

Discipline for Minor Referrals in Middle Schools:
An Analysis of Disproportionality by Race, Ethnicity, Gender, and Special Education Status

by

Tony Daza

A dissertation accepted and approved in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in Special Education

Dissertation Committee:

Dr. Rhonda Nese, Chair/Advisor

Dr. Cengiz Zopluoglu, Core Member

Dr. John Seeley, Core Member

Dr. Maithreyi Gopalan, Institutional Representative

University of Oregon

Fall 2025

© 2025 Tony Daza

This work is openly licensed via a [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/)



DISSERTATION ABSTRACT

Tony Daza

Doctor of Philosophy in Special Education

Title: Discipline for Minor Referrals in Middle Schools: An Analysis of Disproportionality by Race, Ethnicity, Gender, and Special Education Status

Challenging student behavior is often managed through staff referrals which culminate in teacher and school administrator responses. Referrals are divided into major and minor referrals, with minor referrals designed for teachers to address low-level behaviors in the classroom and major referrals designed for school administrators to address elevated and repeated behaviors outside of the classroom. Prior research has demonstrated disproportionate responses to major referrals, often leading to exclusion from the learning environment for students of color and students receiving special education services (those with Individualized Education Plans/IEPs). This study explored patterns of discipline for minor referrals by examining data across middle schools in four states to understand the risk of receiving exclusion for students from minoritized and underserved groups.

The study incorporated multiple measures of disproportionality including Risk Ratios, Raw Differential Representation, and Odds Ratios from a multi-level model to determine the presence of disproportionate responses. Consistent with prior research on major referrals, results showed that Black students as well as those with IEPs received exclusion for minor referrals at higher rates than their peers, but when accounting for student and school level predictors, the odds of exclusion for minor referrals were not significantly different than White male students without IEPs. In fact, referrals for students with IEPs were at reduced odds of exclusion based on the model. Moreover, exclusionary discipline rates varied significantly between schools and by state, with referrals for

students in Georgia having significantly higher odds of resulting in exclusionary discipline and referrals for students in Massachusetts having lower odds of resulting in exclusionary discipline. Although the state differences remained, the addition of PBIS Tier 1 fidelity of implementation as a predictor showed schools with higher fidelity PBIS implementation were significantly less likely to lead to exclusion in Georgia. These findings demonstrate the potential for PBIS to reduce exclusionary discipline practices and disparities by student characteristics for minor referrals.

CURRICULUM VITAE

Tony Daza:

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
St Thomas University, St. Paul, MN
Carleton College, Northfield, MN

DEGREES AWARDED:

Doctor of Philosophy, Special Education, 2025, University of Oregon
Master of Education, Special Education, 2018, St. Thomas University
Academic Behavior Strategist (License), Special Education, 2018, St. Thomas University
Bachelor of Arts, Biology, 2014, Carleton College

AREAS OF SPECIAL INTEREST:

Transition Supports

Transition Assessments & Interventions for Culturally Diverse Students

Positive Behavior Interventions and Supports (PBIS)

Culturally Responsive School Wide Positive Behavior Practices

Teacher Preparation

Secondary Teacher Preparation for Behavior Support

PROFESSIONAL EXPERIENCE:

University of Oregon

Student Life Assessment & Research Extern, *Office of Student Life Research*, 11/23 – Current

ISLA Facilitator, *Salem-Keizer*, 9/23 – 5/24

ISLA Facilitator, *Oregon City*, 9/22 – 6/23

Student Coach Project Elevate, *Eugene*, 1/23 – 6/23

Data Collector Automated Performance Evaluation of Early Language (APEL), 9/21 – 5/24

Data Collector & Trainer Computer Adaptive Storybook Assessment (CASA), 9/22 – 5/24

Monterey County Office of Education

Teacher Induction Program Content Creator, *virtual*, 9/24 – 10/24

Portland State University

Program Support & Design National Technical Assistance Center on Transition Collaborative (NTACT:C), *virtual*, 12/23 – 5/25

University of Arizona

Data Analyst Canopy Project, *virtual*, 4/23 – 8/23

Minneapolis Public Schools

Special Education Resource Teacher, *Sanford Middle School*, 9/18 – 8/21

Behavior Specialist - 6th Grade Dean, *Sanford Middle School*, 8/16 – 6/17

Special Education Resident Teacher, *Andersen United Community School*, 6/17 – 8/18

Northfield Public Schools

8th Grade Summer Blast Teacher, *Northfield Summer Blast Program*, 6/18 - 8/18

TORCH Program AmeriCorps Promise Fellow, *Northfield Middle School*, 9/14 - 8/16

Afterschool Site Assistant, *Greenvale Community School*, 9/14 - 5/16

GRANTS, AWARDS, AND HONORS:

Mary Wiese Endowed Prize, *Carleton College*, 2014

Awarded to a student who demonstrates a respect for education, a love for students, and the understanding of the place of education in their future. In addition, the student must embody the following qualities: cultural pride, kindness, perseverance, self-reliance, discipline, and respect and care for other people.

PUBLICATIONS:

Avelar, J., Zyskind, K., & Daza, T. (2023). *Prioritizing multilingual learners: How inclusive schools meet the needs of bilingual students and English learners*. Canopy Schools.

<https://canopyschools.org/research-2023>

Nese, R. N.T., Nese, J. F.T., Santiago-Rosario, M.R., Izzard, S., Newson, A, Pimentel-Mannan, I, Cohen Lissman, D, Daza, A, Triplett, D, Malose, S. (2025). *Improving Classroom Engagement and Reducing Exclusionary Discipline: A Quasi-Experimental Study of the Inclusive Skill-Building Learning Approach During COVID-19*. *Journal of Education*. 205(2), 114-129. <https://doi.org/10.1177/00220574241309893>

Santiago-Rosario, M. R., Izzard, S., Newson, A., Daza, T., & Nese, R. N. T. (2025). *Student experiences with the Instructional Skill-building Learning Approach: An alternative to exclusion*. *Adolescents* 2025, 5, 19. <https://doi.org/10.3390/adolescents5020019>

ACKNOWLEDGMENTS

I would like to give thanks and acknowledge all those who have helped me along the way. From my first foray into the world of education through community tutoring through a high school Spanish class I have been surrounded by role models who exemplify the values of being an educator. In particular, my time volunteering under the guidance of Susan Sanderson of the TORCH Program sparked my passion for working with youth. After moving to Minneapolis Public Schools, I had the opportunity to learn the role of a behavior specialist under the guidance of Ms. Melissa Becker. The next step in my journey led me to Dr. Shelley Neilsen Gatti who taught me the ins and outs of how to be a special educator. Back in Minneapolis Public Schools I was guided in my first years as a teacher by my co-teacher Ms. Janell Momsen. My next leap brought me to the University of Oregon to pursue a doctorate in Special Education.

