

NEPC REVIEW: THE ACADEMIC AND BEHAVIORAL CONSEQUENCES OF DISCIPLINE POLICY REFORM: EVIDENCE FROM PHILADELPHIA (FORDHAM INSTITUTE, DECEMBER 2017)



Reviewers:

Yolanda Anyon, University of Denver
Kathryn Wiley, University of California at San Diego

February 2018

National Education Policy Center

School of Education, University of Colorado Boulder
Boulder, CO 80309-0249
(802) 383-0058
nepc.colorado.edu



Kevin Welner
Project Director

William Mathis
Managing Director

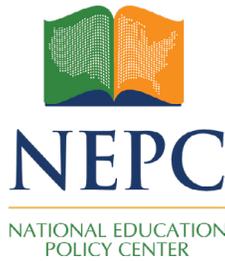
Alex Molnar
Publishing Director

Publication Type: NEPC Reviews are expert third-party reviews of selected non-peer-reviewed publications. Using academic peer review standards, reviewers consider the quality and defensibility of a report's assumptions, methods, findings, and recommendations. Written in non-academic language, reviews are intended to help policymakers, reporters, and others assess the social science merit of a reviewed report and to judge its value in guiding policy.

Funding: This review is one of a series made possible in part by funding from the [Great Lakes Center for Education Research and Practice](#).

Suggested Citation: Anyon, Y., & Wiley, K. (2018). *NEPC Review: "The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia"* (Fordham Institute, December 2017). Boulder, CO: National Education Policy Center. Retrieved [date] from <http://nepc.colorado.edu/thinktank/review-discipline-reform>.

This material is provided free of cost to NEPC's readers, who may make non-commercial use of the material as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at nepc@colorado.edu.



**NEPC REVIEW: THE ACADEMIC AND BEHAVIORAL
CONSEQUENCES OF DISCIPLINE POLICY REFORM:
EVIDENCE FROM PHILADELPHIA
(FORDHAM INSTITUTE, DECEMBER 2017)**

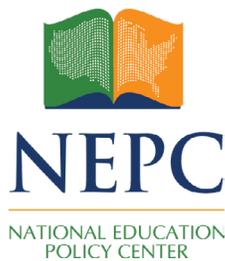
Reviewers:

Yolanda Anyon, University of Denver
Kathryn Wiley, University of California at San Diego

February 2018

Summary of Review

A recent report from the Fordham Institute investigated the impact of a reform in the School District of Philadelphia that eliminated suspensions for certain low-level misbehaviors. The report considered whether the policy change was associated with any of the following: (a) district-wide out-of-school suspension rates, (b) academic and behavioral outcomes for students (looking separately at students who had a record of prior suspensions and those with no prior suspensions), and (c) racial disparities in suspensions. While the report concluded that the reform was a failure, the actual results were mixed, with the positive trends for students who were earlier suspended being much stronger in magnitude than evidence of negative outcomes for students who were not. A strength of the report is the use of advanced statistical methods and a longitudinal dataset to answer the questions of interest. However, the report is plagued by logical fallacies, overly simplified interpretations of findings, and inflammatory language. Moreover, the report uses misleading causal (“consequences”) language in the title and to describe study results, even though the study design is limited by unmeasured confounding factors and inappropriate comparison groups. Thus, while the analyses upon which the report is based have some technical merits, the narrative seems more of an attempt to advance a political agenda opposed to the reform studied than to improve understanding of complex policy issues.



NEPC REVIEW: THE ACADEMIC AND BEHAVIORAL CONSEQUENCES OF DISCIPLINE POLICY REFORM: EVIDENCE FROM PHILADELPHIA (FORDHAM INSTITUTE, DECEMBER 2017)

Reviewers:

Yolanda Anyon, University of Denver
Kathryn Wiley, University of California at San Diego

February 2018

I. Introduction

Research spanning more than three decades has consistently raised concerns about the use of out-of-school suspensions. Youth of color, particularly Black students, are significantly more likely than their peers to be suspended from school, even for the same behaviors.¹ Students who are suspended are at increased risk for a wide range of negative developmental outcomes, from school pushout to contact with the criminal justice system.² Additionally, schools that use punitive and exclusionary practices more often than others tend to have larger achievement gaps and are also perceived by students, teachers, and parents to be less safe.³

