

Local Implementation of State-Level Discipline Policy: Administrator Perspectives and Contextual Factors Associated With Compliance

Kaitlin P. Anderson 

Lehigh University

Sarah McKenzie

University of Arkansas

Many states and school districts are implementing reforms to reduce reliance on exclusionary discipline such as out-of-school suspension and expulsion. This article uses survey and administrative data to study the implementation of a state-level policy limiting elementary school out-of-school suspension and expulsions. While the results are limited in sample size and generalizability, we find that survey respondents from relatively disadvantaged schools reported greater difficulties and challenges with implementation and tended to comply at lower rates. Policy makers seeking more equitable implementation of these types of reforms should work with schools to better understand and attend to local preferences, buy-in, and capacity for implementation.

Keywords: *educational policy, educational reform, elementary schools, policy analysis, principals, regression analyses, survey research*

Introduction

In recent years, many school districts and states have sought to reduce exclusionary discipline such as out-of-school suspension (OSS) and expulsion due to growing awareness of the academic and life consequences of excluding students from school (e.g., Anderson, Ritter, & Zamarro, 2019; Bacher-Hicks et al., 2019; Chu & Ready, 2018; Davison et al., 2021; Fabelo et al., 2011; Nicholson-Crotty et al., 2009; Wolf & Kupchik, 2017), as well as concern about disproportionalities in exclusionary discipline by race and disability status (e.g., Anderson, 2021; Anderson & Ritter, 2017; Liu et al., 2021; Losen et al., 2015; Shi & Zhu, 2021; Skiba et al., 2011; Skiba et al., 2014; Sullivan et al., 2013; Vincent et al., 2012).

In this article, we study the school-level implementation of an Arkansas state law restricting the use of exclusionary discipline in elementary schools. Act 1059 stated that districts “shall not use out-of-school suspension or expulsion for a student in kindergarten through grade five” except when the behavior “poses a physical risk” or “causes a serious disruption that cannot be addressed through other means.” We combine administrative data with survey responses from principals and counselors to ask the following research questions:

Research Question 1: What school and district characteristics are associated with principal and counselor perspectives about Act 1059 and its implementation?

Research Question 2: What school and district characteristics are associated with compliance with Act 1059 as reported in the administrative data?

We use empirical evidence of similar discipline reforms, as well as theoretical perspectives on policy implementation to frame our study. Similar student discipline reforms have been incompletely implemented (Anderson, 2018; Steinberg & Lacoé, 2018) due to a variety of reasons that may include lack of staffing and space for in-school options (A. N. Gray et al., 2017), unawareness or impracticality of alternative approaches, or lack of buy-in due to local preferences and needs. School leaders and support staff are in a unique position to influence the implementation of discipline policy, so understanding their perspectives on reforms such as Act 1059 is important for assessing implementation.

While the results are limited in terms of sample size and generalizability, we find that survey respondents from relatively disadvantaged schools reported greater difficulties and challenges with implementation, and less success post-policy. There is also suggestive evidence that relatively disadvantaged schools complied at lower rates. Our findings



indicate inequities with respect to the policy implementation challenges schools were facing and the potential impact on the school community. Policy makers seeking more equitable implementation of these types of reforms should work with schools to better understand and attend to local preferences, buy-in, and capacity for implementation.

In the sections that follow, we discuss relevant literature on the issue of exclusionary discipline in schools, the research related to reforms in this area, as well as theories related to policy implementation that help frame our study. Then, we describe Act 1059 and the Arkansas policy context, the data and methods, our results, and the implications for policy and practice.

Literature Review and Theoretical Framework

Exclusionary Discipline in Schools

Exclusionary discipline removes students from the learning environment and is associated with worse academic outcomes, including lower achievement, grade retention, and dropout (Anderson, Ritter, & Zamarro, 2019; Balfanz et al., 2014; Chu & Ready, 2018; Cobb-Clark et al., 2015; Fabelo et al., 2011; Noltemeyer et al., 2015; Suh & Suh, 2007; Swanson et al., 2017). Students excluded from school for disciplinary reasons are also more likely to be arrested or involved in the criminal justice system (Davison et al., 2021; Fabelo et al., 2011; Mowen & Brent, 2016; Nicholson-Crotty et al., 2009; Wolf & Kupchik, 2017).

Black students in particular are at higher risk of being suspended (e.g., Anderson & Ritter, 2017; Losen et al., 2015; Skiba et al., 2011; Skiba et al., 2014; Skiba et al., 2002; Sullivan et al., 2013; Vincent et al., 2012). The racial gaps in student discipline are not simply due to differences in rates of misbehavior, as Black students tend to receive more exclusionary consequences, controlling for the type of infraction (Anderson & Ritter, 2017; Ritter & Anderson, 2018), and even when comparing outcomes for students involved in the same incident, such as a fight between students from different racial or ethnic backgrounds (Barrett et al., 2019; Liu et al., 2021; Shi & Zhu, 2021). One reason that Black students are disproportionately suspended is that they attend more punitive, exclusionary schools (Anderson & Ritter, 2017; Christie et al., 2004; Gopalan & Nelson, 2019; Payne & Welch, 2010; Sheldon & Epstein, 2002; Skiba et al., 2014; Welch & Payne, 2010). When students are assigned to high suspension schools, this harms their educational attainment and increases their likelihood of arrest and incarceration as an adult (Bacher-Hicks et al., 2019).

Of particular concern is the exclusion of young students, although there is a lack of causal evidence in this area. Exclusionary discipline may be developmentally inappropriate for young children still learning school norms (Jacobsen et al., 2019). If formal sanctions become stressors, they might result in behavioral problems, and younger students

may be less able to cope appropriately (Agnew & Brezina, 2010). Furthermore, students suspended in elementary school are also more likely to be suspended in middle school (Raffaale Mendez, 2003).

Reforms to School Discipline Policy and Practice

Many states and school districts have made efforts to reduce reliance on exclusionary discipline and incorporate alternative approaches, with mixed results. Broadly, these efforts can be categorized as policy reforms or programmatic responses. Policy-based suspension reforms include reducing the length of suspensions, reducing the type of infractions for which suspension is legal, and eliminating mandatory suspensions. Some of these reforms have focused on elementary grades. As of 2018, at least 16 states, plus Washington D.C., limited the use of suspension or expulsion in early grade levels (Rafa, 2018).¹ In many cases, these laws allow exceptions, particularly for safety concerns.²

A growing body of evidence indicates discipline policy reforms have had mixed effects depending on the outcomes and group of students assessed (Anderson, 2020; Anderson, Egalite, & Mills, 2019; Hashim et al., 2018; Hinze-Pifer & Sartain, 2018; Lacoë & Steinberg, 2018; Sartain et al., 2015; Steinberg & Lacoë, 2018). The way such policies are implemented is important for interpreting their effects. For example, a 2013 ban on OSS for truancy in Arkansas led to slight increases in attendance, with suggestive evidence of achievement gains and reductions in disciplinary referrals (Anderson, 2020), but schools with more non-White students were less likely to comply with the policy (Anderson, 2018). Steinberg and Lacoë (2018) studied a reform in Philadelphia that prohibited OSS for classroom disorder infractions and found that students who had previously been suspended had improved attendance and reduced classroom disorder OSS after the reform. Where the policy was implemented fully, nonsuspended peers did not experience spillover effects, but where the reform was not implemented fully, nonsuspended students experienced decreased attendance and math achievement.

There is still relatively little evidence on discipline reforms focused on elementary schools, a gap this study begins to fill. One exception is an analysis of Maryland's 2017 law prohibiting suspension or expulsion in preK through second grade, with some exceptions (Maryland State Department of Education, 2020). The report showed, after this law, OSS was not eliminated, but decreased significantly, with declines in each grade PK–2 and for most groups of students. While most local school systems reduced their use of suspensions, a few showed no statistically significant difference postreform. The current study adds to the currently limited literature on the implementation and impact of school discipline reforms in elementary grades.

Administrator and Educator Roles and Perspectives on Student Discipline

A variety of individual-, school-, and system-level factors may influence school personnel's perspectives on and implementation of student discipline reform. Our investigation into these potential relationships is informed by the role of individual school personnel in student discipline policy and practice, as well as the literature on policy change, dissemination, and implementation.

We seek to understand principal and counselor perspectives on Act 1059, as their role and perspectives are relevant to the fidelity of reform implementation. Teachers' perspectives would also be informative, as they are often involved at the disciplinary referral stage, but gaining contact information for a representative set of teachers across the state was not feasible.

Principal perspectives are particularly important, as principals tend to be responsible for making decisions about how to respond to referrals, within the bounds of district policies (e.g., codes of conduct). As such, principals are in a unique position to influence school climate and discipline practices. For example, principals' propensity to remove students for misbehavior has an impact on important outcomes; greater principal propensity to remove reduces the number of reported offenses—a possible deterrence effect—but increases use of exclusionary discipline conditional on referral, juvenile justice referrals, and student dropout (Sorenson et al., 2021). Also, principals with more favorable attitudes toward exclusionary discipline use it at higher rates (Mukuria, 2002; Skiba et al., 2014).

School counselors have training to prevent disruptive behavior and collaborate with other school personnel to establish policies or programs that prevent disruptive behaviors. Typically, for the purposes of maintaining positive relationships with students, school counselors are not involved with administering discipline, yet they may be a "significant contributor to the development of the prevention and intervention programs through which problem behaviors are managed and positive behaviors are nurtured" (American School Counselor Association, 2020). As a result, we include counselor perspectives on this reform as well.

Policy Adoption and Implementation

Literature on policy implementation is particularly relevant here. There may be local factors that prevent full implementation of state policy. Successful policy implementation requires resources and capacity, and if a policy's demands outweigh the ability or will of local actors, implementation fidelity may suffer (Cohen et al., 2007; Hill & Hupe, 2003; Matland, 1995; McLaughlin, 1987). Disconnect between policy intent (at the state-level) and perception (at the local-level) may create implementation issues, as policy makers have been characterized as "ignorant of practice" (Cohen

et al., 2007, p. 522). Top-down theorists have suggested that policies should have clear, consistent goals and limit the amount of change necessary from local actors (Mazmanian, & Sabatier, 1983; Van Meter & Van Horn, 1975).

Policy ambiguity—although sometimes necessary to build a coalition to get legislation passed (Baier et al., 1986; Matland, 1995)—may prevent policy from being implemented as intended (Baier et al., 1986; Cohen et al., 2007; Firestone, 1989; Sabatier & Mazmanian, 1979; Weatherly & Lipsky, 1977). Ambiguous policy language or goals may mean practitioners do not have a clear vision for implementation (Cohen et al., 2007; Matland, 1995). In the case of Act 1059, there is relatively ambiguous language about the allowable exceptions to ban, which creates an opportunity for misalignment between original intent and eventual implementation.