At UO I received invaluable experience in conducting research and analyzing data. I had the chance to experience research studies firsthand with Dr. Lillian Duran and Dr. Rhonda Nese. I not only learned how to design and collect data but through opportunities at UO I learned how to conduct statistical analysis of data through Dr. Brian Clark. I would not have been able to prepare this manuscript without the support of my adviser Dr. Rhonda Nese along with the support of my committee, Dr. Cengiz Zopluoglu, Dr. John Seeley, and Dr. Maithreyi Gopalan. An additional thank you to Dr. Sean Austin who was critical in the brainstorming and planning stage of the study.

I want to acknowledge the support of my family who have given me their love and support throughout my educational career. When I was young, my mom taped a cutout of the phrase “Knowledge is power,” to my mirror and she has been my biggest champion throughout my career. Thank you as well to my dad, who immigrated to the US to study, and imbued his

respect for the field of education to transform lives. Finally, thank you to my wife, Monica for her continued patience, love, and support. Without your support I would not be able to complete this challenging journey.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION AND LITERATURE REVIEW	16
History of Behavior Management.....	16
The Rise of PBIS	17
The Impact of NCLB and ESSA on Suspension	19
Measuring Inequality	20
Disproportionality Metrics.....	20
Exclusionary Response to Behavior	22
Office Discipline Referrals (ODRs).....	24
ODRs by Student Race and Ethnicity	24
ODRs by Student Gender.....	26
ODRs by Student Disability Status.....	27
The Impact of ODRs on Students	28
Academic Impact	28
Social Emotional Impact.....	29
Exclusionary Discipline	31
Exclusion by Student Race and Ethnicity	33
Exclusion by Student Gender.....	34
Exclusion by Student Disability Status.....	35
The Impact of Exclusionary Discipline on Students.....	36
Academic Impact of Exclusion	36
Social Impact of Exclusion	37

Moving Away from Exclusion.....	38
The Rise of Suspension Bans.....	38
The Impact of Alternatives to Exclusionary Discipline.....	39
Minor Behavior Referrals	41
Limited Research on Minor Behavior Referrals	43
The Impact of Teacher Implicit Bias in Minor Referrals	43
Purpose of Study	45
Research Questions	46
II. METHODOLOGY	48
Settings.....	48
Sample Selection.....	48
Sample Preparation	49
Participants.....	51
PBIS Subsample for RQ5	54
Key Variable Processing	56
Student Race and Ethnicity	56
Challenging Behavior Variables	57
Defining Exclusionary Discipline.....	57
PBIS Fidelity Scores	58
Missing Data	62

Chapter	Page
Analytic Software	64
PBIS Subsample for Analysis	66
Measurement.....	67
Minor Referral Identification.....	67
Exclusionary Discipline Variable	68
Dependent Variables by Research Question.....	68
Data Analytic Plan	70
Overview of Data Analysis.....	70
Risk Ratio Calculation	71
Raw Differential Representation Calculation	71
RQ4 Multi-level Logistic Regression Model.....	72
RQ5 PBIS Moderation Multi-level Logistic Regression Model.....	74
III. RESULTS	78
RQ1: Patterns of Challenging Behaviors in Minor Referrals	79
RQ2: Patterns of Administrative Actions in Minor Referrals	84
RQ3: Differences in Major and Minor Referral Patterns.....	92
Overview of Differences in Exclusionary Discipline	92
State Differences in Exclusionary Discipline	93
Student Demographic Differences in Receiving a Major Referral	95
State Differences in Receipt of a Major Referral.....	97
Intersectionality by Race, Gender, and IEP Status	100
RQ4: Differences in Exclusionary Discipline for Minor Referrals	102

Exclusionary Discipline for Minor Referrals by State.....	103
Intersectional Differences in Exclusionary Discipline	107
State Level Differences in Odds of Exclusionary Discipline	108
Student Demographic Differences in Odds of Exclusionary Discipline	109
RQ5: Moderation of Exclusionary Discipline from PBIS Implementation.....	111
Tier 2 PBIS Fidelity of Implementation	114
IV. DISCUSSION.....	116
The Impact of Teacher Implicit Bias on Subjective Referrals.....	118
Differences Between Major and Minor Referrals.....	120
Intersectionality.....	123
Minor Referral Exclusionary Discipline	124
PBIS Implementation.....	127
Implications for Research	129
Implications for Practice	131
Limitations	132
Future Research	136
Recommendations.....	138
Summary	139
APPENDIX: R Package Citations	141
REFERENCES CITED.....	144

LIST OF FIGURES

Figure	Page
1. Differences in Exclusion Rates by Ethnicity and Gender, 2020-2021 OCR	33
2. School-level Distribution of Tier 1 and Tier 2 Fidelity scores	60
3. School-level Distribution of PBIS schools by state and tier	61
4. Distribution of Exclusionary discipline rates by school and state	64
5. Frequency of Challenging Behaviors for Minor Referrals	80
6. Frequency of Administrative Actions for Minor Referrals.....	85
7. Ratio of Major to Minor Referral Risk by Race and State	100
8. Risk Ratios for Major Referral and Major and Minor Referral Exclusion by State and Student Group	106

LIST OF TABLES

Table	Page
1. Sample demographic characteristics by referral, student, and school	53
2. Frequency of challenging behaviors for minor referrals by State.....	80
3. Challenging Behaviors for Major and Minor Referrals by State	82
4. Frequency Administrative Responses to Minor Referrals	85
5. Administrative Response to Major and Minor Referrals by State	87
6. Exclusionary Discipline from Minor Referrals by State.....	88
7. Descriptive Statistics of Student Characteristics by State	89
8. Examining exclusion rate for the interaction of student characteristics for minor referrals.....	91
9. Exclusionary Responses by Student Group for Major and Minor Referrals	93
10. Exclusion rates of major and minor referrals by student demographics by state	94
11. Combined overall RR and RDR by Student Demographic for Receipt of Major or Minor Referral	96
12. Risk Ratio and Raw Differential Representation (RDR) by Race and State for Risk of Major or Minor Referral	99
13. Risk Ratio and Raw Differential Representation by Race, Gender, and IEP Status Interaction for Major Referrals	101
14. Disparity in Exclusionary Discipline for Minor Referrals by Race, Gender, and IEP Status	103
15. Disproportionality in Exclusionary Discipline for Minor Referrals by State	105
16. Risk Ratios & RDR for Exclusion for Minor Referrals by Race, Gender, and IEP Status	107
17. Predicting Exclusion by Race, IEP Status, and Gender with State Fixed Effects	110

18. PBIS Tier 1 Fidelity: Predicting Exclusion with and without State Interaction	113
19. PBIS Tier 2 Fidelity: Predicting Exclusion with and without State Interaction	115

Chapter I: Introduction and Literature Review

Students in schools across America continue to receive out-of-school suspensions as a method of managing unwanted behaviors. A substantial body of evidence has demonstrated that suspensions have a detrimental impact on student academic performance and sense-of-belong (Del Toro & Wang, 2022; Gale & Nepomnyaschy, 2024; Gregory et al., 2010; St-Amand et al., 2022). An additional concern is the presence of disparities in school discipline practices, with minoritized students receiving disproportionately more suspensions than their White peers for the same unwanted behaviors (Carter et al., 2016; Gopalan & Nelson, 2019; Leung-Gagné et al., 2022; Skiba et al., 2014; Welsh & Little, 2018). Disproportionate exclusionary discipline negatively affects students' academic and social outcomes, increasing the likelihood of school failure and incarceration (Bergh & Cowell, 2013; Lacoë & Steinberg, 2019; Losen & Martinez, 2020). Prior research has primarily focused on the patterns of exclusionary discipline for major referrals, but this study examines disproportionality in the response to minor behavior referrals.

