.64	1	7.33*
Dissertation	204.15*	.24	.01	.23	.26	5	38.74*
Conference publication	0.00	.17	.00	.09	.24	1	4.34*
Level of analysis	280.92*						
National	153.60*	.23	.01	.21	.24	3	30.47*
State	52.15*	.25	.05	.20	.30	4	36.90*
District	75.17*	.27	.01	.25	.28	4	36.90*
School	0.00	.54	.00	.42	.64	1	7.33*

Note. Q statistic tests were performed to assess homogeneity of variance. CI = confidence interval; LL = lower limit; UL = upper limit; SES = socioeconomic status.
*$p < .01$.

the sample, $Q_{between}(5) = 52.05$, $p < .001$, with the largest ES emerging in studies with 5% to 50% male participants and the weakest ES emerging in studies in which the gender composition could not be determined. In addition, the overall ES for race varied across groups, $Q_{between}(6) = 1{,}428.61$, $p < .001$. Specifically, the strongest association was found for the study using a sample comprising more than 60% ethnic minority participants, and the weakest association was found for the studies using samples comprising more than 60% Hispanic participants. The overall ES for socioeconomic status varied across groups, $Q_{between}(4) = 1{,}985.16$, $p < .001$, with the ES being strongest in studies in which the populations were 26% to 50% low socioeconomic status and weakest in the study in which the population was more than 75% low socioeconomic status. In addition, the overall ES was statistically significant across the different publication types, $Q_{between}(4) = 15{,}062.28$, $p < .001$, with technical reports providing the strongest effects and dissertations providing the weakest effects. Finally, statistical significance also emerged when we considered the

level of analysis, $Q_{between}(3) = 7.96$, $p = .05$, with national studies evidencing the strongest ESs and district-level studies evidencing the weakest ESs. In sum, the measure of ES varied across all of the moderators for academic achievement. Therefore, these moderators explain some of the variation across the studies.

The effect of suspension on dropout was evaluated with respect to gender, socioeconomic status, suspension type, publication type, and level of analysis. See Table 5 for the estimated ESs by group. The overall ES varied based on the gender composition of the sample, $Q_{between}(3) = 75.78$, $p < .001$, with the largest ES emerging in studies using samples with 50% male students and the smallest ES emerging in the study using a sample with 5% to 50% male students. In addition, the overall ES for race varied across groups, $Q_{between}(5) = 161.90$, $p < .001$. The largest ESs were observed in studies using samples that were racially mixed and those in which the racial composition could not be determined; conversely, the smallest ESs were observed in studies using samples that were more than 60% White. The overall ES for socioeconomic status varied across groups, $Q_{between}(3) = 32.50$, $p < .001$. Although the strongest ES was found in the 26% to 50% low–socioeconomic status population and the weakest effects were found in the 0% to 25% low–socioeconomic status population, these results should be interpreted with caution because socioeconomic status was unspecified in a number of studies. The overall ES was also statistically significant across the different publication types, $Q_{between}(3) = 22.83$, $p < .001$, with the highest ES emerging in a professional newsletter and the lowest ES emerging in a conference publication. Finally, significance was found regarding the level of analysis, $Q_{between}(3) = 35.12$, $p < .001$, with the strongest effects emerging in the school-level studies and the weakest effects emerging in the national-level studies. For the examination between suspension and dropout, the moderators explained some of the variability between the studies. Therefore, the relationship between suspension and dropout varied depending on the level of the moderator.

DISCUSSION

The current meta-analysis was designed to quantify and synthesize the relationship between school suspension and achievement, as well as between school suspension and dropout. The results revealed a statistically significant inverse relationship between each type of suspension (i.e., ISS, OSS, combined, and not reported) and academic achievement. In addition, although there were insufficient studies to analyze data on each suspension type individually, a statistically significant positive relationship between overall suspension rate and dropout rate emerged for OSS. Moderator effects were also documented, suggesting that some of the variance across studies is accounted for by the sample and study characteristics examined. This study adds to the extant literature by using evidence aggregated from empirical sources that, in combination, demonstrate trends and trajectories related to the use of school suspensions. Specifically, this study filled a gap by quantifying the overall strength of the relationship between suspension and outcomes across a variety of locales and populations. Although causality cannot be inferred, the meta-analytic techniques used provide an evidentiary base for reconsidering schools' reliance on suspension as a means for addressing misbehavior. Practice and policy need to be informed by an evidence base derived from multiple studies showing consistent trends and ESs; this study contributes preliminarily to such evidence.