Finally, practitioner beliefs and attitudes, which may be highly context dependent, also play an important role in how policies will be implemented. Principals filter policy through their own beliefs and worldviews, affecting interpretation (Spillane et al., 2002). Given the relationship between administrator attitudes and exclusionary discipline use (Mukuria, 2002; Skiba et al., 2014), such beliefs and attitudes are nontrivial.

In addition to theories of policy implementation (discussed above), there are theories related to policy adoption and diffusion that are useful for exploring why certain laws are enacted in certain states at certain times. There are schools of thought related to whether policy is largely static, unchangeable, and path dependent, more dynamic and adaptable, or a combination, in which policy is slow to change at times, and very responsive to issues and pressure to reform in others, depending on contextual factors (see Cohen-Vogel & McLendon, 2012, for discussion). Within this third perspective, theories of punctuated equilibrium (Baumgartner & Jones, 1991, 1993), multiples streams (Kingdon, 1984), policy regimes (Skowronek, 1993), and policy innovation and diffusion (Berry & Berry, 1990; V. Gray, 1994; Mintrom, 1997; Walker, 1969), have all sought to explain the nature of policy change, and may be useful for understanding the enactment of state education policy. Even though a test of these theories is out of the scope of the current study, these theories would be useful for future research in this area.

Arkansas Act 1059

Arkansas Act 1059 of 2017 stated: *The school district shall not use out-of-school suspension or expulsion for a student in kindergarten through grade five (K–5) except in cases when a student's behavior: (A) Poses a physical risk to himself or herself or to others; or (B) Causes a serious disruption that cannot be addressed through other means.*

Prior to Act 1059, Arkansas' school districts had a high degree of flexibility regarding student discipline,³ except that

suspensions for truancy offenses were banned by Act 1329 of 2013. While Act 1059 did not completely ban OSS or expulsion, the apparent intent was to significantly reduce exclusionary discipline in grades K-5 except in extreme circumstances. Act 1059 did not limit in-school suspension (ISS) and did not specifically mention alternative approaches, although Act 1329 of 2013 encouraged the use of “evidence-based strategies” such as “positive behavior interventions and support systems” and “restorative justice.” Effective implementation of such approaches requires time and resources, and top-down policy reforms may conflict with local perspectives on how disciplinary issues should be handled based on local needs. As a result, it is important to understand how this law was perceived and implemented across different school contexts. This study is important given that a 2013 Arkansas law banning OSS for truancy was not implemented fully or equitably (Anderson, 2018), and as a result, had limited impact on student outcomes (Anderson, 2020).

Data and Survey Methods

We use data from three sources: (1) restricted-use data provided by the Arkansas Department of Education, (2) publicly available data from the National Center for Education Statistics (NCES), and (3) survey responses from elementary school principals and counselors.

Administrative and Publicly Available Data

The administrative data include 5 years (2014–2015 through 2018–2019) of student demographics, test scores, and infraction-level disciplinary incident records for Arkansas public school students in grades K–5, including charter schools. Student-by-year level data are used to create school-by-year level measures of demographic information (e.g., total school enrollment, school percent FRL (free or reduced-price lunch), school percent non-White students, and school percent special education). Similarly, student-by-year level test scores on state mathematics and reading/English language arts (ELA) test for students in Grades 3 to 8 for all tests administered without accommodations⁴ are standardized by subject, grade level, and academic year to have a mean of zero and standard deviation of one (i.e., *z*-scores), and then averaged at the school-by-year level. We hypothesize that school size, student demographics, and achievement will be statistically related to our outcome measures, and these serve as our main variables of interest. School percent special education is used primarily as a control variable.

The discipline data include infraction and consequence types for each incident. Table 1 presents the frequency of each infraction and consequence type, by year. Beginning in 2016–2017, the state reported new categories that were previously included in an “other” category.⁵ Over time, the

number of “other” infractions and consequences, including these new groups, generally increased. This does not necessarily mean that these “other” misbehaviors were on the rise, as reporting of these incidents may have been increasing as well. Furthermore, there is uncertainty about what is included in this category, and it may still include relatively serious incidents. From 2014–2015 to 2018–2019, the most common infraction types were disorderly conduct (34.5% of K–5 infractions) and insubordination (16.3%). For our analyses, we group these together as “subjective” infractions. Given Act 1059’s exception for behaviors that pose a risk of physical harm, we expect to see more declines among these subjective infractions. Panel B shows the consequences reported during this period. Across all 5 years, the most common consequences were “other” (41.1%), ISS (28.4%), OSS (16.0%), and corporal punishment (13.9%).

Table 2 reports the types of infractions for which OSS and expulsion were given, in Grades K–5, by year. The first four columns indicate counts, by type. Most elementary OSS/expulsion is for subjective infractions like disorderly conduct (31.7% of OSS and expulsion over this 3-year period) and insubordination (12.5%), as well as fighting (19.6%), and “other” (24.0%).⁶ The last four columns show the share of infractions resulting in OSS or expulsion, by infraction type. Violence/weapon infractions result in OSS/expulsion at a higher rate (35.8%), as do substances (49.5%), gang-related incidents (57.9%), and terroristic threats (53.0%). Subjective infractions only result in OSS/expulsion 13.6% of the time.

Despite the low OSS/expulsion rate for subjective infractions (13.6%), the sheer number of those infractions means they contribute a large share of OSS/expulsions (44.2%). These incidents may be relatively minor and could plausibly be handled through less exclusionary responses. Indeed, the use of OSS/expulsion for subjective categories declined from 15.7% in 2016–2017 to 13.2% in 2018–2019. We also see decreased OSS and expulsions for other infraction types, with increases only occurring for staff assault and public displays of affection (based on incident counts). Although overall, exclusionary discipline decreased postreform, OSS/expulsion were used in almost 9,000 incidents in 2018–2019, representing approximately 5,000 elementary-age students suspended or expelled at least once that year.

For our analysis, these incident-level data are used to create both predictors of interest (for Research Question 1), and outcome variables (for Research Question 2). For Research Question 1, we use 3 years of discipline data (2014–2015 to 2016–2017) to calculate two measures of school-level baseline discipline practices as predictor variables: (1) the average number of infractions, per 100 students, per year, averaged over the last three prepolicy years, and (2) the overall share of infractions reported during the 3-year baseline period that resulted in exclusionary consequences (OSS, expulsion, or referral to an ALE). Using 3 years of data

TABLE 1

Frequency of Infractions and Consequences, Grades K–5, by Type (2014–2015 to 2018–2019)

Infractions/consequences	Pre-Act 1059			Post-Act 1059		Total	% of total
	2014–2015	2015–2016	2016–2017	2017–2018	2018–2019		
<i>Panel A: Infractions, by type</i>							
Disorderly conduct	23,518	25,845	26,360	26,582	21,700	124,005	34.5
Insubordination	11,885	13,234	12,917	12,347	8,299	58,682	16.3
Fighting	5,570	6,630	6,827	7,261	5,321	31,609	8.8
Bullying	2,589	2,635	2,264	2,306	1,787	11,581	3.2
Student assault	1,292	1,291	1,386	1,919	1,320	7,208	2.0
Staff assault	300	309	324	424	417	1,774	0.5
Vandalism	376	365	385	358	289	1,773	0.5
Truancy	163	221	247	288	176	1,095	0.3
Knife	196	159	187	181	127	850	0.2
Tobacco	49	46	33	47	28	203	0.1
Drugs	33	41	37	41	14	166	0.0
Explosives	16	8	20	10	17	71	0.0
Club	31	13	14	1	2	61	0.0
Guns	23	7	9	7	4	50	0.0
Gangs	11	3	8	3	8	33	0.0
Alcohol	6	6	5	7	6	30	0.0
Total other	20,983	22,083	26,811	27,724	22,350	119,951	33.4
Other	20,983	22,083	24,825	25,346	20,499	113,736	
Stealing/theft			929	1,078	850	2,857	
Harassment/sexual harassment			490	477	473	1,440	
Terroristic threats			256	396	306	958	
Cellphone/electronic devices			267	365	166	798	
Public display of affection			29	36	36	101	
Cyberbullying			15	26	20	61	
Total	67,041	72,896	77,834	79,506	61,865	359,142	100.0
<i>Panel B: Consequences, by type</i>							
Expulsion	34	31	16	41	27	149	0.0
Referral to ALE	141	113	103	128	100	585	0.2
Out-of-school suspension	11,609	12,641	13,373	11,110	8,767	57,500	16.0
In-school suspension	18,453	20,971	21,325	23,627	17,476	101,852	28.4
Corporal punishment	11,068	11,033	10,447	9,737	7,532	49,817	13.9
No action	435	341	211	387	252	1,626	0.5
Total other	25,301	27,766	32,359	34,476	27,711	147,613	41.1
Other	25,301	27,766	17,563	16,870	12,835	100,335	
Warning			5,141	6,572	5,949	17,662	
Detention			4,641	5,970	4,410	15,021	
Bus suspension			3,071	3,077	2,523	8,671	
Parent conference			1,547	1,692	1,666	4,905	
Saturday school			340	238	154	732	
More than one type of “other”			56	57	174	287	
Total	67,041	72,896	77,834	79,506	61,865	359,142	100.0

Note. Beginning in 2016–2017, some additional categories began being separately reported. About 1% of infractions resulted in more than one consequence type. The numbers reported indicate the most-exclusionary/severe consequence type. Thus, each category may include some incidents that resulted in the reported category plus some less exclusionary/severe consequences. Act 1059 was passed in April of 2017, near the end of the school year, so we include all years 2014–2015 to 2016–2017 as pre-Act 1059, and the years 2017–2018 and 2018–2019 as post-Act 1059. ALE = alternative learning environment.