Of particular concern is the use of out-of-school suspension for minor misbehaviors, which include categories such as “disruption” or “defiance.” These categories are typically on the low end of schools’ behavior ladders, as compared to more extreme categories related to weapons or physical aggression. Minor misbehaviors account for the bulk of suspensions assigned in schools, and are subject to a greater degree of racially disproportionality than other incident categories, raising questions about the appropriateness of suspensions as a response to these nonviolent incidents.⁴ In light of this research and in reaction to sustained parent and youth organizing campaigns, a number of schools districts have recently prohibited or limited the use of out-of-school suspensions for low-level misconduct. For example, in 2014, California eliminated suspensions for “minor misbehaviors” and similarly policy changes have occurred in Miami-Dade, Denver, and Los Angeles.⁵ These reforms have naturally led to questions about the impact of limiting suspensions. Yet few districts have used advanced statistical methods to assess the relationship between new discipline policies and student, school, or district outcomes.

The Fordham Institute report, *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*, published December 2017, aimed to fill this gap. It examined how patterns in suspension, standardized test scores, and attendance changed after the School District of Philadelphia (SDP) passed a policy limiting the use of out-of-school suspensions, particularly for minor misbehaviors.⁶ The Fordham Institute report presents a synthesis of two research papers authored by Matthew Steinberg, University of Pennsylvania, and Johanna Lacoë, Mathematica Policy Research, which assessed the relationship between the policy change and subsequent trends in the district.⁷ In this review, we evaluate the methods, findings, and conclusions presented in the Fordham Institute report and their potential contribution to policy decision-making.

II. Findings and Conclusions of the Report

The SDP policy mandated an end to using out-of-school suspension for “conduct offenses,” which include failing to follow classroom rules and cursing or using obscene gestures. The new policy also encouraged schools to use suspension as a last resort for students’ public displays of affection and unsanctioned use of electronic devices. SDP’s discipline reform changed the harshest punishment for some more serious offenses (e.g. robbery, simple assault, and bullying, called “non-conduct offenses”) from expulsion to suspension. Finally, the reform encouraged schools to use in-school interventions rather than exclusionary approaches to resolving all types of student misbehavior and conflict. The policy did not introduce any new supports or resources for schools.

The Fordham Institute report addressed four distinct, but interrelated research questions about the impact of the SDP policy reform on district, school, and student outcomes. Each question was answered using a difference-in-differences statistical approach that compares change over time for distinct groups. All of the statistical models also used fixed effects, which means that school- and district-level contexts were held constant. This approach accounts for unmeasured variables that do not change over time.

First, the report considered whether policy reform in SDP led to a reduction in district-wide suspension rates, comparing the district to most others in Pennsylvania over a ten-year period. Results indicated that the policy was not related to lower overall suspensions in the three years following discipline reform, relative to other districts. Accounting for pre-policy trends (suspensions were already declining prior to reforms), SDP’s overall suspension rate did not change substantially during the first two years following the reform, then increased somewhat by the third year of policy implementation, relative to other districts. This pattern varied depending on the type of suspension being considered, with non-conduct (more serious) offenses consistently rising slightly over the three-year period and no clear pattern evident for conduct offenses, which were the primary target of the policy reform. The magnitudes of the associations between time and suspension rates were very small and are arguably negligible, though statistically significant. The report also documents modest and statistically significant increases in truancy rates, relative to other districts, in the years following discipline reform. The report’s takeaway is that the policy led to initial reductions

in conduct suspensions that did not persist after the first year of policy implementation, but also resulted in an increase in suspensions for more serious, non-conduct, infractions.

Second, the report examined whether the policy reform was associated with changes in several student outcomes (suspension, standardized test scores, and attendance) for youth who had received suspensions prior to the reform. In this case, students who were suspended before the reform were compared to those who had not been suspended previously. Results suggested that students with a history of suspensions before the policy change received fewer numbers and days of suspension for both conduct and non-conduct offenses after the reform's passage. It is also reported that the policy change was not associated with changes in standardized test scores for these students (though results suggest otherwise in the supplementary papers), but was significantly and strongly related to improved attendance the year following the reform. These findings are not revisited in the report's conclusions.