History of Behavior Management

Pinpointing the beginning of exclusionary discipline in schools is challenging, but out-of-school suspension (OSS) as a practice to manage behavior is believed to have started in the 1960s in the United States (Allman & Slate, 2011). Prior to that time, behavior management was teacher specific and could include corporal punishment or other consequences as school administrators deemed appropriate. During the 1960s suspension became a prominent behavior management tool, despite evidence showing that the practice did not improve student behavior (Anderson, 2018; Craig & Martin, 2023; Freiberg & Lapointe, 2006). Although *Gross vs Lopez*, a case heard by the supreme court in 1975, gave students the right to due process in suspensions under the 14th amendment (Oyez, n.d.), the use of suspension rose steadily through the 1980s and

1990s. This rise was in part due to zero-tolerance policies, which mandated suspensions for specific behaviors (Leung-Gagné et al., 2022). For example, in 1994 the Gun-Free Schools Act was signed, mandating expulsion for a minimum of a year for students who brought firearms to school (Sone, 2023). The zero-tolerance policies with strict predetermined consequences, often leading to suspension, were implemented in response to rising concerns about school safety (Kang-Brown et al., 2023; Schiff, 2018). However, with the prevalence of zero-tolerance policies there was heightened scrutiny on the long-term consequences and fairness of the approach. A large body of evidence showing detrimental impacts of suspensions on students led to a more recent push toward banning the use of suspensions for certain behaviors. As of 2021, fifteen states and the District of Columbia (Kelley et al., 2021) have laws limiting the use of OSS, primarily for elementary school students – though no state has banned the use of suspensions completely (Mata, 2024). In 2011 Los Angeles Unified School District issued a ban on suspension for willful defiance and a subsequent study found the ban demonstrated the ability of such policies to decrease suspension rates (Hashim et al., 2018). Although there was a decline in the overall suspension rate, suspension gaps between Black and non-Black as well as between students receiving special education and those without persisted according to the authors.

The Rise of PBIS

Around this time, Positive Behavioral Interventions and Supports (PBIS), a prevention framework based on the public health model, began gaining traction to preemptively manage challenging behaviors. In 1997, amendments to the Individuals with Disabilities Education Act (IDEA) cemented the terms positive behavioral interventions and supports (PBIS) and positive behavior supports (PBS) into the practices of American schools (Sugai & Horner, 2002). PBIS stems from behavioral science and applied behavior analysis to proactively address student

challenging behavior through school-wide systems and classroom management strategies rather than reactive forms of discipline that have been ineffective at reducing challenging behaviors, such as zero tolerance policies, the use of suspension, and alternative placement programs (Sugai & Horner, 2002). PBIS has become a widely adopted framework for improving school climate and promoting equity through school-wide practices (McIntosh et al., 2021b). Effective PBIS implements a multi-tiered system of supports (MTSS) that incorporates interventions at varying levels of intensity to meet student needs (Center on PBIS, 2021). Tier 1 focuses on universal prevention using school-wide behavioral expectations and reinforcing positive behaviors in the classroom and other school settings. Tier 2 provides specific interventions for students demonstrating greater need than Tier 1 alone, often this is in the form of targeted group interventions, and Tier 3 offers individualized supports for students with more significant needs. Across all tiers, PBIS promotes school-wide practices that focus on positive reinforcement, cultural responsiveness, family engagement, and the use of data for continuous improvement.

Due to increased implementation of PBIS across the United States as an alternative approach to behavior management, researchers have examined the impact of the intervention on exclusionary discipline for students. Multiple studies have demonstrated the ability of PBIS school-wide intervention systems to reduce the use of exclusionary discipline and lead to more equitable outcomes (Gage et al., 2019; McIntosh et al., 2021a,b). For example, Austin and colleagues (2024) found a PBIS intervention aimed at changing teacher perceptions of minoritized students led to a reduction in exclusionary discipline referrals for Black students. The impact of PBIS interventions to support schools and teachers to proactively manage challenging student behavior suggests that PBIS implementation may moderate discipline disparities.

The Impact of NCLB and ESSA on Suspension

Fast forward to the late 1990s and early 2000s suspension rates continued to rise until reaching a peak and subsequent decline in the early 2010s (Gopalan et al., 2024; Leung-Gagné et al., 2022). In 2001, No Child Left Behind (NCLB) was signed into law resulting in increased pressure on teachers and schools to improve test scores for students with penalties in place for underperforming schools. Research suggests that this shift in policy may have led to increased student behaviors and disproportionate punishment of minoritized students (Holbein & Ladd, 2017) by increasing the pressure on students and an incentive for teachers to remove students from the classroom to focus on content rather than addressing challenging behavior.

In 2016 in the wake of NCLB, the Every Student Succeeds Act (ESSA) allowed states to shift the focus from solely academic metrics to include measures of success based on student socio-emotional well-being, such as school climate (Gayl, 2017), paving the way for increased implementation of PBIS. Although PBIS implementation has increased in the subsequent years and suspension rates overall have decreased, suspension and office discipline referral (ODR) rates continue to show disproportionately high rates for Black students and students receiving special education services (Brobbe, 2018). More recently, the COVID-19 pandemic disrupted the status-quo of education in America through the implementation of online learning. The impact of the shift to online learning is still being studied, but early research suggests the pandemic increased mental health concerns for students with a rise in internalizing behaviors such as anxiety and depression observed and evidence for an increase in externalizing behaviors for students with pre-existing behavior disorders (Bera et al., 2022). The US education system continues to adapt to the post COVID-19 era with schools focusing on supporting student mental

health and academics with a continued effort to implement PBIS while reducing discipline disparities for minoritized students (McIntosh et al., 2021a,b).