The unfavorable relationship between suspensions and both outcome variables is consistent with a plethora of recent calls for shifts away from the use of exclusionary discipline (e.g., Losen, 2011; Noltemeyer & Fenning, 2013). These findings are particularly concerning given that low-income and urban schools, those which often face greater challenges related to achievement and dropout, use school suspensions at significantly higher rates than other schools (Noltemeyer & Mcloughlin, 2010). This suggests that students who may experience heightened risk from the outset may be doubly disadvantaged by their schools' use of disciplinary practices that may

further exclude them from instruction that they need to progress educationally and alienate them from the school setting. Furthermore, the use of suspensions may initiate or exacerbate a process of student disengagement, the end result of which could be dropout or poor achievement. The study findings are even more concerning given that many school disciplinary policies require suspension for minor nonviolent infractions such as tardiness (Fenning et al., 2008).

The finding that OSS is more strongly associated with poor achievement than ISS makes intuitive sense. Multiple chapters and books critiquing the use of OSS have been written (e.g., Noltemeyer & Fenning, 2013), whereas ISS has largely avoided such focused scrutiny. This study provides quantitative empirical data aggregated across many studies to undergird this concern levied at OSS. The pathways through which OSS exerts a stronger negative influence than ISS can only be hypothesized. However, students receiving OSS may be more likely to miss academic instruction, may be provided with no opportunities to make up missed assignments, may experience negative influences from other delinquent peers in an unsupervised community setting, may experience a negative attitude toward school, and may lack educator support compared with peers receiving ISS. Although it is likely that OSS does exert more negative effects than ISS, an alternative explanation that should also be considered is that students assigned to ISS have likely engaged in less severe behaviors and therefore may experience less risk for academic disruption. Additional research is needed to examine this competing explanation for the findings.

Interestingly, several moderating effects emerged and suggested that the effects of suspension on outcomes vary based on several factors. Type of suspension; racial, gender, and socioeconomic status composition of the sample; type of publication; and level of data analysis were all significant moderators of both outcome variables. Although we can have confidence that these variables do moderate the effects of suspension on achievement and dropout, we urge readers to exercise caution in interpreting the specific role of any one level of the moderators because of the small number of studies included in different levels of the moderator variable. For example, although it appears that ISS is related to higher dropout rates than OSS, this finding was based on only one ES that examined ISS in relation to dropout rate and the confidence interval was quite wide. It is important to focus interpretations on those moderators that we can have more confidence in, such as the role of suspension type on achievement (because of the acceptable level of effects analyzed and small confidence intervals).

In addition to synthesizing the relationship between suspension and outcomes, this study also revealed the topography of the empirical work on this topic. We found no research at the preschool level and very little research at the elementary level. This is an important gap given findings from the Yale University Child Study Center (Gilliam, 2005) that the national expulsion rate of preschool students is 3.2 times higher than that of K–12 students. Moreover, dissertations represented a large number of the cases used in the study. Additional work published in peer-reviewed journals could help broaden the exposure of these issues to the general educational community. Finally, we found less research on ISS as compared with OSS; consequently, we were unable to examine the relationship between ISS and school dropout.

Limitations

There are several limitations of this study that should be considered when one is interpreting the results. First, we cannot definitively conclude that suspensions caused the outcomes. It could be that students who are low achievers, for example, are more likely to misbehave in the classroom because of frustration and therefore receive suspensions. In addition, most studies did not control for presuspension academic or behavioral difficulties that might influence the outcome variables, so we were not able to include these controls in the analyses. There are also several measurement issues that should be considered. For

example, some studies measured the variables at the school level and others at the individual-student level; we used estimated total numbers of participants for some studies; we averaged *r* values if more than one indicator of achievement or dropout was used in any given case; and we estimated dropout ESs based on school-completion ESs for some studies. Moreover, as previously mentioned, moderator analyses must be interpreted with caution given the small number of studies available for some levels of the moderator variables. Finally, there was not enough information on variables about the suspensions themselves to analyze the suspensions' causes, timing, and length. It is possible that these variables could moderate the relationship between suspensions and student outcomes.