The third research question considered whether the policy reform was associated with changes in the same student outcomes, but this time for students who were *not* suspended before discipline reform was passed in the district. To answer this question, a different methodology was employed. Several groups of students were constructed based on the school that they attended and the degree to which the school changed its conduct suspension rates after the reform passed. Schools were categorized as full compliers (18%), partial compliers (60%), non-compliers (17%), or a comparison group that had no conduct suspensions before or after reform passage (5%). Full compliers eliminated their use of conduct suspensions after the reform, whereas partial compliers reduced their conduct suspension rates but still had some, and non-compliers increased conduct suspensions. Results indicated that students attending full complier schools did not experience any changes in achievement or attendance after the policy reform. However, test score and attendance outcomes worsened slightly for students in schools that partially complied, or did not comply at all, with the new mandates. The results are interpreted to mean that the policy had negative and unintended consequences for non-suspended students. The report also suggested that the existence of partial and non-complier schools indicated that implementation of the policy was uneven and depended on school's prior suspension rates, student composition, and achievement levels. It concluded with a recommendation that district reforms need to be coupled with additional resources, such as professional development programs to train teachers in Positive Behavioral Interventions and Supports (PBIS), for schools that are highly segregated and low-achieving.

Finally, the report considered whether racial disproportionalities in out-of-school suspension declined after SDP's discipline policy reform. Results indicated that racial gaps in conduct suspensions shrunk very slightly post-reform, but that racial disparities in non-conduct offenses grew more rapidly, leading to larger overall suspension gaps for Black and Latino students. The report attributed this trend to several possible factors: schools with more students of color were less likely to comply with the policy reform, low-level misbehaviors may have been reclassified as higher-level incidents in order to avoid the mandates of the new policy, or students of color may have escalated their behavior during this time period.

III. The Report's Rationale for Its Findings and Conclusions

The findings and conclusions of the report were based exclusively on several advanced statistical analyses of discipline, attendance, and achievement data. The authors relied on this original research to conclude that the policy did not have a sustained impact on district-wide suspension rates, but did have unintended and negative “spillover” consequences for non-suspended students. Findings regarding the modest, but positive, changes in attendance and discipline for students who had been suspended prior to the reform were essentially ignored and were not discussed in the report's conclusions. Based on the results indicating that implementation of the policy reform was uneven across the district, and that the schools who were least likely to comply served more disadvantaged students, the report recommended that policy mandates need to be coupled with additional resources to help schools implement effective alternatives to suspension without negatively impacting students who do not misbehave. These conclusions are arrived at without consideration of the limitations of the study design.

IV. The Report's Use of Research Literature

The report's use of the research literature is quite weak (just 11 peer-reviewed articles are referenced) and very selective, particularly in the foreword, which was not written by the study authors. Although a few prominent and accepted works were referenced in the background section, several key claims in the foreword were not supported with any evidence. In some cases, these assertions actually contradicted the preponderance of available research.

For example, it is stated in the report that “it should not shock us to discover that in some circumstances and communities, minority students misbehave at ‘disproportionate’ rates.”⁸ The authors provided no references in support this claim, which stands in stark contrast to the consensus achieved by a collaborative of 26 scholars convened to review the extant literature on school discipline disparities. These researchers concluded that there is “no evidence” that differential rates of behavior by race explain disparities.⁹ Instead, they cited research suggesting that in some cases students of color report lower rates of risk behaviors that are punished at school (e.g. drug and alcohol use) than White youth.¹⁰ This conclusion is also supported by several peer-reviewed studies that have accounted for student behavior and still find pronounced racial disparities in office discipline referrals and suspensions.¹¹

The report also stated that “everyone knows that changing a district's policy on suspensions is unlikely to alter the underlying issues in tough schools.”¹² While it is true that a new policy mandate may not shape the material conditions in a school, there is evidence from other types of educational reforms - even those uncoupled with additional resources - that policy directives can change the behavior and attitudes of adult staff members.¹³ Some scholars have argued, with empirical support, that these actions and beliefs of school staff are more meaningful drivers of school discipline outcomes than student behavior.¹⁴ Another example of an unsupported claim is found in the report's recommendation that discipline reform is “best initiated at the school level rather than the district level, where the law of unintended

consequences is more apt to prevail.”¹⁵ Again, no evidence is provided - from the study report or the broader literature - to support such an argument.