Measuring Inequity

The increase in exclusionary discipline and the detrimental impact on minoritized students has led to research on disproportionality of office discipline practices on various groups of students such as students of color, males, students with disabilities, socioeconomic status, and LGBTQ youth. Disproportionality is defined as the difference in treatment of students from different race, ethnic, and identity backgrounds, in this case, specifically in terms of school discipline (Schollenberger, 2015). In the literature, disproportionality has been measured in numerous ways and for different outcomes, leading to discrepancies in findings reporting on disproportionality. For discipline outcomes there could be disproportionality in the rate of out-of-school suspension (OSS), in-school suspension (ISS), office discipline referrals (ODRs), or exclusionary discipline outcomes. Due to the long history of exclusionary discipline practices in the US, there is a robust literature base which has repeatedly demonstrated a disproportionate rate of office discipline referral and suspensions for students of minoritized races, particularly for Black students (Barrett et al., 2021; Girvan et al., 2017; Gopalan & Nelson, 2019; Liu et al., 2022; Losen et al., 2014; Skiba et al., 2002, 2011, 2014).

Disproportionality Metrics

Despite the research on the presence of disproportionality; there are differing opinions on the definition and measurement of disproportionality by researchers (Curran, 2020; Girvan et al., 2019). There are several metrics commonly used in the literatures regarding disproportionality. The most often reported are risk ratios (RR) as they are intuitive to interpret, easy to calculate, and can be used to compare across school contexts (Curran, 2020). For example, a risk ratio of

1.5 means for Black students compared to White students for suspension can be interpreted as a 50% increase in risk for Black students to be suspended. Risk ratios have been codified as the metric of choice by the department of education (Department of Education 2017), but they are sensitive to small numbers of students and do not provide information on absolute rates or numbers of individuals impacted (Girvan et al. 2019). Furthermore, risk ratios alone are inadequate for assigning racial disproportionalities as they demonstrate a risk for a specific group but not a comparative risk (Bottiani et al., 2023; Girvan et al., 2019). The researchers outline a need for a wider range of reported metrics on disproportionality to uncover the nuances of disproportionality, which is necessary for informing education policy and practice. These other metrics include risk difference (RD), difference in standardized risks (probit d'), raw differential representation (RDR), rate of exclusionary discipline events, and alternate rate of risk exposure.

Girvan and colleagues (2019) recommend reporting multiple metrics and highlight raw differential representation as the primary measure of disproportionality. RDR allows for comparisons of a target group to a reference group, such as Black students to White students, in an easy to interpret number. Although the measure does not inherently account for differences across schools and districts, it may be standardized to show the difference between groups per 100 students, allowing for across context comparison (Girvan et al., 2019).

Although the department of education has used risk ratios as the metric of choice for reporting disproportionality, new research has demonstrated that risk ratios do not capture the full picture of disproportionality and is not stable over time as a measure to examine trends in disproportionality (Bottiani et al., 2023; Girvan et al. 2019). Instead, risk ratios can be one of a set of metrics reported along with raw differential representation (RDR) and discipline event rate, which provide minimal overlap and maximum information on disproportionality in the data

(Girvan et al. 2019). According to Bottiani and colleagues (2023) one of the benefits of RDR is the correlation between the metric and a measure of student connectedness and equity which was not the case for other measurements of disproportionality, suggesting RDR may be better suited for research examining the effects of positive school wide behavior interventions.

Another option for standardizing the risk ratio is through an alternate risk ratio (ARR), which is calculated using the number of students in the target group receiving the treatment versus those in the target group not receiving the treatment instead of total number of students in the target group, to create a measure of risk of exposure to exclusion across different schools (Bottiani et al., 2023). However, ARR cannot be calculated if the total number of students in the population is unknown, for example in data sets containing all students who received a referral. These metrics allow for more robust examination of disproportionality but require the number of enrolled students by demographic category as well as total enrollment.

In contrast, a logistic regression model will output the log odds of the outcome which can be transformed into an odds ratio for reporting disproportionality for target groups. Odds ratios are similar to risk ratios, but with a subtle difference. The risk of suspension for example, compares the likelihood of suspension between groups, whereas the odds ratio compares the odds of an event occurring vs not occurring. In short, the odds ratio is the ratio of probability of suspension to the probability of not being suspended. In practice, risk ratios and odds ratios are often similar, but the odds ratio is a natural outcome of the logistic regression that allows for a comparison of different student groups.

Exclusionary Responses to Behavior

Exclusionary responses to behavior are often conceptualized as suspensions, usually OSS and expulsion, but exclusion exists as a continuum of severity (R.N.T. Nese et al., 2021).

Exclusion includes any response to behavior that removes a student from the classroom instructional environment such as time in the office, detention, and in-school suspension. In addition to the exclusionary responses to behavior, there are also differences in the severity of referral based on the challenging behavior with referrals often being categorized as major and minor referrals and occasionally as subjective and objective behaviors. Major referrals, often referred to as office discipline referrals (ODRs), are more likely to result in the commonly researched exclusionary discipline consequences such as OSS, ISS, and expulsion, whereas minor referrals are more often handled in the classroom or given consequences such as a conference with a teacher, parent contact, or detention. Though parent communication and conferencing with school staff are integral responses to major referrals as well. This study aims to examine the patterns of discipline for minor referrals. For an examination of the impact of exclusion from academic instruction due to minor behavior referrals we must first review the findings of suspensions and ODR research to understand the consequences of exclusionary discipline.

The research on exclusionary discipline often focuses on two outcomes, ODRs and suspensions, which clearly show a negative impact on student academics as a result of suspension. ODR research is less clear in terms of long-term outcomes but demonstrates disparities for rate of referral for Black students, males, and students with IEPs, with interactions between the groups as well. In particular, Black boys and students with an emotional behavioral disorder (EBD) designation are more likely to receive office discipline referrals (ODRs) suspensions. Additionally, it is hypothesized that students who receive an ODR are more likely to be suspended in the future, as the authors described it, “an ODR is like a ticket to the suspension lottery,” (Welsh & Rodriguez, 2024). The authors also highlight that disparities may

be even greater for receipt of ODR for students from certain groups than the disparity in suspension rates by student characteristics.

Office Discipline Referrals (ODRs)

The concerning impact of exclusion on student well-being has led to extensive research focused on school suspensions as the primary unit of analysis in ODR research (Noltmeyer et al., 2015). However, the majority of discipline referrals do not result in OSS, instead focusing on in-school consequences. The long term consequences of ODRs, and more specifically, minor referrals, on student outcomes remain largely underexplored, though minor referrals are likely to result in exclusion through a short removal from class due to a conference with behavior staff, time in a break room or in the office, or longer removal in the form of detention or ISS (R.N.T. Nese et al., 2021; Wang et al., 2023). These responses to referrals, though not as severe as OSS, can disrupt and limit a student's learning. Missed instructional time compounds throughout a student's educational career leading to increased chance of academic failure and feelings of not-belonging in the classroom environment (R.N.T. Nese et al., 2021). Combined with the lack of skill building opportunities present in commonly used exclusionary discipline practices, ODRs can increase the likelihood of receiving a suspension and missing instructional time (Amemiya et al., 2020). Understanding whether the punitive disciplinary measures of minor referrals support or hinder students' success in school is a critical, yet unanswered, question. Furthermore, understanding the disparate impact of minor referrals on students from different backgrounds and identities is important for developing interventions to support all students.