Future Directions for Research

The findings and limitations reveal several directions for future research on this topic. First, additional empirical examinations of early childhood populations and research controlling for presuspension difficulties appear warranted. In addition, more research on variables that might moderate the relationships between suspension and outcomes is recommended, considering that moderators can reveal how individual ecology affects outcomes (Burns, 2011). For example, studies on the differential impact of suspensions based on their length, timing, and cause could inform policy, theory, and practice. Similarly, more research examining the moderators studied here is warranted, considering some findings that must be interpreted with caution. Finally, researchers should further examine the causal impact of OSS versus ISS.

Implications for Practice

These findings have implications for policy and practice, revealing that suspension is associated with deleterious outcomes. Simply eliminating suspension as a practice without other changes to support positive student outcomes is unlikely to be a panacea. Therefore, it is important for school psychologists to advocate for alternatives to suspension that incorporate prevention and early intervention. Noltemeyer and Fenning (2013) discuss several school-wide systems-change initiatives that could be considered to transform the culture of exclusionary discipline to a more proactive and preventive model. These include tiered models of support such as School-Wide Positive Behavior Support (SWPBS), which has been associated with a reduction in ISSs and office disciplinary referrals (e.g., Barrett, Bradshaw, & Lewis-Palmer, 2008). Another emerging initiative is restorative justice, which focuses on creating supportive schools that encourage students to take responsibility for their actions and repair harm that they may have caused rather than focusing on using exclusionary discipline as a means of retribution for misbehavior (Sumner, Silverman, & Frampton, 2010). Although the research evidence supporting the effectiveness of restorative justice is nascent and quite limited, preliminary case study evidence is promising (e.g., Sumner et al., 2010).

Despite preventive efforts, there will be circumstances in which suspension is warranted (e.g., if a student presents an immediate threat that makes his or her presence at school unsafe). When school suspension must be used, ISS should be considered when safe and appropriate (e.g., nonviolent offenses). ISS may have fewer negative consequences than OSS, although this assertion requires further evidence. Robinette (2012) discussed additional alternatives to OSS specified in California educational law, which include community service, requirements of parents to attend classes, temporary removal from class, and other corrective means (e.g., limitations on recess, behavior support plan, after-school programs). Chin et al. (2012) also proposed an alternative-to-suspension model that promotes prosocial outcomes by engaging students in a manner that fulfills their skill deficits and emotional needs.

It is important for all educational professionals to receive high-quality professional development on ways to promote prosocial behavior and address misbehavior. Educators who often begin the disciplinary process in the classroom should be given opportunities to

learn and practice these techniques with feedback. This is particularly important given findings that office disciplinary referrals initiated in the classroom are strong predictors of school suspension rates (Mcloughlin & Noltemeyer, 2010). In addition, school administrators who ultimately make suspension decisions should be well trained on the impact of these decisions and be provided guidance in crafting their schools' disciplinary policy. School psychologists can facilitate these types of professional development offerings while continuing to build their own knowledge on behavior supports.

To minimize the use of potentially harmful discipline practices, we also urge school psychologists to accept leadership in emphasizing the important role of data in the change process. School discipline data should be analyzed by a building-leadership team on a regular basis and disaggregated by demographic characteristics in an effort to monitor suspension trends and develop effective supports. To encourage schools to seriously tackle the issue of suspensions, we concur with Losen (2011) that suspension rate data could be considered in measures of school efficacy and that schools should be incentivized to support educators in improving classroom and behavior management in settings with high suspension rates.

Changing the disciplinary culture of schools is a complex challenge that requires vast shifts in educational practice and policy. Prevention-focused initiatives, alternatives to suspension, high-quality professional development and coaching, data-based decision making, and supportive disciplinary policy will all be essential to effectively transform this vision into a reality.

REFERENCES

*Ahn, S. (2010). *The effects of school membership on academic and behavioral performance of at-risk students* (Doctoral dissertation). Retrieved from ERIC database. (ED517291)

*Albertus, M. D. (2011). *The relationship between student literacy and student discipline: An analysis of student data in the Gwinnett public school system* (Doctoral dissertation). Retrieved from EBSCOhost PsycINFO database. (2011-99011-084)

*Allman, K. L. (2011). *Students with disabilities: The impact of student discipline on academic achievement* (Doctoral dissertation). Retrieved from EBSCOhost PsycINFO database. (2011-99210-168)

*Alspaugh, J. W. (1996). *The longitudinal effects of socioeconomic status on elementary school achievement.* Retrieved from ERIC database. (ED397120)

Arcia, E. (2006). Achievement and enrollment status of suspended students: Outcomes in a large, multicultural school district. *Education and Urban Society, 38*, 359–369. doi:10.1177/0013124506286947

Barrett, S. P., Bradshaw, C. P., & Lewis-Palmer, T. (2008). Maryland statewide PBIS initiative: Systems, evaluation, and next steps. *Journal of Positive Behavior Interventions, 10*, 105–114. doi:10.1177/10983 00707312541

*Boon, H. J. (2008). Risk or resilience? What makes a difference? *Australian Educational Researcher, 35*(1), 81–102. doi:10.1007/BF03216876

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2007). *Comprehensive meta-analysis* (Version 2) [Computer software]. Englewood, NJ: Biostat.