TABLE 2

Types of Infractions for Which K–5 Students Are Expelled or Suspended (2016–2017 to 2018–2019)

Infraction type	Counts of infractions resulting in OSS/ expulsion				Percentage of infractions resulting in OSS/expulsion			
	2016–2017	2017–2018	2018–2019	3 years total	2016–2017	2017–2018	2018–2019	3 years total
Subjective	6,168	4,624	3,963	14,755	15.7	11.9	13.2	13.6%
Disorderly conduct	4,275	3,313	2,992	10,580	16.2	12.5	13.8	14.2%
Insubordination	1,893	1,311	971	4,175	14.7	10.6	11.7	12.4%
Violence/weapons	3,426	3,449	2,357	9,232	39.1	35.2	32.7	35.8%
Fighting	2,516	2,485	1,541	6,542	36.9	34.2	29.0	33.7%
Student assault	523	574	428	1,525	37.7	29.9	32.4	33.0%
Staff assault	225	248	293	766	69.4	58.5	70.3	65.8%
Knife	137	125	81	343	73.3	69.1	63.8	69.3%
Club	8	1	1	10	57.1	100.0	50.0	58.8%
Explosives	10	9	9	28	50.0	90.0	52.9	59.6%
Guns	7	7	4	18	77.8	100.0	100.0	90.0%
Substances	38	53	17	108	50.7	55.8	35.4	49.5%
Drugs	22	29	9	60	59.5	70.7	64.3	65.2%
Tobacco	12	18	6	36	36.4	38.3	21.4	33.3%
Alcohol	4	6	2	12	80.0	85.7	33.3	66.7%
Bullying	418	347	287	1,052	18.5	15.0	16.1	16.5%
Vandalism	69	52	37	158	17.9	14.5	12.8	15.3%
Truancy	32	13	3	48	13.0	4.5	1.7	6.8%
Gangs	8	1	2	11	100.0	33.3	25.0	57.9%
Total other	3,234	2,620	2,147	8,001	12.1	9.5	9.6	10.4%
Other	2,757	2,145	1,816	6,718	11.1	8.5	8.9	9.5%
Harassment/sexual harassment	146	120	100	366	29.8	25.2	21.1	25.4%
Cyberbullying	4	5	2	11	26.7	19.2	10.0	18.0%
Cellphone/electronic devices	12	11	4	27	4.5	3.0	2.4	3.4%
Public display of affection	3	3	8	14	10.3	8.3	22.2	13.9%
Stealing/theft	160	121	76	357	17.2	11.2	8.9	12.5%
Terroristic threats	152	215	141	508	59.4	54.3	46.1	53.0%
Total	13,393	11,159	8,813	33,365	17.2	14.0	14.2	15.2%

Note. Total number in each year is higher than the sum of OSS and expulsion in Table 1 because this table includes a small number of infractions that included OSS along with a referral to an ALE, which in Table 1, were included within the row for ALE. OSS = out-of-school suspension; ALE = alternative learning environment.

creates more stable measures of baseline discipline use than a single year's data. To alleviate the influence of outliers, these two measures are Winsorized (Dixon, 1960; Locker, 2001) to replace observations below the fifth percentile with the fifth percentile value.⁷

For Research Question 2, we use the discipline data to create four outcome measures related to compliance with the law. We hypothesize declines in OSS/expulsion use will be concentrated among subjective infractions, so the four measures capture the 2016–2017 to 2017–2018 percentage change and the 2016–2017 to 2018–2019 percentage change in the share of infractions resulting in OSS/expulsion, both overall and for subjective infractions. Percentage changes

relative to very small base rates (small denominators) are often outliers that could skew the data. As such, we Winsorize these variables, replacing values above the 95th percentile with the 95th percentile value.

We also incorporate publicly available NCES data on district-level per-pupil instructional expenditures as an additional independent variable of interest, as we expect funding will be related to capacity for compliance.

Survey Design and Measures

Administrative data are supplemented with survey measures, collected through a retrospective survey design,

asking respondents to recall the year prior to the policy and to reflect on changes since that year. Retrospective surveys have limitations, such as memory distortion or recall (Pearson et al., 1992), and benefits, particularly with regard to measuring perceptions of change over time (Little et al., 2020). Nonretrospective baseline survey data are also imperfect, as response shift bias (Bray et al., 1984) may occur when an individual's frame of reference or their internal rating system changes between the pre- and posttests. While retrospective data are subject to misremembering (Pearson et al., 1992), they are less subject to response shift bias.

Survey items were adapted from a variety of sources with the goal of assessing principal and counselor perspectives on Act 1059. We used principal components analysis (PCA) rather than factor analysis, because PCA does not assume an underlying causal relationship between the factors and the observed variables (Floyd & Widaman, 1995). Given that the included items are noncontinuous Likert-type items, we used polychoric correlations (Olsson, 1979) to estimate the correlation matrix used in the PCA. We follow the Kaiser–Guttman rule (Kaiser, 1991) to estimate the number of principal components (those with an eigenvalue of greater than one) and used oblique (promax) rotation to allow the identified components to be correlated (Crocker & Algina, 1986). Descriptive statistics for each individual item and summary construct are in the Supplemental Appendix Table A (available in the online version of this article). Next, we describe these constructs, and their related measures, used as outcomes in Research Question 1.

We asked respondents to indicate the extent to which they agreed or disagreed with 14 statements about this law (on a scale from 1 to 4). Seven out of 14 items loaded to a single component that reflected *Support for the law* (e.g., whether it's reasonable, feasible, whether it will help students, and whether it will hurt school climate). The other seven loaded to a second component, *Clarity for implementation*. Items include, for example: “the wording of this law is ambiguous,” “expectations with respect to this law are clear to me,” and “I have a clear vision for how to successfully implement this law in my school.”

We combine two items indicating the school's reported difficulty with implementation in each of the first two outcome years. We asked, “To what extent did your school experience difficulties implementing this law during the two most recent school years?” (in 2017–2018 and 2018–2019). Respondents responded on a scale from one (“no difficulties”) to four (“substantial difficulties”). For ease of interpretation, we created a single binary indicator of whether a respondent indicated at least a three (moderate or higher), on average, across these 2 years.

Two measures relate to challenges/barriers for implementation, based on a total of 11 items. Four items loaded to one principal component (*Challenges related to politics, policies, and leadership*), and seven to another (*Challenges*

related to lack of resources and capacity). Both measures are on a scale from one to four, with four indicating greater challenges.

We use a measure of the change in reported school climate from 2016–2017 (prepolicy) to 2018–2019 (the second postpolicy year). We reviewed many school climate measures, but no existing scale perfectly met our needs, so we developed a brief instrument to incorporate many key components (see online Supplemental Appendix Table A).⁸ Specifically, we asked items related to orderliness, teacher–student relationships, leadership, and school mission. We collected information on 14 items, separately, by year, and for each year, all items loaded to a single component, on a scale from 1 to 4, where 4 indicates more positive climate. For our analyses, we calculated the change in reported *School climate* from the baseline year to the second outcome year.

Our final measure, *Postpolicy success*, asked respondents “How has your school's success with each of the following changed, as a result of this new law?” All seven items (reducing racial gaps in student discipline, reducing reliance on exclusionary discipline, improving attendance, reducing misbehavior, ensuring a safe and orderly environment, improving student achievement overall, and reducing student achievement gaps) loaded to a single component. These items were on a scale from 1 to 5 (where 5 indicates *much more successful than before*), which allowed for the inclusion of a middle category for *about the same*. As a result, the interpretation of the results for this outcome is slightly different than for the other outcomes.

Survey Data Collection

Between May 28 and November 4 of 2019,⁹ we sent an email-based Qualtrics survey to all elementary school principals (611) and counselors (707) in Arkansas (some schools have more than one counselor). We received 235 responses from principals (an initial response rate of 38.5%), and 155 responses from counselors (an initial response rate of 21.9%).

For Research Question 1, our primary interest was to learn from individuals who at the time of the survey were working in the same school since the baseline year, 2016–2017, so we exclude responses from individuals who are not able to compare conditions pre- and postpolicy and make a few other sample restrictions.¹⁰ There was nontrivial item-level missingness, so to make our samples comparable across models, we focus on a sample of 110 schools for whom we have data on all measures. We primarily use principal responses, only incorporating a counselor's response where a principal response was not available (14 of the 110 cases). The results are generally robust to the exclusion of these counselor responses. Due to these sample restrictions, we have a sample of individuals with relative job stability

and low item-level nonresponse. These individuals, and the schools they represent, may not be representative of the state.

Relatedly, for Research Question 2, the sample is restricted to schools for which we were able to calculate the outcome measures, requiring discipline infractions overall, as well as subjective infractions, to be reported in the baseline year, 2016–2017, as well as each of the first two outcome years, 2017–2018 and 2018–2019. As a result, Research Question 2 is only estimated among the set of schools who were regularly reporting disciplinary data, and specifically subjective infractions, across this time period.

In Table 3, we compare the characteristics of sample schools with other Arkansas elementary schools. The schools for which we have complete survey measures are less likely to be located in cities, have a greater share of White students, and lower shares of Hispanic and limited English proficient students. Infraction counts are higher in the responding schools. To address these differences, our Research Question 1 analysis weights the survey data by the inverse probability of response (see the Analytic Approach section). There may be unobservable factors not accounted for in the weights. For example, principals concerned about job security or accountability might be more reluctant to respond or more subject to social desirability bias. To minimize these concerns, we communicated that survey responses would be “kept confidential to the extent allowed by law and University policy,” that “all responses will be summarized . . . such that no individuals, schools, or districts can be identified,” and that participation was voluntary.

Table 3 also shows that the schools in the Research Question 2 sample tend to be larger, are less likely to be in the Northwest, and more likely to be in the Central region. They tend to serve a greater share of Black and FRL students but fewer students of other races (i.e. not White, Black, or Hispanic) and limited English proficient students. Average achievement in the sample schools is slightly below the state average. Since the sample restriction is based on the availability of infraction reports across all 3 years, the schools included in this sample tend to report higher numbers of infractions and consequences of all types. The included schools use corporal punishment at lower rates than the excluded schools.

Analytic Approach

To address our research questions, we use a series of multivariate regression models. All findings are descriptive, not causal.

Research Question 1: What School and District Characteristics Are Associated With Principal and Counselor Perspectives About Act 1059 and Its Implementation? For Research Question 1, we test which school- and district-level

characteristics are associated with seven key measures based on survey responses. Two measures are attitudes/beliefs about the law (*Support for the law* and *Clarity for implementation*), one is a binary¹¹ indicator of moderate or substantial difficulties with implementation, two are reports of *Challenges* with implementation, one is a measure of change in *School climate*, and one is reported *Postpolicy success*. See online Supplemental Appendix Table A. We predict these outcomes as a function of observable baseline school characteristics, following:

$$Y_s = \beta_0 + \mathbf{X}_s \beta_1 + \varepsilon_s \quad (1)$$

where Y_s is one of the seven outcomes, and \mathbf{X}_s is a vector of 2016–2017 school characteristics, for school s , including the log of K–5 student enrollment, school percent non-White, school percent FRL-eligible, school percent special education, average math and ELA test scores (estimated as z-scores and averaged together), and district-level per-pupil expenditures. This vector also includes baseline discipline outcomes, calculated over three baseline years for added stability: the number of infractions per 100 students per year, and the share of infractions that resulted in exclusionary consequences (OSS, expulsion, or referral to an alternative learning environment). All of these variables—except school percent special education—are theorized to relate to the degree of noncompliance, and are considered variables of interest. Standard errors, clustered at the district level, are indicated by ε_s . Observations are weighted by the inverse propensity to respond to the survey.¹²

Research Question 2: What School and District Characteristics Are Associated With Compliance With Act 1059 as Reported in the Administrative Data? Next, we estimate a model similar to Equation (1), but where Y_s is one of four compliance-related outcomes, each based on the share of infractions (overall or subjective) that result in OSS or expulsion. Again, we theorize that most predictor variables—except school percentage of special education—would relate to the degree of noncompliance, and are considered variables of interest. We hypothesize declines will be concentrated among subjective infractions.