ODRs by Student Race and Ethnicity

Research suggests that Black students are more likely to receive ODRs, specifically for subjective behaviors, and suspension than their White peers (Bradshaw et al., 2010; Legette &

Anyon, 2023; Welsh & Rodriguez, 2024). Data show that despite initiatives to eliminate disproportionality, Black students are given ODRs at a 2-3 times higher rate than White students (Skiba et al., 2011). Although not as pronounced, data also suggests the existence of disproportionate discipline of students from Latiné and Native American backgrounds (Gage et al., 2021; Gion et al., 2018). Studies suggest these three groups are all disciplined at a higher rate than White and Asian students (Liu et al., 2022). Although some have argued differential involvement in challenging behavior can explain the difference in office discipline referrals, an analysis of high school data by Huang (2018) found that Black students and White students exhibited similar levels of attitudes supporting deviant behavior, indicating student behavior alone does not explain the difference in referral rates. One author (McConnell, 2022) argues instead that cultural disconnect and implicit bias contribute to the disproportionate ODR rate for students of color.

An analysis of the mechanism by which racial disparities are perpetuated suggest teacher perception of student behavior by race may play a role (Kunesh & Noltemeyer, 2019; Okonofua & Eberhardt, 2015). Studies suggest implicit bias, which is an underlying, often subconscious bias, may lead to teachers paying closer attention to students of a different race, leading to disparities in discipline. For example, one study found that when prompted to look for challenging behavior teachers spent more time watching Black students (Gilliam et al., 2016). Other research has confirmed minoritized students are subject to increased surveillance from teachers, potentially leading to increased behavior discipline (Zimmermann, 2018, 2024). These underlying biases can lead to disproportionate discipline for minoritized students and the disproportionate rate of ODRs for students by race leads to negative impacts on student academic, social, and school outcomes.

ODRs by Student Gender

In addition to racial disparities, there are also documented differences in rates of referral by gender as well as sexual orientation (Samimi et al., 2024). In elementary and middle school, boys consistently receive ODRs at a higher rate than girls, though the exact rate varies by school and state context (Kaufman et al., 2010). Kaufman and colleagues (2010) posited one theory for the difference in ODR rate. They speculated that the higher rates of externalizing behavior exhibited by boys compared with girls, led to more overt behavior that is easier for teachers to observe and results in higher levels of disciplinary consequences. Less is known about the impact of ODRs on LGBTQ youth. A scoping review of the literature found girls in the LGBTQ community were more likely to experience exclusionary discipline, but the study highlighted the limited research on the topic (Samimi et al., 2024). Another study examining the risks faced by students based on sexual orientation found that same-sex attraction had higher odds of discipline for girls, but no increased risk for boys (Mittleman, 2018).

The interaction between race and gender may also lead to different rates of ODRs. A study examining the predicted probability of ODRs was significantly higher for boys compared to girls with Black boys being the most at risk (Lehmann, 2023). A study by Bradshaw and colleagues (2010) found Black boys were significantly more likely to receive an ODR compared to White boys. Other research suggests Black girls and Latinas are subject to stereotypes of being aggressive and loud (Morris & Perry, 2017) and Black and Latino boys are stereotyped as dangerous and lazy (Reny & Manzano, 2016). These dangerous stereotypes may explain racial bias in ODRs leading to discipline gaps (Legette & Anyon, 2023). In a study examining race and gender Morris and Perry (2017) found Black boys in middle and high school were twice as likely to receive ODRs for subjective behaviors compared to White boys. Black girls were three times

more likely to receive ODRs compared to White girls. Research has shown gender may play a role in the likelihood of receiving an ODR, with boys more likely than girls to receive a referral, as well as the potential for different rates of ODRs based on student race and gender interaction as well as membership to minority groups.

ODRs by Student Disability Status

Students with IEPs are a proxy used to represent students with disabilities as the IEP is an indicator of a student receiving special education services. Although students with IEPs often have social and academic needs, they are also more likely to receive ODRs as a result of their label (Gage et al., 2017). Research examining disproportionality based on IEP status for ODR rate, but students with disabilities have been shown to be at greater risk of suspension, though the rate does vary by disability category. The disproportionality may be partly explained by the wide array of characteristics by disability type. For example, there is evidence in the literature describing the increased risk of suspension and discipline for students with emotional and behavioral disorder (EBD) (Forman & Markson, 2022).

In fact, in a disaggregated examination of suspension disproportionality by disability status, researchers found students with EBD and those with externalizing behaviors associated with ADHD were more likely to be suspended compared to non-disabled peers, but other disability categories did not have increased risk of suspension (Krezmien et al., 2017). This suggests the disproportionality present in discipline for students with IEPs may be primarily a result of the increased risk for students with EBD. A longitudinal study examining the association between special education designation and discipline outcomes, unsurprisingly given the criterion for EBD qualification, found a close link between prior externalizing behaviors and EBD designation, but also noted a potential decrease in exclusionary discipline after designation

(Hurwitz et al., 2021). This decrease after designation may be due to increased leniency for students with IEPs or increased interventions by case managers to address challenging behaviors. The research highlights the need to increase protections for students with disabilities and to shift from a model that requires students to fail before being referred for additional support.

Impact of ODRs on Students

Office Discipline Referrals (ODRs) are used in response to challenging student behavior, leading to a variety of consequences for students ranging from expulsion to a conference with behavior staff. A common outcome of ODRs is the removal from classroom instruction, though the lengths of time removed vary based on the referral, meaning the impact of an ODR will differ depending on contextual factors. Because the consequences of an ODR vary based on behavior—and likely based on student race/ethnicity, referring teacher, and other factors—ODRs should not be considered equal units of measurement, nevertheless the removal from academic instruction coincides with negative outcomes for students receiving ODRs.

Academic Impact

It is not surprising challenging behavior and receipt of an ODR have a negative impact on student academics. In an examination of the relationship between ODRs and academic performance, Whisman and Hammer (2014) found students with one or more ODR had a 70% probability of scoring below proficiency for math. In fact, even one ODR put students at increased odds to score below proficient in math. Office discipline referrals often result in removal from class, as well as additional consequences depending on the behavior, which lead to loss of classroom instructional time (R.N.T. Nese et al. 2021). This loss of instructional time can have negative long-term effects on student academic scores. One study by Kremer and

colleagues (2016) found externalizing behaviors had a negative relationship with academic performance with enduring effects over time.

One potential mediating factor between the negative impact of ODRs on academic performance is student perception of school climate. Students who received a suspension for a minor referral—which Black students were significantly more likely than White students to receive—had lower academic scores in the subsequent years (Del Toro & Wang, 2022; Wang et al., 2023). However, when the authors examined student perception of school climate, they highlighted two important findings: students who received a suspension for a minor referral had lower perceptions of school climate and students with higher perceptions of school climate were protected from the negative impact on academics. This suggests both that school climate is critical to student success and that ODRs may negatively impact student perception of school climate. According to the study, students who receive an ODR and perceive it as unjust are more likely to rate school climate negatively and as unfair.