*Bruce, T. E. (2010). *Selected correlates of suspension of students in a regional secondary school in southeastern New England* (Doctoral dissertation). Retrieved from EBSCOhost ERIC database. (ED517082)

Burns, M. K. (2011). School psychology research: Combining ecological theory and prevention science. *School Psychology Review, 40*, 132–139.

*Carpenter, D. M., II, & Ramirez, A. (2007). More than one gap: Dropout rate gaps between and among Black, Hispanic, and White students. *Journal of Advanced Academics, 19*, 32–64.

Chapman, C., Laird, J., & KewalRamani, A. (2010). *Trends in high school dropout and completion rates in the United States: 1972-2008* (NCES 2011-012). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2011/2011012.pdf

Chin, J. K., Dowdy, E., Jimerson, S. R., & Rime, J. (2012). Alternatives to suspension: Rationale and recommendations. *Journal of School Violence, 11*, 156–173.

*Christle, C. A., Jolivette, K., & Nelson, C. M. (2007). School characteristics related to high school dropout rates. *Remedial and Special Education, 28*, 325–339. doi:10.1177/07419325070280060201

*Christle, C. A., Nelson, C. M., & Jolivette, K. (2004). School characteristics related to the use of suspension. *Education and Treatment of Children, 27*, 509–526.

Eckstrom, R. B., Goertz, M. E., Pollack, J. M., & Rock, D. A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record, 87*, 357–373.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634.

Fenning, P., Golomb, S., Gordon, V., Kelly, M., Scheinfield, R., Morello, T., & Banull, C. (2008). Written discipline policies used by administrators: Do we have

References marked with an asterisk indicate studies included in the meta-analysis.

sufficient tools of the trade? *Journal of School Violence, 7*, 123–146. doi:10.1300/J202v07n02_08

Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Education Evaluation and Policy Analysis, 13*, 189–202. doi:10.3102/01623737013002189

*Franklin, B. J., & Crone, L. J. (1992, November). *School accountability: Predictors and indicators of Louisiana school effectiveness.* Paper presented at the annual meeting of the Mid-South Educational Association, Knoxville, TN. (ERIC Document Reproduction Service No. ED 354 261)

Gifford-Smith, M., Dodge, K. A., Dishion, T. J., & McCord, J. (2005). Peer influence in children and adolescents: Crossing the bridge from developmental to intervention science. *Journal of Abnormal Child Psychology, 33*(3), 255–265. doi:10.1007/s10802-005-3563-7

Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems. Early education.* New Haven, CT: Yale University Child Study Center.

*Gonzalez, P., & Institute of Education Sciences (2006). *School behavior and disciplinary experiences of youth with disabilities. Facts from the National Longitudinal Transition Study-2 (NLTS2).* Washington, DC: U.S. Department of Education.

*Goran, L. G., & Gage, N. A. (2011). A comparative analysis of language, suspension, and academic performance of students with emotional disturbance and students with learning disabilities. *Education & Treatment of Children, 34*, 469–488. doi:10.1353/etc.2011.0035

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Thousand Oaks, CA: Sage.

Hyman, I. A. (1997). *School discipline and school violence: The teacher variance approach.* Boston, MA: Allyn & Bacon.

*Johnston, J. S. (1989). High school completion of in-school suspension students. *NASSP Bulletin, 73*(521), 89–95. doi:10.1177/019263658907352117

*Kim, J. (2011). *Relationships among and between ELL status, demographic characteristics, enrollment history, and school persistence* (CRESST Report 810). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

*Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools, 43*, 701–712. doi:10.1002/pits.20177

*Lee, T., Cornell, D., Gregory., A., & Fan, X. (2011). High suspension schools and dropout rates for Black and White students. *Education & Treatment of Children, 34*, 167–192.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Lochner, L., & Moretti, E. (2002). The effect of education on crime: Evidence from prison inmates, arrests and self-reports. *The American Economic Review, 94*, 155–189. doi:10.1257/000282804322970751