Student discipline policy is often set at a district level through a district code of conduct. As a result, we estimate another set of results with district fixed effects (FE), rather than region FE, to explore whether our results differ when identifying the relationships using only within-district variation. The models with region FE remain our preferred analyses, as only 57 of the 175 districts represented in our preferred analyses had more than one elementary school represented, so within-district variation is limited and only available for a distinct set of larger school districts.

TABLE 3
Characteristics of Sample Schools, Compared With Other Schools

	Research Question 1 preferred samples	Other schools	Difference	Research Question 2 preferred samples	Other schools	Difference
<i>School/district context</i>						
Number of schools	110	514		342	282	
K–5 student enrollment	373.6	362.9	10.7	383.6	342.0	41.6***
Northwest region	33.6%	34.8%	–1.2%	27.2%	43.6%	–16.4%***
Northeast region	20.9%	19.6%	1.3%	20.5%	19.1%	1.3%
Central region	23.6%	29.2%	–5.5%	34.2%	20.9%	13.3%***
Southwest region	14.5%	10.3%	4.2%	10.2%	12.1%	–1.8%
Southeast region	7.3%	6.0%	1.2%	7.9%	4.3%	3.6%*
City	18.2%	30.1%	–11.9%**	26.3%	30.0%	–3.7%
Suburb	11.8%	10.0%	1.9%	12.3%	7.9%	4.4%*
Town	25.5%	21.7%	3.8%	24.6%	19.6%	4.9%
Rural	44.5%	38.3%	6.3%	36.8%	42.5%	–5.7%
Per pupil district expenditures in 2016–2017	5598.8	5514.1	84.7	5559.8	5491.5	68.3
<i>Student characteristics and achievement</i>						
School % White	68.6%	60.7%	7.9%**	59.1%	65.7%	–6.6%***
School % Black	18.0%	21.5%	–3.5%	25.7%	15.1%	10.5%***
School % Hispanic	8.5%	12.5%	–4.1%***	10.5%	13.4%	–2.9%***
School % other races	5.0%	5.3%	–0.3%	4.8%	5.8%	–1.0%**
School % FRL	65.3%	67.7%	–2.4%	69.8%	64.2%	5.6%***
School % special education	13.7%	13.0%	0.7%	13.3%	13.0%	0.3%
School % LEP	5.0%	9.0%	–4.0%***	6.9%	10.0%	–3.1%***
Average math z-score	0.007	–0.055	0.062	–0.077	0.004	–0.080**
Average ELA z-score	0.012	–0.047	0.059	–0.071	0.016	–0.087***
<i>Number of infractions per 100 students (2016–2017)</i>						
All infractions	47.1	34.7	12.5***	48.3	23.1	25.2***
Subjective infractions	23.0	17.3	5.7*	25.3	9.9	15.4***
Violence-weapon-related infractions	5.2	3.9	1.3**	5.2	2.8	2.4***
All remaining infractions	18.9	13.5	5.4**	17.8	10.4	7.4***
<i>Number of consequences per 100 students by infraction type (2016–2017)</i>						
OSS or expulsion for any infraction	6.7	6.2	0.5	8.8	3.2	5.6***
OSS or expulsion for subjective infractions	3.2	2.7	0.4	4.2	1.1	3.0***
OSS or expulsion for violence/weapon infraction	1.7	1.7	0.0	2.3	1.0	1.3***
ISS for any infraction	13.1	9.9	3.2	14.5	5.7	8.8***
ISS for subjective infractions	5.6	4.9	0.7	7.6	2.0	5.6***
ISS for violence/weapon infractions	1.6	1.0	0.6**	1.4	0.7	0.7***
<i>Share of 2016–2017 infractions resulting in consequence types</i>						
OSS or expulsion	25.0%	28.2%	–3.2%	28.6%	26.4%	2.2%
ISS	26.4%	23.4%	3.0%	24.1%	23.7%	0.4%
ALE	0.1%	0.1%	0.0%	0.2%	0.1%	0.1%
Corporal punishment	16.8%	15.2%	1.6%	14.0%	17.7%	–3.7%**
No action/warning	4.1%	4.9%	–0.8%	4.7%	4.9%	–0.2%
Other action	27.5%	28.0%	–0.5%	28.5%	27.3%	1.2%

Note. “Research Question 1 preferred samples” are the 110 schools for which all survey constructs used in the Research Question 1 analyses were available (Table 5). More schools responded partially, but we restricted the sample so that it is consistent and comparable across models. “Research Question 2 preferred samples” are the 342 schools for which all measures used for Research Question 2 (Tables 6 and 7) were available. LEP = limited English proficiency; ISS = in-school suspension; OSS = out-of-school suspension. ELA = English language art; FRL = free or reduced-price lunch; ALE = alternative learning environment. * $p < .1$. ** $p < .05$. *** $p < .01$.

Results

Research Question 1: What School and District Characteristics Are Associated With Principal and Counselor Perspectives About Act 1059 and Its Implementation?

We report bivariate correlations in Table 4. School percent non-White is negatively correlated with *Support for the law* ($r = -0.36$), and *Clarity for implementation* ($r = -0.24$). Baseline use of exclusionary discipline is negatively correlated with *Clarity for implementation* ($r = -0.27$). Both types of *Challenges with implementation* are positively correlated with school percent non-White ($r = 0.36$ and $r = 0.44$, respectively) and baseline use of exclusionary discipline ($r = 0.38$ and $r = 0.29$, respectively). In other words, whiter schools and schools that relied less on exclusionary discipline indicated more *Support for the law*, greater *Clarity for implementation*, and fewer *Challenges*. Support and capacity for implementation may be unevenly distributed in ways that exacerbate educational inequities for students of color, in particular Black students, who attend more exclusionary schools (Anderson & Ritter, 2017). In the set of results that follow, we assess the relationships between these and other variables using multivariate regression, controlling for observable school characteristics.

In Table 5, we present the results of our preferred specifications investigating Research Question 1. Column 1 shows that smaller schools indicated greater *Support for the law*. A 10% greater school enrollment is associated with -0.026 unit (-0.04 *SD*) lower *Support for the law*. Furthermore, while imprecise and not statistically distinguishable from zero, the effect size on school percent FRL-eligible indicates that respondents from less advantaged schools reported less *Support for the law*.

The outcome in column 3 is a binary indicator of *Moderate or substantial difficulties* with implementation (relative to minimal or no difficulties) in the first two outcome years. Larger schools, lower performing schools, more exclusionary schools, and schools in lower spending districts were more likely to report difficulties. Some of these estimates are imprecise. For example, a 10 percentage point higher baseline reliance on exclusionary discipline is associated with a 3.96 percentage point higher likelihood of reporting moderate or substantial difficulties, although the 95% confidence interval implies a relatively wide range from 0.3 percentage points to 7.6 percentage points. Similarly, the next two columns show how school characteristics related to implementation *Challenges*. Schools with a greater share of non-White students, schools with greater baseline use of exclusionary discipline, lower performing schools, and schools in lower spending districts indicated greater *Challenges* of various types, reiterating that relatively disadvantaged schools reported more difficulties.

Finally, column 7 indicates that schools serving more FRL students reported less *Postpolicy success*, although this

estimate is relatively imprecise. None of the independent variables were significantly related to *Clarity for implementation* (column 2) or *Improvement in School Climate* (column 6), controlling for other factors.

Specification Checks. We test the robustness of our results to a variety of specifications: (1) using non-Winsorized versions of the baseline discipline use variables as well as the measures of change in exclusionary discipline use, (2) including district FE instead of region FE, (3) restricting the analysis to only principal respondents (see online Supplemental Appendix Table B), and (4) using all available data and allowing the sample size to vary by model (see online Supplemental Appendix Table C). The models with the non-Winsorized versions of these variables are nearly identical.¹³ The district FE models indicate no statistically significant relationships, as there is little within-district variation among the 110 schools included in this sample (i.e., only 49 of these 110 schools were in a district with more than one school in the sample), and as a result, statistical power suffered.¹⁴

The results in online Supplemental Appendix Table B are substantively similar to the main results in Table 5. In online Supplemental Appendix Table C, while the directional effects were generally similar, there are several cases in which statistical significance is different from Table 5, due to relatively imprecise estimates in both sets of models. In online Supplemental Appendix Table C, we find stronger evidence that high-FRL schools reported lower *Support for the Law* and that high exclusionary schools (at baseline), reported lower *Clarity for implementation*. In column 3, school size and baseline use of exclusionary discipline are no longer statistically significant. In column 4, baseline test scores and per-pupil expenditures are no longer significant. In column 5, per-pupil expenditures was no longer significant, and we find significantly higher *Challenges* for schools with greater baseline exclusionary discipline use. Finally, in column 6, we find that schools with greater baseline use of exclusionary discipline reported larger *School climate* gains (or smaller climate declines). Notably, all these differences were in terms of statistical significance, and not the direction of the relationship, so overall these generally support the main findings in Table 5: relatively disadvantaged schools (based on a variety of indicators), typically had greater difficulties and *Challenges* with implementation, and less *Postpolicy success*.