Social Emotional Impact

Although most research has examined the impacts of suspensions on student social well-being, some studies suggest that persistent ODRs may also be detrimental to student's social and emotional health. For example, in a study by Bottiani and colleagues (2017) examining school perceptions of Black and White students at schools with disproportionate use of referrals, they found Black students were significantly more likely to have increased negative perceptions of school equity and lower levels of sense of belonging compared to white students.

Students from minoritized gender roles such as students who identify as LGBT+ may experience greater social impacts from ODRs. In a recent study (Herry & Mulvey, 2025) the authors found students who identified as gay, lesbian, or bisexual had lower perceptions of

school climate and were more likely to report being discriminated against by staff and peers. In particular, boys were more likely to report experiencing discrimination from teachers leading to negative perceptions of teacher relationships. The study noted students from the LGB community expressed lower perceptions of equality in school discipline, fewer opportunities for engagement, academic support, and school connectedness. These findings suggest students from minoritized backgrounds due to sexual orientation may experience the effects of ODRs more acutely due to discriminatory practices.

The impact of ODRs on social emotional outcomes based on gender and race suggest students who are from minoritized backgrounds may experience bias differently based on gender. For example, Black boys are perceived as dangerous and are referred at higher rates than White boys leading to increased disconnection from school and ultimately an increased likelihood of the school to prison pipeline (Monahan et al., 2014). Although boys in generally tend to receive more referrals for disruption and fighting compared to girls (Kaufman et al., 2010), Black and Latina girls receive higher rates of referrals for fighting and subjective referrals compared to White girls (Legette & Anyon, 2023). Additionally, girls may be more likely to experience social and psychological bullying compared to boys who have higher rates of referral for physical aggression, leading to negative social consequences (Card et al., 2008). In an examination of Black girls' perception of teacher decisions surrounding subjective behaviors, the students noted a lack of voice and disengagement from the learning environment as a result of negative interactions with teachers (Griffith, 2023). Additionally, research suggests social emotional competency interventions may benefit students and limit challenging behavior, but social emotional competency alone did not explain differences in discipline by race and gender

(Davis et al., 2024). This suggests that despite differences in student social emotional skills, ODRs were still impactful and disproportionately harmed students from minoritized groups.

Exclusionary Discipline

Existing evidence underscores the negative impact of suspension as a form of discipline (Gerlinger et al., 2021). Although these effects are well-documented for suspensions, the impact of minor referrals on students is less well-studied. Referrals that do not result in suspension may lead to less severe but still exclusionary consequences, such as removal from class for time in the office, detention, or in-school suspension (R.N.T. Nese et al., 2021). These responses disrupt students' learning, reduce instructional time, and may foster feelings of disconnection from the classroom environment (Amemiya et al., 2020). When compounded over time, such disruptions can increase the risk of academic failure and exacerbate the cycle of exclusionary discipline, making students more likely to receive a suspension in the future. The negative effects are most pronounced for minoritized students —Black, male, and students receiving special education services—who are more likely to be suspended and less likely than their white peers to complete high school (Girvan et al., 2017; Noltemeyer et al., 2015). Research on exclusionary discipline practices has demonstrated the negative impact of suspension, which has led to policies banning the use of suspension gaining popularity and an overall decrease in the practice of suspension. Other forms of exclusionary discipline likely impact students negatively as well due to loss of instructional time, yet there is a dearth of research on the topic of the impact of non-suspension responses to ODRs.

Both ISS and OSS, collectively referred to as suspension, negatively impact students and disproportionately effect minoritized students, particularly those from minority racial and ethnic groups as well as those with IEPs. Suspension rates are lower today than in the past fifty years,

but disparities for rates of suspension by student group persist (Gregory et al., 2018; Leung-Gagné et al., 2022). Examining the historical OSS data from the Office for Civil Rights (OCR) a pattern of rising suspension rates can be observed from 1973 until 2010 —spurred by an increase in zero tolerance policies —with a peak around 7 percent (Welsh & Little, 2018). In subsequent years the suspension rate has fallen to 5 percent as of 2017, and most recently, the out-of-school suspension rate for K-12 students was down to nearly 1 percent (Civil Rights Data Collection, 2023; Leung-Gagné et al., 2022). Despite the drop, the suspension rate continues to vary significantly by student group. For example, according to the Office of Civil Rights (OCR), the national suspension rate for White students in 2021 was 1.42 percent while the rate for Black students was 2.3 percent. This difference fluctuates by states with some states, such as Iowa and South Dakota, having suspension rates for Black students at around 6 percent compared to White students maintaining a rate of about 1 percent (National Center for Education Statistics, 2024). In addition, students who received services under IDEA—students receiving educational services for a disability— represented 14 percent of enrolled K-12 students but comprised almost 25 percent of the students receiving OSS. Students with disabilities and Black students were also overrepresented in referrals to law enforcement and school related arrests (Civil Rights Data Collection, 2023). The data highlight the reduction in OSS for students over the years and the persistence of disproportionality based on student race and disability status. In figure 1 below data of exclusionary discipline for K-12 students split by gender and race are shown, demonstrating the higher rate of Black male and female suspension given their enrollment compared to other racial groups. In particular, Black students had a disproportionately high rate compared to other minority groups such as Latiné, Asian, and Pacific Islander students.

Figure 1

Differences in Exclusion Rates by Ethnicity and Gender 2020-2021 OCR

Figure 3. Percent of K-12 students who were disciplined in public schools, by sex and race/ethnicity

		Hispanic or Latino of any race	American Indian or Alaska Native	Asian	Native Hawaiian or Other Pacific Islander	Black or African American	White	Two or more races
Boys	Enrollment	15%	<1%	3%	<1%	8%	24%	2%
	In-School Suspensions	13%	1%	<1%	<1%	15%	37%	3%
	Out-of- School Suspensions	12%	1%	1%	<1%	18%	36%	4%
	Expulsions	14%	1%	1%	<1%	18%	34%	3%
Girls	Enrollment	14%	<1%	3%	<1%	7%	22%	2%
	In-School Suspensions	6%	<1%	<1%	<1%	8%	14%	2%
	Out-of- School Suspensions	5%	<1%	<1%	<1%	9%	13%	2%
	Expulsions	6%	<1%	<1%	<1%	8%	13%	2%

Note: The image is sourced from the Student Discipline and School Climate in U.S. Public Schools report highlighting the findings of 2020-2021 Civil Rights Data from the Office of Civil Rights. (Civil Rights Data Collection, 2023)