Losen, D. J. (2011). *Discipline policies, successful schools, and racial justice.* Boulder, CO: National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/discipline-policies

Losen, D. J., & Skiba, R. J. (2010). *Suspended education: Urban middle schools in crisis.* Retrieved from http://civilrightsproject.ucla.edu/research/k-12-education/school-discipline/suspended-education-urban-middle-schools-in-crisis

Mac Iver, M. A. (2011). The challenge of improving urban high school graduation outcomes: Findings from a randomized study of dropout prevention efforts. *Journal of Education for Students Placed at Risk, 16*, 167–184. doi:10.1080/10824669.2011.584497

Mcloughlin, C. S., & Noltemeyer, A. L. (2010). Research into factors contributing to discipline use and disproportionality in major urban schools. *Current Issues in Education, 13*(2), 1–21.

*Moore, J. E. (1987, April). *Middle school folklore: It may be true!* Paper presented at the annual meeting of the American Educational Association, Washington, DC. Retrieved from ERIC database. (ED283902)

*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 & Treatment of Children, 24*, 276–293.

*Muyskens, P., Marston, D., & Reschly, A. L. (2007). The use of response to intervention practices for behavior: An examination of the validity of a screening instrument. *California School Psychologist, 12*, 31–45.

Noltemeyer, A., & Fenning, P. (2013). Systemic school discipline: Issues of equity from a social justice perspective (pp. 91–117). In D. Shriberg, S. Y. Song, A. H. Miranda, & K. M. Radliff (Eds.), *School psychology and social justice: Conceptual foundations and tools for practice.* New York, NY: Routledge.

Noltemeyer, A. L., & Mcloughlin, C. S. (2010). Patterns of exclusionary discipline by school typology, ethnicity, and their interaction. *Perspectives on Urban Education, 7*, 27–40.

Northeastern University–Center for Labor Market Studies. (2009). *Left behind in America: The nation's dropout crisis.* Boston, MA: Center for Labor Market Studies Publications. Retrieved from http://hdl.handle.net/2047/d20000598

Quintana, S. M., & Minami, T. (2006). Guidelines for meta-analyses of counseling psychology research. *The Counseling Psychologist, 34*, 839–877. doi:10.1177/0011000006286991

*Raffaele Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development, 99*, 17–33.

Raffaele Mendez, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education & Treatment of Children, 26*, 30–51.

*Rausch, M. K., & Skiba, R. (2004). Unplanned outcomes: Suspensions and expulsions in Indiana. *Education Policy Briefs, 2*(2), 1–8. Retrieved from http://ceep.indiana.edu/projects/PDF/PB_V2N2_UnplannedOutcomes.pdf

*Ridings, K. R. (2010). *The place of at-risk factors among students graduating or dropping out of high school: A study of path analyses* (Doctoral dissertation). Retrieved from EBSCOhost ERIC database. (ED521726)

Robey, R. R., & Dalebout, S. D. (1998). A tutorial on conducting meta-analysis of clinical outcome research. *Journal of Speech, Language, and Hearing Research, 41*, 1227–1241.

Robinette, D. (2012). Alternatives to student suspension. *Leadership, 42*, 32–36.

*Safer, D. J. (1986). Nonpromotion correlates and outcomes at different grade levels. *Journal of Learning Disabilities. 19*, 500–503. doi:10.1177/002221948601900812

Shadish, W. R., Jr., & Sweeney, R. B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *Journal of Consulting and Clinical Psychology, 59*, 883–893. doi:10.1037/0022-006X.59.6.883

Sharkey, J. D., & Fenning, P. A. (2012). Rationale for designing school contexts in support of proactive discipline. *Journal of School Violence, 11*, 95–104. doi:10.1080/15388220.2012.646641

*Shepard, C. (2010). *An investigation of middle school suspensions and the relationship of middle school suspensions to district-wide test scores* (Doctoral dissertation). Retrieved from EBSCOhost PsychINFO database. (2010-99171-137)

Shirley, E. L. M., & Cornell, D. G. (2011). The contribution of student perceptions of school climate to understanding the disproportionate punishment of African American students in a middle school. *School Psychology International, 33*, 115–134. doi:10.1177/0143034311406815

*Silvey, D. F. (1995). The effect of in-school suspension on the academic progress of high school science and English students. Retrieved from ERIC database. (ED389069)

Skiba, R. J., Horner, R. H., Chung, C-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review, 40*, 85–107.