Similarly, the report made several assertions that were only supported by a very limited body of research. It cautioned that disruptive students can negatively impact the learning of their peers, but only one peer-reviewed study was cited in support of this claim. Moreover, the study referenced involved children who had experienced domestic violence, which is not representative of all students with discipline issues. This argument also ignored research by one of the report authors, conducted in Chicago, which indicated that reforms to suspension length did not have negative peer effects.¹⁶

Finally, several misleading and inflammatory comments were made in the introduction regarding recent federal policy guidance on school discipline.¹⁷ First is that this guidance implied that most school staff “have racist tendencies and may be deliberately violating students’ civil rights.”¹⁸ Second, that the “unspoken assumption” behind the concerns about racial disparities in suspension rates is that gaps between student groups “must be attributable to such bias.”¹⁹ This is an inaccurate characterization of the federal policy guidance, which drew attention to racial disparities in suspension, encouraged practitioners to address them, and provided examples of when the federal government might find cause to intervene.²⁰ Such concerns about discrimination appear to be valid in light of growing experimental evidence that finds evidence of racial bias in discipline decisions.²¹

However, the guidance also discussed many other factors beyond bias that contribute to disproportionality and is very careful to point out that disparate impact does not always mean intentional discrimination. In fact, some scholars have critiqued these documents for *not* addressing the issue of racism enough, particularly in terms of recommended interventions, which largely focus on race-neutral strategies to promote a positive school climate.²²

V. Review of the Report’s Methods

The studies at focus in this report used two econometric statistical approaches – fixed effects and difference-in-difference - to answer the research questions of interest. Although these approaches are not experimental, the findings were interpreted to mean that the policy change caused the trends described in the report. This is especially problematic because the studies are limited by two key methodological issues – unmeasured confounding variables and inappropriate comparison groups.

The strength of a fixed effects approach is that it accounts for unmeasured variables that are constant or unchanging across time. On the other hand, this method does not control for unobserved student-, school-, or district-level characteristics that do vary over time and could influence discipline outcomes. To a certain extent, this is acknowledged in the report: “school discipline is an extraordinarily difficult subject to study, in part because we do not observe student behavior directly, but only documented responses to it.”²³ Two examples of

student characteristics that were not controlled for in the study that may influence outcomes are described in the report, including a death in the family or exposure to neighborhood violence that might lead to a student acting out. Several other unmeasured student-level variables are outlined in the papers that supplement the report, such as lack of controls for receipt of suspensions after the passage of the policy reform. (It is unclear why the latter were not included in the analyses, as the study authors had access to this data.)

The validity of study findings regarding the impact of the reform on suspension rates was weakened by more than just a lack of observable data about student behavior. Research indicates that variable school and district characteristics – not just student factors – are also associated with the use of exclusionary outcomes.²⁴ In particular, the report ignores organizational contexts that may have changed over this time, such as a new principal or teachers, or the introduction of new academic or behavioral initiatives. Indeed, turnover among principals and teachers is high in urban districts like SDP that serve predominantly socially and economically disadvantaged students.²⁵ Moreover, a principal’s discipline philosophy may be one of the strongest determining factors in a school’s discipline outcomes, but was not measured in this study.²⁶ Equally problematic is the lack of data about other behavioral interventions implemented in SDP that likely varied with time, such as in-school prevention programs or alternatives to suspension like restorative practices. Qualitative evidence indicates that schools modify and experiment with different approaches to addressing student misbehavior, and that these efforts are highly contingent on the vision of school leadership.²⁷ In short, study findings must be interpreted with caution because of lack of data capturing dynamics in schools that could have influence the outcomes of interest.

The validity of study findings regarding the impact of the reform on suspension rates was weakened by more than just a lack of observable data about student behavior.

Similarly, changes in the national policy context might also explain some of the study results. For example, findings regarding changes in district-level suspension rates after the passage of policy reform are mixed until the 2014-2015 school year, when suspension rates in comparison districts decline. Rather than indicating a failure of SDP’s policy reform,

this trend may be partially attributable to national policy guidance issued in January 2014 that encouraged schools to minimize their use of exclusionary practices, focus on providing students with support services, and improve school climate. It seems possible that suspension rates in SDP appeared to rise more significantly during the 2014-2015 school year, relative to other districts, because these localities were implementing reforms of their own.

With respect to the difference-in-difference methodology used, the validity of this approach hinges on the matching of participants to non-participants with the same observed characteristics. This assumption was not met for the most controversial analysis in the report, which suggested negative spillover effects of the reform on non-suspended students. In this case, attendance and academic outcomes for students in schools that had no conduct suspensions before or after the reform were compared to students in schools that had different degrees of compliance with the policy mandates. The paper that supplements the report noted that “the academic achievement of students in partial compliance schools was significantly lower than in the full compliance schools in the pre-reform period.”²⁸ These schools also tended to

have higher suspension rates and served students from less advantaged backgrounds. Such schools are likely to be meaningfully different on several dimensions that are related to test scores and attendance, but were not accounted for in the study. For example, high-poverty and racially segregated schools tend to have lower teacher quality and offer weaker opportunities for academic engagement than schools serving more advantaged students.²⁹ Thus, the comparison group used to answer research questions about non-suspended students used was meaningfully different from the other groups in the study, in ways that likely influenced the outcomes of interest.