Testing Interactions Between School Demographics and District-Level Per-Pupil Expenditures. In general, we find that schools in districts with lower per-pupil instructional expenditures, schools with greater proportions of non-White or FRL-eligible students, and schools with higher baseline use of exclusionary discipline reported more *Challenges*. It is possible that there is a compounding or interaction effect of disadvantage for less well-resourced schools who also serve more low-income students and students of color, or

TABLE 4
Correlations Between Survey Constructs and Observable School and District Characteristics at Baseline

Measure/characteristic	Survey measures															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>Survey measures</i>																
1. Support for the law	1.00															
2. Clarity for implementation	0.63	1.00														
3. School climate (2016–2017)	0.08	-0.03	1.00													
4. School climate (2018–2019)	0.28	0.30	0.59	1.00												
5. Change in school climate (2016–2017 to 2018–2019)	0.21	0.36	-0.46	0.45	1.00											
6. Challenges: lack of resources and capacity	-0.44	-0.47	-0.35	-0.43	-0.08	1.00										
7. Challenges: politics, policies, and leadership	-0.45	-0.44	-0.17	-0.22	-0.05	0.59	1.00									
8. Postpolicy success	0.45	0.45	-0.15	0.29	0.48	-0.24	-0.19	1.00								
<i>Administrative data</i>																
9. K–5 Enrollment	-0.17	-0.04	0.03	-0.09	-0.14	0.08	0.15	-0.02	1.00							
10. % non-White	-0.36	-0.24	-0.21	-0.20	0.02	0.36	0.44	-0.14	0.26	1.00						
11. % FRL	-0.24	-0.05	-0.27	-0.19	0.09	0.10	0.08	-0.18	-0.26	0.31	1.00					
12. % special education	0.14	0.16	-0.25	-0.10	0.17	-0.10	0.00	0.12	-0.19	-0.20	0.27	1.00				
13. Average z-score (math and ELA combined)	0.34	0.12	0.39	0.24	-0.17	-0.33	-0.29	0.06	0.09	-0.52	-0.69	-0.14	1.00			
14. District per-pupil instructional expenditures	0.00	-0.04	-0.06	0.02	0.09	0.00	-0.03	-0.10	-0.02	0.26	0.18	0.16	-0.01	1.00		
15. Infractions per year per 100 students (past 3 years)	-0.21	0.06	-0.24	-0.19	0.05	-0.02	0.03	0.01	-0.09	0.17	0.30	0.05	-0.27	-0.10	1.00	
16. Share of infractions resulting in OSS, expulsion, or ALE (past 3 years)	-0.18	-0.27	-0.05	-0.01	0.05	0.38	0.29	-0.10	0.28	0.47	-0.16	-0.21	-0.08	0.28	-0.36	1.00

Note. Correlation shown are for the set of 110 schools with complete survey-data used for Research Question 1. All administrative data are from the prepolicy baseline year (2016–2017), except for the disciplinary measures (Nos. 15 and 16), which are summarized across 3 baseline years (2014–2015 to 2016–2017). Administrative data are all school-level except for district-level per-pupil instructional expenditures. OSS = out-of-school suspension; ELA = English language art; FRL = free or reduced-price lunch; ALE = alternative learning environment.

TABLE 5
Predicting Survey Responses as a Function of School and District Characteristics

Characteristic	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Support for the law	Clarity for implementation	Binary: moderate/substantial difficulties (17–18 and 18–19 combined)	Challenges: lack of resources, capacity	Challenges: politics, leadership	Improvement in School Climate (16–17 to 18–19)	Postpolicy success
Log of school enrolled in Grades K–5	-0.257*** (0.097)	0.033 (0.097)	0.156** (0.065)	0.096 (0.111)	0.062 (0.088)	-0.043 (0.099)	-0.019 (0.100)
School % non-White	0.235 (0.318)	-0.134 (0.324)	-0.364 (0.225)	0.337 (0.330)	0.982*** (0.355)	-0.121 (0.216)	-0.101 (0.352)
School % FRL-eligible	-0.838 (0.543)	-0.392 (0.519)	-0.056 (0.339)	0.036 (0.530)	0.217 (0.488)	-0.356 (0.463)	-1.258** (0.549)
School % special education	2.465* (1.345)	1.367 (1.264)	-2.451** (0.986)	-2.408 (1.775)	2.086 (1.347)	0.959 (0.819)	2.865*** (1.366)
Average z-score (ELA and math combined)	0.306 (0.325)	-0.038 (0.323)	-0.689*** (0.229)	-0.645** (0.319)	-0.072 (0.314)	-0.287 (0.249)	-0.256 (0.294)
Number of infractions per year per 100 students	-0.002 (0.001)	0.000 (0.001)	0.001 (0.001)	0.002 (0.002)	0.000 (0.002)	0.000 (0.001)	0.001 (0.001)
Share infractions resulting in OSS, ALE, or expulsion	-0.213 (0.219)	-0.177 (0.172)	0.396** (0.186)	0.700** (0.294)	0.267 (0.303)	0.179 (0.170)	0.043 (0.348)
Per-pupil instructional expenditures (thousands)	0.017 (0.060)	0.059 (0.082)	-0.084* (0.046)	-0.246** (0.098)	-0.211** (0.091)	0.029 (0.045)	-0.072 (0.087)
Constant	4.190*** (0.616)	2.252*** (0.836)	0.165 (0.439)	3.508*** (0.956)	1.546** (0.738)	0.240 (0.717)	3.998*** (0.669)
Observations	110	110	110	110	110	110	110
R ²	.280	.160	.503	.403	.373	.069	.125
Adjusted R ²	.191	.056	.441	.329	.295	-.046	.017

Note. Includes region fixed effects for the five regions in the state. Heteroskedastic-robust standard errors, clustered at the district level, in parentheses. All models are weighted by the inverse probability of survey response. The number of infractions per year per 100 students and the share of infraction resulting in OSS, ALE, or expulsion are both based on a 3-year average (the last three years prepolicy). OSS = out-of-school suspension; ELA = English language art; FRL = free or reduced-price lunch; ALE = alternative learning environment.
 * $p < .1$. ** $p < .05$. *** $p < .01$.

those with higher baseline use of exclusionary discipline. We tested this by including interactions, one at a time, in the same models shown in Table 5, for the *Challenges* outcomes. The results are available in online Supplemental Appendix Table D. Most of these interactions were not statistically significant, but there is an interaction regarding *Challenges related to lack of resources or capacity* in which we find that schools serving more non-White students and located in lower spending districts reported these types of *Challenges* to a greater degree, providing some support for an interaction effect.

Research Question 2: What School and District Characteristics Are Associated With Compliance With Act 1059 as Reported in the Administrative Data?

We model compliance as measured by four measures of change in the use of OSS/expulsion from the administrative data (see the section on administrative data measures). To be included in the preferred sample (held constant across all models), schools had to report disciplinary incidents, including subjective infractions, in the baseline year and both post-policy years. The results are in Table 6. The first four columns include region FE, and the final four include district FE. Positive coefficients indicate a factor is associated with increases in OSS/expulsion (i.e., a greater degree of noncompliance).

In columns 1 to 4 (with region FE), but not when only comparing within district (columns 5–8), we find that schools serving a greater share of non-White students complied at lower rates. Within district (columns 5–8), however, there is evidence that schools serving more FRL-students tended to comply at lower rates. The most consistently significant predictor of compliance in this table is higher reliance on exclusionary discipline at baseline. This result might reflect simply that high-exclusionary schools had more room for reducing this practice. Specifically, although the outcome variables were already Winsorized at the 95th percentile, we theorized that there could be a floor effect whereby schools with very low baseline use of exclusionary discipline have nearly no room to decrease their use, but could easily increase use at very high rates, due to the low base rate. Accordingly, we identified a set of 25 outliers with reported increases in use of OSS/expulsion of greater than 400%. These schools tend to have very low baseline use of exclusionary discipline. When these outliers are dropped, the coefficient on baseline exclusionary discipline use is not statistically significant, indicating this result was driven by these outliers, rather than by previously high-exclusion schools complying at higher rates. These results are available in Table 7. Another notable difference from Table 6, is that when dropping these outliers, we have less evidence that school percent non-White is negatively correlated with compliance, and more evidence that school percent FRL-eligible

is negatively correlated with compliance in the district FE models. On the whole, the results indicate that relatively disadvantaged schools were complying at lower rates.

Specification Checks. We test the robustness of our results in Table 7 (after dropping outliers) using alternative specifications. We loosen the restriction to keep the samples consistent across models, and the results (see online Supplemental Appendix Table E) are substantively similar to those in Table 7, with some differences in statistical significance. In particular, online Supplemental Appendix Table E more frequently finds a statistically significant relationship between school percentage of non-White and noncompliance, in the models with region FE. We also test the results to using non-Winsorized versions of baseline discipline use and the results are nearly identical.¹⁵

Discussion

In this article, we investigated perspectives on a discipline policy reform, and its implementation, across varying school contexts. Although the results are often noisy and imprecise, we generally find that relatively disadvantaged schools and more exclusionary schools reported greater difficulties and challenges with implementation. Overall, about one third of survey respondents indicated moderate or substantial difficulties with implementation. The most commonly cited challenges were a lack of mental health supports/counseling resources, significant discipline challenges, a lack of feasible disciplinary alternatives, and a lack of resources and/or time.

This study identified some equity concerns regarding the outcomes of this policy. Schools with a greater share of FRL students were less likely to report postpolicy success, and schools with greater shares of disadvantaged students reduced their use of exclusionary discipline at lower rates, controlling for other factors. Notably, although more exclusionary schools reported greater difficulties and challenges with implementation, there was no evidence to suggest they were complying at lower rates. Still, with more support or training to address some of these challenges, the reductions in exclusionary discipline may have been even greater. Although exclusionary discipline decreased overall, OSS and expulsion was still the response to almost 9,000 incidents in 2018–2019, representing approximately 5,000 elementary-age students suspended or expelled at least once that year.

Policy makers seeking to reduce disproportionalities in student discipline should take these findings to heart. There are key equity concerns for low-income students and students of color who were concentrated in schools that were facing greater challenges with implementation. In particular, schools in lower spending districts with more students of color were reporting greater challenges related to capacity

TABLE 6
Predicting Compliance Based on Administrative Data as a Function of School and District Characteristics

Characteristic	Percentage change in share resulting in OSS/expulsion							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log of school enrolled in Grades K-5	0.040 (0.101)	0.081 (0.092)	-0.050 (0.099)	0.056 (0.099)	-0.098 (0.179)	-0.035 (0.228)	-0.127 (0.171)	0.017 (0.214)
School % non-White	0.217 (0.246)	0.592** (0.286)	0.600** (0.266)	0.769*** (0.288)	-0.990 (1.147)	0.192 (0.740)	-1.339 (0.869)	0.133 (0.969)
School % FRL-eligible	-0.036 (0.411)	-0.287 (0.424)	-0.204 (0.480)	-0.800* (0.434)	1.306 (0.837)	1.250 (0.808)	2.124** (0.884)	1.641* (0.856)
School % special education	-0.226 (1.050)	-1.077 (1.263)	-1.004 (1.132)	-0.167 (1.396)	-0.366 (1.725)	-1.661 (2.261)	-1.038 (1.682)	-1.643 (2.232)
Average z-score (ELA and math combined)	0.094 (0.220)	-0.154 (0.243)	0.103 (0.224)	-0.305 (0.248)	0.383 (0.407)	0.689 (0.458)	0.262 (0.470)	0.621 (0.451)
Number of infractions per year per 100 students	0.001 (0.001)	-0.002* (0.001)	0.001 (0.001)	-0.001 (0.001)	0.003 (0.002)	-0.001 (0.003)	0.001 (0.003)	-0.002 (0.003)
Share infractions resulting in OSS, ALE, or expulsion	-0.483***-0.204	-0.790***-0.269	-0.350-0.242	-0.482*-0.263	-0.701-0.507	-0.577-0.54	-0.945**-0.471	-0.900**-0.403
Per-pupil instructional expenditures (thousands)	0.012 (0.070)	0.040 (0.085)	0.079 (0.082)	0.079 (0.090)				
Constant	-0.219 (0.761)	-0.438 (0.792)	-0.207 (0.819)	-0.572 (0.888)	-0.559 (1.262)	-0.008 (1.712)	-0.390 (1.229)	-0.754 (1.582)
Region FE	Y	Y	Y	Y	Y	Y	Y	Y
District FE					Y	Y	Y	Y
Observations	342	342	342	342	342	342	342	342
R ²	.057	.116	.053	.103	.715	.678	.690	.664
Adjusted R ²	.023	.084	.019	.070	.393	.314	.340	0.284

Note. Heteroskedastic-robust standard errors, clustered at the district level, in parentheses. The outcome measures in columns 1 to 4 are percentage changes such that a one unit increase in the outcome corresponds to a 100% increase. The number of infractions per year per 100 students and the share of infraction resulting in OSS, ALE, or expulsion are both based on a 3-year average (the past 3 years prepolicy). OSS = out-of-school suspension; ELA = English language art; FRL = free or reduced-price lunch; ALE = alternative learning environment; FE = fixed effects.
 * $p < .1$. ** $p < .05$. *** $p < .01$.