Exclusion by Student Race and Ethnicity

Despite the reduction in suspension rates, some studies show Black and Latiné students are suspended at higher rates compared to White students (Lehmann, 2023; Losen & Martinez, 2020). The literature base has established the increased risk of suspension for Black students, but analysis on the disparities between Latiné and White students is mixed with some research suggesting Latiné students are suspended at a higher rate than White students and other studies suggesting there is no difference in suspension rates between White and Latiné students (Morgan & Wright, 2018). Furthermore, the authors found Asian students were less likely to be suspended compared to White students and other races. These findings were replicated by Nguyen and

colleagues (2019), though they highlighted the diversity within the Asian student category, often referred to as Asian and Pacific Islander, AAPI. Asian students are a diverse group with students from China, Japan, and Korea often being ascribed to the model minority stereotype but students from other parts of Asia such as the Hmong community were not included in that stereotype (Wu, 2014). The study found Asian students broadly were less likely to receive school discipline than White students, and Pacific Islander students were more likely to be disciplined (Nguyen et al., 2019). In addition, Latiné students may be reported as Latiné or Hispanic ethnicity but may present as White, Black, or another race, leading to potentially different school discipline experiences. Overall, the study reinforced the literature on school discipline showing higher risk of discipline for students from non-White backgrounds with Black and Latiné students being the most likely to be punished.

Exclusion by Student Gender

The research on exclusionary discipline by gender is clear that boys tend to be more likely to be suspended than girls. There is also evidence to suggest an interaction between gender and race with Black boys being more likely to be suspended than White boys, particularly for low level offenses (Gagnon et al., 2017; Smith et al., 2023). A corroborating study found not only are boys more likely to receive suspensions, but they are also more likely to receive suspensions in the future, potentially compounding differences in suspension rates by race and gender (Wilkerson & Afacan, 2022). Differences in disciplinary action may go beyond just male vs female differences as a meta-analysis of exclusionary discipline research found LGBTQ+ students were at-risk of disproportionate exclusionary discipline actions (Welsh & Little, 2018). Exclusionary discipline disparities put an already at-risk group of students at a further

disadvantage due to the negative impact of suspension on student academics and social well-being.

Exclusion by Student Disability Status

Students with disabilities, often quantified in studies as students with IEPs, are more likely to receive exclusionary discipline than their non-disabled peers (Cunningham et al., 2024; Smith et al., 2023; Welsh & Little, 2018). By one estimate, students with disabilities are 50 percent more likely to receive ISS than their peers, when controlling for race, gender, and previous referrals (Cunningham et al., 2024). Students with disabilities are also at an increased risk of receiving OSS (Krezmien et al., 2017) with one report suggesting students served by IDEA are twice as likely to receive OSS compared to non-disabled peers (Barton & Nishioka, 2014). In an analysis of exclusionary discipline for students with disabilities by Sullivan and colleagues (2014) they found suspension risk varied based on disability category. Students with EBD were significantly more likely to receive a suspension compared to other disability categories. The authors also recognized the interaction between race and disability status as Black students with disabilities were three times more likely to be suspended compared to other races (Sullivan et al., 2014). Although this difference may also be caused by the overrepresentation of Black boys labeled with the EBD disability category (Bal et al., 2019). Krezmien and colleagues (2017) examined suspension rate differences among students with autism and intellectual disability and found a decreasing trend in suspension across time, but an increased risk of suspension for Black students with intellectual disabilities compared to White students with the same label. They did not find any differences for students with autism. Combined, the data suggest there is an interaction between race and disability leading to an increased risk of suspension for some students.

The Impact of Exclusionary Discipline on Students

The negative effects of exclusionary discipline have led researchers to examine student outcomes associated with ISS, OSS, and expulsion, collectively referred to as exclusionary discipline. The research has often focused on OSS and the consequences for student academic performance, social relationships, and long-term school outcomes. The results of years of studies suggest suspension is not an effective method of shaping student behavior and does not lead to a reduction in challenging behaviors. Instead, research suggests suspension leads to negative academic and social outcomes for students receiving exclusionary discipline.

Academic Impact of Exclusion

Research has shown negative academic outcomes—including lower grades, risk of retention and course failure— associated with exclusionary discipline practices (Noltemeyer et al., 2015). In a meta-analysis by Noltemeyer and colleagues (2015) examining the impact of suspension on student academic outcomes, the authors found a significant decrease in academic performance linked to suspension. A more recent study confirmed the impact of suspensions extends beyond grades. The study found suspension negatively impacts student attendance, credits earned, graduation rate, and standardized test scores as well (Chu & Ready, 2018). The elevated risk of suspension by race, gender, and disability interacts with the negative impacts of suspension to exacerbate the achievement gap for at-risk students (Lacoe & Steinberg, 2019). Specifically, Black students, males, and students with IEPs are at increased risk of suspension leading to higher risk of academic failure and potentially a wider achievement gap.

The negative impacts of suspension and disparity by race, gender, and IEP status widen the achievement gap because if a group of students is at greater risk of suspension, they are more likely to feel the negative impacts of exclusion (Pearman et al., 2019). This is supported by

research as higher rates of ISS and OSS for students with IEPs is associated with decreased GPA and increased risk of high school dropout (Cholewa et al., 2018). A study examining the interaction between race and academic outcomes argued the mechanism of suspension leading to lower academic outcomes may be due to missed instructional time, missing assignments, and educator perception of the suspended students (Bell & Puckett, 2023). Additionally, suspension in the early grades puts students at greater risk for future suspensions, leading to a compounding impact of suspension on future academics (Wilkerson & Afacan, 2022). Furthermore, suspensions have the long-term risk of pushing students out of school and increasing the likelihood of student involvement in the juvenile justice systems (Leung-Gagné et al., 2022). In addition to negative effects for the excluded students, some research suggests exclusionary discipline may also negatively impact non-suspended students' academic performance by shaping school climate (Perry & Morris, 2014).

Social Impact of Suspension

Academics are not the only area where students are impacted as suspensions has a social emotional toll as well. Reduced engagement in class and lack of a feeling of belonging are also detrimental impacts of exclusion. For example, students from minoritized groups, in particular Black students, have increased chance of suspension, which leads to lower perception of school climate, less engagement in school, and reduced sense of belonging (Bottiani et al., 2017). Students with disabilities, who are more likely to receive exclusionary discipline, are also at increased risk for negative social impacts as researchers have demonstrated a link between higher exclusion rates and lower school engagement (Cunningham et al., 2024). Fortunately, some research suggests developing more welcoming school environments may be a protective

factor to reduce the disparities between Black and White students in terms of suspension rates (Fisher et al., 2020).

With the abundance of evidence for negative effects of suspension it begs the question: Does suspension work and is it worth it? The evidence suggest suspension is ineffective at reducing challenging behaviors (Brobby, 2017). The detrimental impacts of suspension and the lack of efficacy in shaping student behavior have provided justification for the implementation of bans on suspensions as a form of behavior management and the search for restorative and proactive behavior management strategies.