Taken together, concerns about biased comparison groups and unaccounted-for factors that change over time suggest that the study findings synthesized in the report are not as strong as the reader is led to believe.

VI. Review of the Validity of the Findings and Conclusions

The most problematic aspects of the Fordham Institute report are the ways in which study findings are interpreted and translated into conclusions and recommendations. Although the analyses used advanced statistical methods, the study design was not experimental, so results can only be interpreted in terms of relationships between different variables. Indeed, most of the evidence presented in the report is classified by the authors as “associational.”³⁰ Yet the report implies otherwise by using causal language throughout, including in the title, in which changes in outcomes are described as “consequences” of the reforms. Such claims are misleading, as it is not possible to use these methods to conclusively determine whether the policy reform did in fact cause the differences observed. This is especially true given the nature of the administrative datasets that were the basis for the analyses, as these lacked observable data about both student and staff behavior.

Equally concerning, there are several instances in which the findings and conclusions presented in the report either contradicted or overlooked results from the original research studies. For example, the conclusion that “most schools did not comply” with the reform belies the actual data, which indicated that the vast majority of schools at least partially adhered to the policy mandates (only 17% were not compliant at all). In assessing whether the reform was associated with changes in suspension, achievement, and attendance for students who had been suspended prior to the reform, the Fordham report claims that these students experienced no changes in achievement. This contradicts the original study, which concluded that previously suspended students made marginal, but statistically significant, gains in math proficiency.³¹ Thus, study findings consistently indicated that outcomes improved in the post-reform period for students who had been suspended prior to the policy change. These results were not considered in the report’s conclusions about the value of the policy change or their recommendations for policy makers. This is surprising, as the magnitude of the positive trends for previously suspended students is much larger than any evidence of negative trends among their unsuspended peers. In general, the pattern of inconsistent reporting of the original study results indicates some “cherry picking” in order to

make a partisan argument against current federal policy.

Additionally, variation in implementation of discipline policies is not surprising and has been documented elsewhere,³² suggesting the most relevant policy question is how to support schools in implementing reforms uniformly. The Fordham report's attribution of compliance variation to student demographics alone oversimplifies the complex dimensions of policy reform that likely have more to do with adult skills and competencies than student characteristics. That implementation was unsustainable does not suggest the policy should be abandoned, but rather that greater efforts should focus on increasing schools' capacity, as is recommended at the end of the report.

Addressing questions of the reform's impact on achievement and attendance among the "rule-abiding majority," the report finds that achievement declined among students who were not suspended prior to the reform but attended schools that continued to use conduct suspensions.³³ In a strange manipulation of reasoning, they argue the decline in achievement among these "rule-abiders" was due to the inability of these schools to suspend misbehaving peers. However, this interpretation contradicts the fact that these schools continued to use suspensions. To clarify, results indicated that achievement declines among non-suspended peers were only observed in schools that did not fully implement the policy, meaning these schools were still suspending students for low-level misbehaviors. The notion that these schools could not suspend students, is, then, inaccurate. The report overlooks a logical explanation for the achievement decline: the continued use of suspensions. For even if the continued use of suspensions did not directly impact rule-abiders (this is unknown, as the study did not account for suspensions after the reform), it may be related to negative school experiences that are associated with achievement, such as student connectedness.³⁴ It is also possible that achievement declines were due to other confounding contextual variables, such that non-complying schools may have also had high administrative turnover or

The report's claim that the achievement of rule-abiding peers was negatively impacted appears to be a stark misinterpretation of the results.

were subject to school closure efforts – two issues SDP faced during the analysis period.³⁵

Additionally, the report's emphasis on achievement declines for "rule-abiders" in non-complying schools overlooks a very important, and understated, finding: in schools that fully implemented the district policy, students who were not suspended prior to the reform

saw no changes in achievement. In contrast to the conclusion drawn by the report, which illogically argues the policy negatively impacted rule abiders, the findings actually indicate that prohibiting suspension was not adversely related to achievement for "rule-abiding" students. Thus, the report's claim that the achievement of rule-abiding peers was negatively impacted appears to be a stark misinterpretation of the results.