TABLE 7
Sensitivity Analysis for Research Question 2: Dropping Outlier Schools With Extremely High Increases in OSS/Expulsion Use Postpolicy

Characteristic	Percentage change in share resulting in OSS/expulsion							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log of school enrolled in Grades K-5	-0.013 (0.084)	0.058 (0.073)	-0.036 (0.085)	0.093 (0.091)	-0.175 (0.155)	-0.116 (0.195)	-0.095 (0.156)	-0.003 (0.204)
School % non-White	0.210 (0.206)	0.389 (0.259)	0.353 (0.237)	0.455* (0.262)	-0.836 (0.849)	-0.180 (0.704)	-1.549** (0.678)	-0.590 (0.754)
School % FRL-eligible	0.150 (0.400)	-0.242 (0.364)	0.308 (0.516)	-0.400 (0.404)	1.570* (0.818)	1.529** (0.755)	2.492*** (0.850)	2.127** (0.883)
School % special education	0.117 (0.918)	-0.848 (1.122)	-0.405 (1.113)	-0.089 (1.277)	1.680 (1.476)	-0.505 (1.585)	0.500 (1.683)	-0.674 (1.884)
Average z-score (ELA and math combined)	0.249 (0.206)	-0.113 (0.219)	0.327 (0.221)	-0.138 (0.231)	0.587* (0.332)	0.714 (0.450)	0.452 (0.429)	0.674 (0.482)
Number of infractions per year per 100 students	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.002)	-0.002 (0.003)	0.001 (0.002)	-0.002 (0.003)
Share infractions resulting in OSS, ALE, or expulsion	-0.140 (0.177)	-0.367* (0.215)	0.013 (0.203)	-0.078 (0.217)	-0.651 (0.517)	-0.465 (0.457)	-0.493 (0.484)	-0.365 (0.542)
District per-pupil instructional expenditures (thousands)	-0.006 (0.052)	0.016 (0.050)	0.016 (0.057)	0.026 (0.062)				
Constant	-0.189 (0.625)	-0.391 (0.603)	-0.437 (0.681)	-0.877 (0.754)	-0.414 (1.084)	0.302 (1.595)	-0.962 (1.214)	-0.950 (1.529)
Region FE	Y	Y	Y	Y				
District FE					Y	Y	Y	Y
Observations	317	317	317	317	317	317	317	317
R ²	.050	.096	.052	.086	.724	.640	.720	.631
Adjusted R ²	.012	.060	.014	.050	.411	.232	.401	.213

Note. Corresponds to Table 6, but dropping 25 outlier schools with very high measures on one or more of the dependent variables. Heteroskedastic-robust standard errors, clustered at the district level, in parentheses. The number of infractions per year per 100 students and the share of infraction resulting in OSS, ALE, or expulsion are both based on a 3-year average (the past 3 years prepolicy). OSS = out-of-school suspension; ELA = English language art; FRL = free or reduced-price lunch; ALE = alternative learning environment; FE = fixed effects.
* $p < .1$. ** $p < .05$. *** $p < .01$.

and lack of resources. Furthermore, there is not a clear mechanism for the state to apply pressure on districts to comply, particularly given the ambiguous language that continues to allow exceptions to the suspension and expulsion ban. As McLaughlin (1987) suggested, fidelity of policy implementation is supported by a balance of pressure (i.e., accountability) and support (i.e., resources and capacity). Policy implementation suffers when there is more conflict with local practice, or when requirements outweigh the capabilities/resources of local actors (Belansky et al., 2009; Cohen et al., 2007; Elmore, 1979–1980; Matland, 1995; Mazmanian & Sabatier, 1983; Van Meter & Van Horn, 1975). Our findings reiterate that policy makers should design policies with these potential challenges and conflicts in mind, which may mean using clear (less ambiguous) language regarding what schools are and are not allowed to do, communicating and collaborating with districts prior to policy change in order to prevent unintended consequences, providing resources, trainings, or support to find alternative approaches, and following up with schools and districts that continue not to comply.

This study has important limitations. First, all analyses are descriptive and do not indicate causal relationships. Second, this is a first look at the implementation of this policy and is not assessing its impact on more distal student outcomes. Third, the administrative data do not include the full extent of the information that likely was recorded at the time of the incident, so other than infraction type (e.g., disorderly conduct or fighting) we do not have textual descriptions to determine whether incidents met the exceptions outlined in Act 1059 related to posing a “physical risk” or behaviors that “cannot be addressed through other means.” Fourth, survey responses reported by principals and counselors may be subject to a variety of biases. The use of retrospective survey data may be subject to misremembering (Pearson et al., 1992), and the responses may be affected by social desirability bias (Nederhof, 1985), even when anonymity is assured. Fifth, several sample restrictions were necessary, resulting in a sample that is not fully representative of the state.

As a further limitation, we focused on principals and counselors as a convenience sample for which contact information was readily available to the researchers. The perspectives of teachers, students, families, and community members are missing from this study and would be a worthy population to include in future work. Broadly, questions remain about how implementation actually takes place in the field, and suggest that mixed-methods research would be an important direction for future research in this area.

Furthermore, our study was unable to dive deeply into the ecological factors that may influence those opinions or preferences, which could be explored in future work. For example, within a bioecological systems framework (Bronfenbrenner & Morris, 1998), the individual (i.e.,

biosystem) is located within a microsystem in which that individual regularly participates (e.g., home or workplace). At the mesosystem level, there are interactions between different microsystems. The exosystem encompasses formal and informal structures surrounding the respondents that have an indirect influence such as school boards, police and court systems, and community organizations. In the macrosystem, broader influences such as social and political ideologies or cultural beliefs can influence perceptions of reform. Finally, changes may occur over time due to life transitions or historical events, represented by the chronosystem (e.g., growing awareness of systemic racism in the United States). As such, future work might incorporate a more comprehensive view of the entire ecological system. Similarly, while it was out of the scope of the current study, future work might apply and test theories of policy adoption and diffusion (e.g., punctuated equilibrium, multiple streams) (see McLendon & Cohen-Vogel, 2012).

While descriptive, this study provides important findings about how school characteristics are related to principal and counselor perspectives on and implementation of discipline reform. High-level policies may be enacted without recognition of local needs and preferences. Greater attention to local needs, buy-in, and clarity for implementation are key. Policy makers seeking more equitable implementation need to communicate with schools and local communities about their discipline practices, the purpose of reform, and the resources and supports needed to make such reform feasible, practical, and impactful.

Acknowledgments

We would like to thank Raquel Sosa for her contributions to this project.

ORCID iD

Kaitlin P. Anderson  <https://orcid.org/0000-0002-8445-7352>

Open Practices

The survey protocols and analysis files for this article can be found at <https://www.openicpsr.org/openicpsr/project/157942/version/V1/view>

Notes

1. Connecticut and New Jersey place limits on exclusionary discipline in pre-K through Grade 2, Maryland, and New Jersey limit the use of suspension or expulsion in Grades K–2, California, Ohio, and Texas place limits in Grades K–3, and Arkansas, Louisiana, and Oregon place limits in Grades K–5. The District of Columbia limits exclusionary discipline from prekindergarten to Grade 8, and Kentucky limits its use in “primary school” (Rafa, 2018).

2. For example, Texas and Connecticut allow suspension if students have engaged in an offense related to weapons or violence, or if they were selling, possessing, giving or under the influence

of drugs (CT Gen. Stat § 10-233c, 2012; Texas Education Code § 37.005. Suspension, 2020). Maryland allows suspension “for not more than 5 school days if the school administration, in consultation with a school psychologist or other mental health professional, determines that there is an imminent threat of serious harm” (MD Educ Code § 7-305.1, 2017). New Jersey allows suspensions for students whose conduct is of violent or sexual nature that endangers others or if they possess a firearm on school property (NJ Rev Stat § 18A:37-2A, 2016).

3. Prior to the passage of Act 1329 of 2013, section 6-18-507 of the Arkansas Code stated: “The board of directors of a school district may suspend or expel any student from school for violation of the school district’s written discipline policies,” but was amended by Act 1329 of 2013 to include: “except that a school district shall not use out-of-school suspension as a discipline for truancy,” and was further amended by Act 1059 of 2017 to additionally state: “The school district shall not use out-of-school suspension or expulsion for a student in kindergarten through grade five (K–5) except in cases when a student’s behavior: (A) Poses a physical risk to himself or herself or to others; or (B) Causes a serious disruption that cannot be addressed through other means.” Each district is required by Section AR Code § 6-18-503 to develop written student discipline policies, and if authorizing corporal punishment, this written policy must include provisions for administration of this type of punishment, “including that it be administered only for cause, be reasonable, follow warnings that the misbehavior will not be tolerated, and be administered by a teacher or school administrator and only in the presence of a school administrator or his or her designee, who shall be a teacher or school administrator employed by the school district.”

4. Although our focus is on disciplinary responses in Grades 3 to 5, we use test scores in Grades 3 to 8 as predictor variable to better capture the overall level of achievement in the school.

5. This “other” category was not researcher created. Instead, reporting categories used at the local level did not have a corresponding category at the state reporting level.

6. “Other nonspecified” infractions historically have not rolled up into a state-level reporting category. Starting in 2016–2017, the state started reporting additional categories (terroristic threats, stealing/theft, harassment/sexual harassment, cellphones/electronic devices, cyberbullying, public display of affection), but for comparability of reporting across years, these remain in the “other nonspecified” category.