*Skiba, R., & Rausch, M. K. (2004). *The relationship between achievement, discipline, and race: An analysis of factors predicting ISTEP scores. Children left behind policy briefs supplementary analysis 2-D.* Bloomington, IN: Indiana University Center for Evaluation and Education Policy.

*Skiba, R., Simmons, A., Staudinger, L., Rausch, M., Dow, G., & Feggins, R. (2003). *Consistent removal: Contributions of school discipline to the school-to-prison pipeline.* School to Prison Pipeline Conference: Harvard Civil Rights Project. Retrieved from http://varj.onefireplace.org/Resources/Documents/Consistent%20Removal.pdf

*Stokes, S. B. S. (2012). *The relationship between suspension, student engagement, and dropout* (Doctoral dissertation). Retrieved from EBSCOhost PsycINFO database. (2012-99030-208)

*Strompolis, M., Vishnevsky, T., Reeve, C. L., Munsell, E. P., Cook, J. R., & Kilmer, R. P. (2012). Educational outcomes in a system of care for children with emotional disturbance. *American Journal of Orthopsychiatry, 82*, 129–136. doi:10.1111/j.1939-0025.2011.01126.x

*Suh, S., & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling, 10*, 297–306.

Sullivan, A. L., Klingbeil, D., & Van Norman, E. (2013). Beyond behavior: Multilevel analysis of the influence of sociodemographics and school characteristics on students' risk of suspension. *School Psychology Review, 47*, 99–114.

Sumner, M. D., Silverman, C. J., & Frampton, M. L. (2010). *School-based restorative justice as an alternative to zero-tolerance policies: Lessons from West Oakland.* Berkeley, CA: Thelton E. Henderson Center for Social Justice, University of California, Berkeley, School of Law.

*Sweeten, G. A. (2004). *School dropout and subsequent offending: Distinguishing selection from causation* (Master's thesis, University of Maryland). Retrieved from http://drum.lib.umd.edu/bitstream/1903/191/1/dissertation.pdf

*Thompson, M. A. (2006). *The relationship between student disciplinary referrals and Illinois Standards Achievement Test scores in a Midwestern urban school district* (Doctoral dissertation). Retrieved from EBSCOhost PsycINFO database. (2011-99091-112)

*Tobin, T. J., & Sugai, G. M. (1999). Using sixth-grade school records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders, 7*, 40–53. doi:10.1177/106342669900700105

Townsend, L., Flisher, A. J., & King, G. (2007). A systematic of the relationship between high school dropout and substance use. *Clinical Child and Family Psychology Review, 10*, 295–317. doi:10.1007/s10567-007-0023-7

U.S. Department of Education Office of Civil Rights. (2014). *Civil rights data collection: Data snapshot (school discipline).* Retrieved from http://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf

Waldfogel, J., Garfinkel, I., & Kelly, B. (2007). Welfare and the costs of public assistance. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay: Economic and social consequences of inadequate education* (pp. 160–176). Washington, DC: Brookings Institution.

*Williams, T. R., Davis, L. E., Miller Cribbs, J., Saunders, J., & Williams, J. H. (2002). Friends, family, and neighborhood: Understanding academic outcomes of African American youth. *Urban Education, 37*, 408–431. doi:10.1177/0042085902373006

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis.* Newbury Park, CA: Sage.

Wu, S., Pink, W., Crain, R. L., & Moles, O. (1982). Student suspension: A critical reappraisal. *Urban Review, 14*, 245–303. doi:10.1007/BF02171974

*Zablocki, M. S. (2009). *Predicting school dropout among youth with disabilities: The role of youth characteristics, academic experiences, and emotional engagement factors.* Available from ProQuest Dissertations and Theses database.

Date Received: January 15, 2014
Date Accepted: May 18, 2014
Associate Editor: Mike Vanderwood ■

Amity L. Noltemeyer, PhD, NCSP, is an associate professor in school psychology at Miami University. Her research interests include school discipline disparities, resilience, multitiered systems of support, and school-based mental health services.

Rose Marie Ward, PhD, is a professor in kinesiology and health at Miami University. Her research interests include college student alcohol consumption, sexual assault, and advanced statistical methods.

Caven Mcloughlin, PhD, NCSP, is a professor in school psychology at Kent State University. He focuses on interpreting policy directives to stakeholders, providing psychological services to children in schools internationally, and building the capacity of faculty research productivity in non-Western university settings.