In contrast to the overwhelmingly negative tone of the report, our examination of the findings suggests that the discipline policy reform in SDP was associated with mixed outcomes. Overlooked or underemphasized results included small gains in achievement and attendance among students who were suspended before the policy change, and the lack of adverse changes in achievement or attendance among "the rule abiding majority" who attended schools that fully implemented the policy. Conversely, attending a school that did not

implement the policy, and continued to issue non-conduct suspensions, was associated with declines in achievement and attendance concerns.

VII. Usefulness of the Report for Guidance of Policy and Practice

Given the Fordham report's selective presentation of the original study findings, we do not recommend that policy makers use its conclusions as the basis for policy decisions. Instead, we encourage stakeholders to review the supplementary papers that were the basis for the report, as these provide a more balanced and cautious interpretation of results. Although the overall findings that discipline reform was associated with uneven implementation and mixed outcomes is not new, the analyses do raise important questions about whether policy mandates that are uncoupled with additional resources are sufficient to reduce the use of punitive and exclusionary practices while raising achievement for all students.

Notes and References

- 1 Anyon, Y., Jenson, J., Altschul, I., Farrar, J., McQueen, J., Greer, E., Downing, B., & Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review, 44*, 379-386;

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

Skiba, R., Michael, R.S., Nardo, A.C., & Peterson, R.L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review, 31*, 317-342;

Wu, S.C., Pink, W., Crain, R., & Moles, O. (1982). Student suspension: A critical reappraisal. *The Urban Review, 14*, 245-303.
- 2 Skiba, R., Arredondo, M.I., & Williams, N.T. (2014). More than a metaphor: The contribution of exclusionary discipline to a school-to-prison pipeline. *Equity & Excellence in Education, 47*, 546-564;

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

Kirk, D.S., & Sampson, R.J. (2013). Juvenile arrest and collateral educational damage in the transition to adulthood. *Sociology of Education, 86*, 36-62.
- 3 Steinberg, M.P., Allensworth, E., & Johnson, D.W. (2014). What conditions support safety in urban schools? The Influence of school organizational practices on student and teacher reports of safety in Chicago. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion*. New York, NY: Teachers College Press;

Morris, E.W., & Perry, B.L. (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems, 63*(1), 68-86.
- 4 Fowler, D., Lightsey, R., Monger, J., & Aseltine, E. (2010). *Texas' School-to-Prison Pipeline: Ticketing, Arrest & Use of Force in Schools*. Texas Appleseed. Retrieved January 9, 2018, from www.texasappleseed.net;

Losen, D., & Martinez, T. (2013). *Out of School & Off Track: the overuse of suspensions in American middle and high schools*. Center for Civil Rights Remedies. Retrieved January 15, 2018, from <https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/federal-reports/out-of-school-and-off-track-the-overuse-of-suspensions-in-american-middle-and-high-schools>;

Skiba, R., Horner, R.H., Chung, C.-G., Rausch, M.K., May, S.L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review, 40*(1), 85-107.
- 5 Frey, S. (2014 September 18). *New law limits student discipline measure*. EdSource. Retrieved January 17, 2018, from <https://edsources.org/2014/new-law-limits-student-discipline-measure/67836>;

The White House (2016). *The Continuing Need to Rethink Discipline*. Retrieved September 1, 2017, from https://www.whitehouse.gov/sites/default/files/docs/school_discipline_report_120916.pdf
- 6 Steinberg, M.P. & Lacoce, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from

<http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>

- 7 Lacoë, J.R., & Steinberg, M.P. (2017). *Rolling back zero tolerance: The effect of discipline policy reform on suspension usage, school climate, and student achievement*. Working paper. Retrieved January 16, 2018, from <http://edex.s3-us-west-2.amazonaws.com/%2817.10.31%29%20Rolling%20Back%20Zero%20Tolerance%20-%20Lacoë%20Steinberg.pdf>;