7. The results are nearly identical with and without Winsorization.

8. Also, there are exceptions (e.g., Bear et al., 2011; Bobbett & French, 1991; Brand et al., 2003; Butler & Rakow, 1995; Johnson et al., 1987; Johnson et al., 1997; Zullig et al., 2010), but often, the instruments are not free and publicly available.

9. The survey was initially sent on May 28th, and approximately 66% were received in May, June, or July. A reminder email was sent October 11th, and the remaining 34% of responses were in October or November.

10. The sample includes traditional public and charter schools but excludes virtual schools and schools operated by the Division of Youth Services, dropping 7 schools. 24 schools that did not enroll students in the 2016–2017 school year are excluded.

11. The model predicting this binary outcome uses linear probability models.

12. Regression weights were created using a probit model predicting whether a school responded to the survey as a function of

observable baseline characteristics of the school including the log of student enrollment, district per-pupil instructional expenditures, school percent Black, School percent Hispanic, school percent other non-White races, school percent FRL-eligible, School percent in special education, school average achievement (math and ELA combined), the number of infractions per 100 students at baseline, and the share of infractions resulting in exclusionary discipline at baseline. Specifically, for our main specifications, we only count as a response the 110 schools with complete survey data, and for alternative specifications, we count as a response any school that responded to the survey. Originally, indicators for region and location type were also included, but these were not jointly significant, so they were dropped from these models creating these weights. The results of the probit model used for the creation of weights are available from the authors by request.

13. Table of results available from the authors by request.

14. Table of results (all null) available from the authors by request.

15. The models without the Winsorized versions of these variables are available from the authors, by request.

References

- Agnew, R., & Brezina, T. (2010). Strain theories. In E. McLaughlin, & T. Newburn (Eds.), *The SAGE handbook of criminological theory* (pp. 96–113). Sage. <https://doi.org/10.4135/9781446200926.n6>
- American School Counselor Association. (2020). *The school counselor and discipline*. <https://schoolcounselor.org/Standards-Positions/Position-Statements/ASCA-Position-Statements/The-School-Counselor-and-Discipline>
- Anderson, K. P. (2018). Inequitable compliance: Implementation failure of a statewide student discipline reform. *Peabody Journal of Education*, 93(2), 244–263. <https://doi.org/10.1080/0161956X.2018.1435052>
- Anderson, K. P. (2020). Academic, attendance, and behavioral outcomes of a suspension reduction policy: Lessons for school leaders and policy makers. *Educational Administration Quarterly*, 56(3), 435–471. <https://doi.org/10.1177/0013161X19861138>
- Anderson, K. P. (2021). The relationship between inclusion, absenteeism, and disciplinary outcomes for students with disabilities. *Educational Evaluation and Policy Analysis*, 43(1), 32–59. <https://doi.org/10.3102/0162373720968558>
- Anderson, K. P., Egalite, A. J., & Mills, J. N. (2019). Discipline reform: The impact of a statewide ban on suspensions for truancy. *Journal of Education for Students Placed at Risk (JESPAR)*, 24(1), 68–91. <https://doi.org/10.1080/10824669.2018.1537794>
- Anderson, K. P., & Ritter, G. W. (2017). Disparate use of exclusionary discipline: Evidence on inequities in school discipline from a U.S. state. *Education Policy Analysis Archives*, 25, 49. <https://doi.org/10.14507/epaa.25.2787>
- Anderson, K. P., Ritter, G. W., & Zamarro, G. (2019). Understanding a vicious cycle: The relationship between student discipline and student academic outcomes. *Educational Researcher*, 48(5), 251–262. <https://doi.org/10.3102/0013189X19848720>
- Arkansas Act 1059. (2017). *An act to amend provisions of Title 6 of the Arkansas Code concerning discipline of students in public schools; And for other purposes*. <https://www.arkleg.state.ar.us/Acts/Document?type=pdf&act=1059&ddBienniumSession=2017%2F2017R>

- Arkansas Act 1329. (2013). *An act to evaluate the impact of school discipline on student achievement; And for other purposes*. <https://www.arkleg.state.ar.us/Acts/FTPDocument?path=%2F ACTS%2F2013%2FPublic%2F&file=1329.pdf&ddBienniumSession=2013%2F2013R>
- Bacher-Hicks, A., Billings, S. B., & Deming, D. J. (2019). *The school to prison pipeline: Long-run impacts of school suspensions on adult crime* (NBER Working Paper 26257). https://www.nber.org/system/files/working_papers/w26257/w26257.pdf
- Baier, V. E., March, J. G., & Saetren, H. (1986). Implementation and ambiguity. *Scandinavian Journal of Management Studies*, 2(3-4), 197-212. [https://doi.org/10.1016/0281-7527\(86\)90016-2](https://doi.org/10.1016/0281-7527(86)90016-2)
- Balfanz, R., Byrnes, V., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, 5(2), Article 13. <https://digitalcommons.library.tmc.edu/cgi/viewcontent.cgi?article=1217&context=childrenatrisk>
- Barrett, N., McEachin, A., Mills, J. N., & Valant, J. (2019). Disparities and discrimination in student discipline by race and family income. *Journal of Human Resources*. Advance online publication. <https://doi.org/10.3368/jhr.56.3.0118-9267R2>
- Bear, G. G., Gaskins, C., Blank, J., & Chen, F. F. (2011). Delaware School Climate Survey—Student: Its factor structure, concurrent validity, and reliability. *Journal of School Psychology*, 49(2), 157-174. <https://doi.org/10.1016/j.jsp.2011.01.001>
- Belansky, E. S., Cutforth, N., DeLong, E., Ross, C., Scarbro, S., Gilbert, L., Beatty, B., & Marshall, J. A. (2009). Early impact of the federally mandated local wellness policy on physical activity in rural, low-income elementary schools in Colorado. *Journal of Public Health Policy*, 30(1), S141-S160. <https://doi.org/10.1057/jphp.2008.50>
- Berry, F. S., & Berry, W. D. (1990). State lottery adoptions as policy innovations: An event history analysis. *American Political Science Review*, 84(2), 395-416. <https://doi.org/10.2307/1963526>
- Bobbett, G. C., & French, R. L. (1991, November 12-15). *Assessing school climate* [Paper presentation]. Annual Meeting of the Mid-South Educational Research Association, Lexington, KY, United States. <https://files.eric.ed.gov/fulltext/ED340738.pdf>
- Brand, S., Felner, R., Shim, M., Seitsinger, A., & Dumas, T. (2003). Middle school improvement and reform: Development and validation of a school-level assessment of climate, cultural pluralism, and school safety. *Journal of Educational Psychology*, 95(3), 570-588. <https://doi.org/10.1037/0022-0663.95.3.570>
- Bray, J. H., Maxwell, S. E., & Howard, G. S. (1984). Methods of analysis with response-shift bias. *Educational and Psychological measurement*, 44(4), 781-804. <https://doi.org/10.1177/00131644844444002>
- Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental process. In R. Lerner (Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (5th ed., pp. 993-1028). Wiley.
- Butler, E. D., & Rakow, J. (1995). *Sample Tennessee school climate profile*. Center for Research in Educational Policy. <https://files.eric.ed.gov/fulltext/ED385942.pdf>
- Christie, C. A., Nelson, C. M., & Jolivet, K. (2004). School characteristics related to the use of suspension. *Education and Treatment of Children*, 27(4), 509-526.
- Chu, E. M., & Ready, D. D. (2018). Exclusion and urban public high schools: Short-and long-term consequences of school Suspensions. *American Journal of Education*, 124(4), 479-509. <https://doi.org/10.1086/698454>
- Cobb-Clark, D. A., Kassenboehmer, S. C., Le, T., McVicar, D., & Zhang, R. (2015). Is there an educational penalty for being suspended from school? *Education Economics*, 23(4), 376-395. <https://doi.org/10.1080/09645292.2014.980398>
- Cohen, D. K., Moffitt, S. L., & Goldin, S. (2007). Policy and practice: The dilemma. *American Journal of Education*, 113(4), 515-548. <https://doi.org/10.1086/518487>
- Cohen-Vogel, L., & McLendon, M. K. (2012). New approaches to understanding federal involvement in education. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of education policy research* (pp. 735-748). Routledge.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart & Winston.
- CT Gen. Stat § 10-233c. (2012). *2012 Connecticut General Statutes Title 10—Education and Culture Chapter 170—Boards of Education Section 10-233c—Suspension of pupils*. <https://law.justia.com/codes/connecticut/2012/title-10/chapter-170/section-10-233c/>
- Davison, M., Penner, A. M., Penner, E., Pharris-Ciurej, N., Porter, S. R., Rose, E., Shem-Tov, Y., & Yoo, P. (2021). *School discipline and racial disparities in early adulthood* (EdWorkingPaper: 21-434). <https://www.census.gov/library/working-papers/2021/adrm/CES-WP-21-14.html>
- Dixon, W. J. (1960). Simplified estimation from censored normal samples. *Annals of Mathematical Statistics*, 31(2), 385-391. <https://doi.org/10.1214/aoms/1177705900>
- Elmore, R. F. (1979-1980). Backward mapping: Implementation research and policy decisions. *Political Science Quarterly*, 94(4), 601-616. <https://doi.org/10.2307/2149628>
- Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. Council of State Governments Justice Center & Public Policy Research Institute.
- Firestone, W. A. (1989). Using reform: Conceptualizing district initiative. *Educational Evaluation and Policy Analysis*, 11(2), 151-164. <https://doi.org/10.3102/01623737011002151>
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Gopalan, M., & Nelson, A. A. (2019). Understanding the racial discipline gap in schools. *AERA Open*. Advance online publication. <https://doi.org/10.1177/2332858419844613>
- Gray, A. N., Sirinides, P. N., Fink, R., Flack, A., & DuBois, T. (2017). *District in context: Suspension, climate, and PBIS in the School District of Philadelphia*. Consortium for Policy Research in Education. https://repository.upenn.edu/cgi/viewcontent.cgi?article=1107&context=cpre_researchreports
- Gray, V. (1994). Competition, emulation, and policy innovation. In L. Dodd, & C. Jillson (Eds.), *New perspectives on American politics* (pp. 230-248). CQ Press.
- Hashim, A. K., Strunk, K. O., & Dhaliwal, T. K. (2018). Justice for all? Suspension bans and restorative justice programs in the Los Angeles unified school district. *Peabody Journal of*