Steinberg, M.P. & Lacoë, J.R. (2017). *Reforming School Discipline: School-Level Policy Implementation and consequences for suspended students and their peers*. Working paper. Retrieved January 16, 2018, from <http://edex.s3-us-west-2.amazonaws.com/%2817.07.17%29%20Reforming%20School%20Discipline%20-%20Lacoë%20Steinberg.pdf>
- 8 Steinberg, M.P. & Lacoë, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 5.
- 9 Skiba, R. J., Arredondo, M. I., & Rausch, M. K. (2014). *New and developing research on disparities in discipline*. Bloomington, IN: The Equity Project at Indiana University.
- 10 Toldson, I.A., McGee, T., & Lemmons, B.P. (2013). Reducing suspensions by improving academic engagement among school-age Black males. In D. Losen (Ed.), *Closing the school discipline gap: Research for policymakers*. New York, NY: Teachers College Press.
- 11 Anyon, Y., Jenson, J., Altschul, I., Farrar, J., McQueen, J., Greer, E., ... Simmons, J. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review*, 44, 379-386;

Bradshaw, C.P., Mitchell, M.M., O'Brennan, L.M., & Leaf, P.J. (2010). Multilevel exploration of factors contributing to the overrepresentation of Black students in office disciplinary referrals. *Journal of Educational Psychology*, 102, 508-520;

Finn, J.D., & Servoss, T.J. (2015). Security measures and discipline in American high schools. In D. Losen (Ed.), *Closing the school discipline gap: Research for policymakers* (pp. 44-58). New York, NY: Teachers College Press;

Huang, F.L. (2016). Do Black students misbehave more? Investigating the differential involvement hypothesis and out-of-school suspensions. *The Journal of Educational Research*, 1-11;

Wallace, J.M., Jr., Goodkind, S., Wallace, C., & Bachman, J.G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991-2005. *Negro Educational Review*, 59, 47-62;

The argument that differential behavior drives racial disparities is thoroughly reviewed in Skiba, R., & Williams, N.T. (2014). *Are Black kids worse? Myths and facts about racial differences in behavior*. The Equity Project at Indiana University;

Schollenberger, T. (2015). Racial disparities in school suspension and subsequent outcomes: Evidence from the National Longitudinal Survey of Youth. In D. Losen (Ed.), *Closing the school discipline gap: Research for policymakers* (pp. 31-43). New York, NY: Teachers College Press.
- 12 Steinberg, M.P. & Lacoë, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from

<http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 8.

- 13 Coburn, C.E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77, 211-244;

Russell, J.L. (2011). From child's garden to academic press: The role of shifting institutional logics in redefining kindergarten education. *American Educational Research Journal*, 48, 236-267;
- 14 Skiba, R.J., Chung, C.G., Trachok, M., Baker, T.L., Sheya, A., & Hughes, R.L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal*, 51, 640-670;

Wu, S.C., Pink, W.T., Crain, R.L., & Moles, O. (1982). Student suspension: A critical reappraisal. *The Urban Review*, 14, 245-303.
- 15 Steinberg, M.P. & Lacoce, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 7.
- 16 Mader, N., Sartain, L., & Steinberg, M.P. (2016). *When suspensions are shorter: The effects on school climate and student learning*. Working paper.
- 17 Lhamon, C., & Samuels, J. (2014). *Dear Colleague Letter on the nondiscriminatory administration of school discipline*. U.S. Department of Education Office of Civil Rights & U.S. Department of Justice Civil Rights Division.
- 18 Steinberg, M.P. & Lacoce, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 15.
- 19 Steinberg, M.P. & Lacoce, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 5.
- 20 Lhamon, C., & Samuels, J. (2014). *Dear Colleague Letter on the nondiscriminatory administration of school discipline*. U.S. Department of Education Office of Civil Rights & U.S. Department of Justice Civil Rights Division.
- 21 Okonofua, J.A., & Eberhardt, J.L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science*, 26, 617-624;

Chang, D.F., & Sue, S. (2003). The effects of race and problem type on teachers' assessments of student behavior. *Journal of Consulting and Clinical Psychology*, 71, 235-242.
- 22 Anyon, Y., Lechuga, C., Ortega, D., Downing, B., Greer, E., & Simmons, J. (2017). An exploration of the relationships between student racial background and the school sub-contexts of office discipline referrals: a critical race theory analysis. *Race Ethnicity and Education*, 1-17;