- Education*, 93(2), 174–189. <https://doi.org/10.1080/0161956X.2018.1435040>
- Hill, M., & Hupe, P. (2003). The multi-layer problem in implementation research. *Public Management Review*, 5(4), 471–490. <https://doi.org/10.1080/1471903032000178545>
- Hinze-Pifer, R., & Sartain, L. (2018). Rethinking universal suspension for severe student behavior. *Peabody Journal of Education*, 93(2), 228–243. <https://doi.org/10.1080/0161956X.2018.1435051>
- Jacobsen, W. C., Pace, G. T., & Ramirez, N. G. (2019). Punishment and inequality at an early age: Exclusionary discipline in elementary school. *Social Forces*, 97(3), 973–998. <https://doi.org/10.1093/sf/soy072>
- Johnson, W. L., Dixon, P. N., & Robinson, J. S. (1987). The Charles F. Kettering Ltd. School Climate Instrument: A psychometric analysis. *Journal of Experimental Education*, 56(1), 36–41. <https://doi.org/10.1080/00220973.1987.10806464>
- Johnson, W. L., Johnson, A. M., Gott, R., & Zimmerman, K. (1997). Assessing the validity of scores on the Charles F. Kei'Ering Scale for the junior high school. *Educational and Psychological Measurement*, 57(5), 858–869. <https://doi.org/10.1177/0013164497057005010>
- Kaiser, H. F. (1991). Coefficient alpha for a principal component and the Kaiser-Guttman rule. *Psychological Reports*, 68(3), 855–858. <https://doi.org/10.2466/pr0.1991.68.3.855>
- Kingdon, J. W. (1984). *Agendas, alternatives, and public policies*. Little, Brown.
- Lacoe, J., & Steinberg, M. P. (2018). Rolling back zero tolerance: The effect of discipline policy reform on suspension usage and student outcomes. *Peabody Journal of Education*, 93(2), 207–227. <https://doi.org/10.1080/0161956X.2018.1435047>
- Little, T. D., Chang, R., Gorrall, B. K., Waggenspack, L., Fukuda, E., Allen, P. J., & Noam, G. G. (2020). The retrospective pretest–posttest design redux: On its validity as an alternative to traditional pretest–posttest measurement. *International Journal of Behavioral Development*, 44(2), 175–183. <https://doi.org/10.1177/0165025419877973>
- Liu, J., Hayes, M. S., & Gershenson, S. (2021). *From referrals to suspensions: New evidence on racial disparities in exclusionary discipline* (EdWorkingPaper: 21-442). <https://www.edworkingpapers.com/ai21-442>
- Locker, L. A. (2001, February). *Some “modern” statistics: A primer and demonstration* [Paper presentation]. Annual Meeting of the Southwest Educational Research Association, New Orleans, LA, United States. <http://files.eric.ed.gov/full-text/ED449221.pdf>
- Losen, D., Hodson, C., Keith, M., Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?* The Civil Rights Project, UCLA. <https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/are-we-closing-the-school-discipline-gap>
- Maryland State Department of Education. (2020). *Policy Impact Analysis: HB425/SB651 (2017)*. Maryland State Department of Education Office of Research and Strategic Data Use. <https://www.marylandpublicschools.org/stateboard/Documents/2020/0623/Prek2SuspensionPolicyReport062320.pdf>
- Matland, R. E. (1995). Synthesizing the implementation literature: The ambiguity-conflict model of policy implementation. *Journal of Public Administration Research and Theory*, 5(2), 145–174. <https://www.jstor.org/stable/1181674>
- Mazmanian, D. A., & Sabatier, P. A. (1983). *Implementation and public policy*. Scott Foresman.
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9(2), 171–178. <https://doi.org/10.3102/01623737009002171>
- MD Educ Code § 7-305.1. (2017). *2017 Maryland Code Education Division II—Elementary and Secondary Education Title 7—Public Schools Subtitle 3—Attendance and Discipline of Students § 7-305.1. Student enrolled in public prekindergarten program, kindergarten, first grade, or second grade—Suspension or expulsion prohibited; exceptions*. <https://law.justia.com/codes/maryland/2017/education/division-ii/title-7/subtitle-3/section-7-305.1/>
- Mintrom, M. (1997). Policy entrepreneurs and the diffusion of innovation. *American Journal of Political Science*, 41(3), 738–770. <https://doi.org/10.2307/2111674>
- Mowen, T., & Brent, J. (2016). School discipline as a turning point: The cumulative effect of suspension on arrest. *Journal of Research in Crime and Delinquency*, 53(5), 628–653. <https://doi.org/10.1177/0022427816643135>
- Mukuria, G. (2002). Disciplinary challenges: How do principals address this dilemma? *Urban Education*, 37(3), 432–452. <https://doi.org/10.1177/00485902037003007>
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15(3), 263–280. <https://doi.org/10.1002/ejsp.2420150303>
- Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly*, 90(4), 1003–1018. <https://doi.org/10.1111/j.1540-6237.2009.00674.x>
- NJ Rev Stat § 18A:37-2A. (2016). *2016 New Jersey Revised Statutes Title 18a—Education Section 18A:37-2A—Conditions for suspension, expulsion of certain students*. <https://law.justia.com/codes/new-jersey/2016/title-18a/section-18a-37-2a/>
- Noltmeyer, A. L., Ward, R. M., & Mcloughlin, C. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review*, 44(2), 224–240. <https://doi.org/10.17105/spr-14-0008.1>
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443–460. <https://doi.org/10.1007/BF02296207>
- Payne, A., & Welch, K. (2010). Modeling the effects of racial threat on punitive and restorative school discipline practices. *Criminology*, 48(4), 1019–1062. <https://doi.org/10.1111/j.1745-9125.2010.00211.x>
- Pearson, R. W., Ross, M., & Dawes, R. M. (1992). Personal recall and the limits of retrospective questions in surveys. In J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 65–94). Russell Sage Foundation.
- Rafa, A. (2018). *50-State comparison: State policies on school discipline*. Education Commission of the States. <https://www.ecs.org/50-state-comparison-state-policies-on-school-discipline/>
- Raffaella Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development*, 2003(99), 17–33. <https://doi.org/10.1002/yd.52>

- Ritter, G. W., & Anderson, K. P. (2018). Examining disparities in student discipline: Mapping inequities from infractions to consequences. *Peabody Journal of Education, 93*(2), 161–173. <https://doi.org/10.1080/0161956X.2018.1435038>
- Sabatier, P., & Mazmanian, D. (1979). The conditions of effective implementation: A guide to accomplishing policy objectives. *Policy Analysis, 5*(4), 481–504. <https://www.jstor.org/stable/42783358>
- Sartain, L., Allensworth, E., Porter, S., Levenstein, R., Johnson, D. W., Huynh, M. H., Anderson, E., Mader, N., & Steinberg, M. P. (2015). *Suspending Chicago's students: Differences in discipline practices across schools*. University of Chicago Consortium on Chicago School Research.
- Sheldon, S. B., & Epstein, J. L. (2002). Improving student behavior and school discipline with family and community involvement. *Education and Urban Society, 35*(1), 4–26. <https://doi.org/10.1177/001312402237212>
- Shi, Y., & Zhu, M. (2021). *Equal time for equal crime? Racial bias in school discipline* (EdWorkingPaper: 21-383). <https://doi.org/10.2139/ssrn.3837758>
- Skiba, R. J., Chung, C., Trachok, M., Baker, T., Sheya, A., & Hughes, R. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal, 51*(4), 640–670. <https://doi.org/10.3102/0002831214541670>
- Skiba, R. J., Horner, R., Chung, C. G., Rausch, M. K., May, S., & Tobin, T. (2011). Race is not neutral: A national investigation of Black and Latino disproportionality in school discipline. *School Psychology Review, 40*(1), 85–107. <https://doi.org/10.1080/02796015.2011.12087730>
- Skiba, R. J., Michael, R., Nardo, A., & Peterson, R. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review, 34*(4), 317–342. <https://doi.org/10.1023/A:1021320817372>
- Skowronek, S. (1993). *The politics presidents make: Leadership from John Adams to George Bush*. Harvard University Press.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387–431. <https://doi.org/10.3102/00346543072003387>
- Steinberg, M. P., & Lacoë, J. (2018). Reforming school discipline: School-level policy implementation and the consequences for suspended students and their peers. *American Journal of Education, 125*(1), 29–77. <https://doi.org/10.1086/699811>
- Suh, S., & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling, 10*(3), 297–306. <https://doi.org/10.5330/prsc.10.3.w26024vww6541gv7>
- Sullivan, A. L., Klingbeil, D. A., & Van Norman, E. R. (2013). Beyond behavior: Multilevel analysis of the influence of sociodemographics and school characteristics on students' risk of suspension. *School Psychology Review, 42*(1). <https://doi.org/10.1080/02796015.2013.12087493>
- Swanson, E., Erickson, H., & Ritter, G. W. (2017). *Examining the impacts of middle school disciplinary policies on ninth-grade retention* (EDRE Working Paper No. 2017-11). <https://doi.org/10.2139/ssrn.2952972>
- Texas Education Code § 37.005. Suspension. (2020). <https://codes.findlaw.com/tx/education-code/educ-sect-37-005.html>
- Van Meter, D. S., & Van Horn, C. E. (1975). The policy implementation process: A conceptual framework. *Administration & Society, 6*(4), 445–488. <https://doi.org/10.1177/009539977500600404>
- Vincent, C. G., Sprague, J. R., & Tobin, T. J. (2012). Exclusionary discipline practices across students' racial/ethnic backgrounds and disability status: Findings from the Pacific Northwest. *Education and treatment of children, 35*(4), 585–601. <https://doi.org/10.1353/etc.2012.0025>
- Walker, J. L. (1969). The diffusion of innovations among the American states. *American Political Science Review, 63*(3), 1174–1185. https://fbaum.unc.edu/teaching/PLSC541_Fall08/walker_1969.pdf
- Weatherly, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review, 47*(2), 171–197. <https://doi.org/10.17763/haer.47.2.v870r1v16786270x>
- Welch, K., & Payne, A. A. (2010). Racial threat and punitive school discipline. *Social Problems, 57*(1), 25–48. <https://doi.org/10.1525/sp.2010.57.1.25>
- Wolf, K. C., & Kupchik, A. (2017). School suspensions and adverse experiences in adulthood. *Justice Quarterly, 34*(3), 407–430. <https://doi.org/10.1080/07418825.2016.1168475>
- Zullig, K. J., Koopman, T. M., Patton, J. M., & Ubbes, V. A. (2010). School climate: Historical review, instrument development, and school assessment. *Journal of Psychoeducational Assessment, 28*(2), 139–152. <https://doi.org/10.1177/0734282909344205>

Authors

KAITLIN P. ANDERSON is an assistant professor of educational leadership in the Department of Education and Human Services at Lehigh University. Her research interests focus on issues of equity in the implementation and impact of education policy.

SARAH MCKENZIE is an assistant professor in the Department of Education Reform at the University of Arkansas and Executive Director of the Office for Education Policy. Her research interests include assessment literacy, academic growth models, K–20 educational outcomes, and using data to drive changes in school systems in order to better serve all students.