- Carter, P.L., Skiba, R., Arredondo, M.I., & Pollock, M. (2017). You can't fix what you don't look at: Acknowledging race in addressing racial discipline disparities. *Urban Education, 52*(2), 207-235.
- 23 Steinberg, M.P. & Lacoë, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 6.
- 24 Bradshaw, C.P., Mitchell, M.M., O'Brennan, L.M., & Leaf, P.J. (2010). Multilevel exploration of factors contributing to the overrepresentation of Black students in office disciplinary referrals. *Journal of Educational Psychology, 102*, 508-520;
- Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for Black and White high school students. *American Educational Research Journal, 48*, 904-934;
- Mattison, E., & Aber, M.S. (2007). Closing the achievement gap: The association of racial climate with achievement and behavioral outcomes. *American Journal of Community Psychology, 40*, 1-12;
- Skiba, R., Chung, C., Trachok, M., & Hughes, R. (2014). Disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal, 51*, 640-670.
- 25 Simon, N.S., & Johnson, S.M. (2015). Teacher turnover in high-poverty schools: What we know and can do. *Teachers College Record, 117*(3), 1-36;
- Miller, A. (2013). Principal turnover and student achievement. *Economics of Education Review, 36*, 60-72.
- 26 Skiba, R.J., Chung, C.G., Trachok, M., Baker, T.L., Sheya, A., & Hughes, R.L. (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.
- 27 Anyon, Y., Wiley, K., Yang, J., Pauline, M., Grapentine, J., Valladares, G., Bell, L., Rosch, A., Cash, D., Downing, B., Greer, E., Kelty, E., Morgan, L. & Pisciotta, L. (2016). *Spotlight on success: Changing the culture of discipline in Denver Public Schools*. Denver, CO: Office of Social and Emotional Learning, Division of Student Services, Denver Public Schools.
- 28 Steinberg, M.P. & Lacoë, J.R. (2017). *Reforming School Discipline: School-Level Policy Implementation and consequences for suspended students and their peers*. Working paper. Retrieved January 16, 2018, from <http://edex.s3-us-west-2.amazonaws.com/%2817.07.17%29%20Reforming%20School%20Discipline%20-%20Lacoë%20Steinberg.pdf>, p. 31.
- 29 Carter, P. L., & Welner, K. G. (Eds.). (2013). *Closing the opportunity gap: What America must do to give every child an even chance*. New York, NY: Oxford University Press.
- 30 Steinberg, M.P. & Lacoë, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 16.
- 31 Steinberg, M.P., & Lacoë, J.R. (2017). Reforming school discipline: school-level policy implementation and the consequences for suspended students and their peers. p. 32. Retrieved January 16, 2018, from <http://edex.s3-us-west-2.amazonaws.com/%2817.07.17%29%20Reforming%20School%20Discipline%20-%20Lacoë%20Steinberg.pdf>

- 32 Advancement Project/Harvard Civil Rights Project. 2000 . *Opportunities suspended: The devastating consequences of zero tolerance and school discipline policies*, 1–48. Retrieved January 17, 2018, from <http://www.eric.ed.gov/PDFS/ED454314.pdf>;
- Anyon, Y., Gregory, A., Stone, S. Farrar, J., Jenson, J., Greer, E., Downing, B., & Simmons, J. (2016). Restorative interventions and school discipline sanctions in a large urban school district. *American Educational Research Journal*, *53*, 1663–1697;
- Christle, C., Nelson, M. & Jolivette, K. (2004). School characteristics related to the use of suspension. *Education and Treatment of Children*. *27*, 509–529.
- 33 Steinberg, M.P. & Lacoë, J.R. (2017). *The Academic and Behavioral Consequences of Discipline Policy Reform: Evidence from Philadelphia*. The Fordham Institute. Retrieved January 2, 2018, from <http://edex.s3-us-west-2.amazonaws.com/publication/pdfs/%2812.05%29%20The%20Academic%20and%20Behavioral%20Consequences%20of%20Discipline%20Policy%20Reform%20Evidence%20from%20Philadelphia.pdf>, p. 5.
- 34 Anyon, Y., Zhang, D., & Hazel, C. (2016). Race, exclusionary discipline, and connectedness to adults in secondary schools. *American Journal of Community Psychology*, *57*, 342–352;
- Gregory, A., Cornell, D., & Fan, X. (2011). The relationship of school structure and support to suspension rates for Black and White high school students. *American Educational Research Journal*, *48*, 904–934.
- 35 McCorry, K. (2015) Philly schools see big shakeup at the top, but principal turnover recedes. *WHYY Public Media*. Retrieved January 16, 2018, from <https://whyy.org/articles/philly-schools-see-big-shakeup-at-the-top-but-principal-turnover-recedes/>;
- Hurdle, J. (2013). Philadelphia officials vote to close 23 schools. *The New York Times*. Retrieved January 16, 2018, from <http://www.nytimes.com/2013/03/08/education/philadelphia-officials-vote-to-close-23-schools.html>