Moving Away from Exclusion

The failure of suspension to reduce challenging behaviors and the detrimental impact on students has led to calls from researchers, policy makers, and practitioners to seek alternatives to exclusionary discipline. In an attempt to shift discipline practices away from suspension, legislators and policy makers have worked to put limitations on the use of suspension. Meanwhile, researchers and practitioners have looked for alternatives to exclusion that have the potential to address the underlying causes of challenging behaviors and reduce the future likelihood of such behaviors from occurring. In addition, schools have sought systematic school-wide approaches, like PBIS, to streamline behavior management and aid in preparing staff to address challenging behaviors without excluding students from the instructional environment.

The Rise of Suspension Bans

In response to the detrimental impacts of suspension and the disproportionality in its usage, suspension bans gained popularity. The increase in suspension bans has contributed to the reduction of OSS but has not eliminated the disproportionality by race and disability status in exclusionary discipline. In 2017, Maryland introduced a ban on suspensions in specific grades

for behaviors that did not pose an immediate threat to students or teachers. Researchers examining the impact of the ban found a reduction in the use of OSS for all students—without replacement of ISS—suggesting a reduction in exclusionary discipline overall (Lincove et al., 2024). However, the ban did not eliminate the disproportionate use of suspension by race and disability status, nor did it have effects for non-targeted grades as Black students continued to be suspended at a higher rate than white students. This trend has been repeated in other studies examining suspension bans, such as a policy in Massachusetts that effectively eliminated the use of OSS in K-3 settings and one examining schools in New York City, both leading to a reduction in overall suspension (Cleveland, 2023; Craig & Martin, 2023). Despite the continued disproportionality, suspension bans may be a useful strategy to lower OSS rates, in fact, 2021 Office of Civil Rights data show OSS rates are currently at their lowest level in 50 years (National Center for Education Statistics, 2024). With a reduction in suspensions, students are at reduced risk of the negative impacts associated with exclusionary discipline. Despite the persistence of disproportionate suspension rates, a reduction in rates can be observed for all groups.

The Impact of Alternatives to Exclusionary Discipline

The reduction in exclusionary discipline through suspension bans and alternative discipline strategies have led to improvements in behavior management. Alternatives to exclusionary discipline such as PBIS and restorative practices have gained traction as practitioners seek alternative forms of behavior management that prioritize positive approaches to school discipline leading to the development of student skills and the maintenance of instructional opportunities for student learning (Sugai & Horner, 2020). Studies have examined the effectiveness of PBIS on exclusionary discipline and found strong evidence that suggest a

reduction in the use of OSS and ISS (Fleming et al., 2024; McIntosh et al., 2021b). From this research, the importance of high-fidelity implementation of PBIS to lower the rate of exclusionary discipline has been highlighted, yet there is not a strong link between PBIS implementation and improved academic performance (James et al., 2019; Kim et al., 2018). In addition, some research suggests PBIS may not always reduce disproportionality in discipline despite reducing suspensions (Zakszeski et al., 2021). This suggests PBIS alone may reduce suspension rates but can still result in lost instructional time, limiting the impact of the model on academic performance. To improve academic outcomes for students in PBIS schools, a framework that intertwines reduced suspension with decreased time out of the classroom and opportunities for student skill building is needed.

Restorative practices, a model for alternatives to suspension that provide students with the opportunity to learn from their challenging behavior and repair harm, has shown the ability to reduce suspension rates and potentially reduce disproportionate discipline (Gregory et al., 2018; Huang et al., 2023; Schiff, 2018). The Inclusive Skill-building Learning Approach (ISLA) is one example of a restorative and instructional intervention that has shown promise for reducing exclusionary discipline and disproportionality. The ISLA model supports schools in developing school wide behavior expectations, training teachers on proactive classroom behavior management, and supporting the implementation of alternative to exclusion that teaches students skills to reduce the future likelihood of challenging behavior. ISLA focuses on reducing the amount of lost instructional time with the goal of ensuring students have high quality academic instruction (R.N.T. Nese et al., 2020, 2022). One study found the ISLA intervention led to a reduction in lost instructional time and importantly, a reduced risk of exclusion for Black students and for males (Pimentel-Mannan et al., 2024). In order to address behavior at school,

ISLA and other restorative frameworks focus not only on major behavior referrals, but also on supporting teachers to address minor behaviors because of their contribution to lost instructional time.

Minor Behavior Referrals

The reduction in suspension rates mask the impact of other forms of exclusionary discipline that result from minor referrals. Out-of-school suspension (OSS) has been well-documented to decrease a student's sense of belonging at school and impair academic learning (Cribb Fabersunne et al., 2023), but the impact of minor referrals is less studied. Given the reduction in OSS, it is time for researchers to more closely examine the impact of other exclusionary discipline strategies such as time-outs in the office, reset rooms, and detention.

Minor behavior referrals are designed to document low level, unwanted behaviors in the classroom and allow teachers to request support in managing those behaviors. Unlike ODRs, which typically have a longer administrative process that may lead to severe consequences like suspension, minor referrals often result in immediate responses such as a mini-conference with the student, or additional exclusion such as time in the reset room, timeout, or detention. These faster forms of exclusion occur on a shorter timeline—from behavior to referral to consequence—possibly within the same class period. Although less severe than out-of-school suspension, these responses still represent forms of exclusion that lead to lost instructional time. These less severe forms of exclusion highlight the continuum of exclusionary discipline with less severe consequences that still result in the removal of students from the instructional environment. Understanding this rapid sequence from behavior to exclusion is critical because even brief removals can accumulate a lot of missed instruction, especially for students already

struggling academically. When disproportionality exists in behavior discipline, these rapid exclusions can exacerbate opportunity gaps between student groups.

In addition to understanding the discipline responses to minor referrals, research also needs to explore the factors that can mitigate harmful discipline practices. PBIS is a proactive prevention framework that supports students and teachers in creating school environments that are consistent, predictable, positive, and safe. The use of PBIS has been shown to reduce the use of ODRs and suspensions (Bradshaw et al., 2010; McIntosh et al., 2021a,b), but the patterns of consequences for minor referrals is unknown. It is possible that PBIS strategies reduce suspensions without reducing non-suspension behavior referrals or eliminating racial and disability status disproportionality (Gallagher et al., 2019; Okonofua et al., 2016; Vera et al., 2018). Understanding the ability of PBIS to reduce referrals is important in reducing exclusion and increasing instructional time for students with lagging academic skills by keeping them in the classroom.

Students with lagging academic skills have been shown to be more likely to engage in disruptive behaviors leading to exclusion but are also the most in need of academic instruction and would benefit the most from increased instructional time (Gregory et al., 2010). Minoritized students by race tend to have lower academic scores due to an opportunity gap, potentially leading to more challenging behaviors (Gregory et al., 2010). The most common types of behavior referrals written are for disruption, defiance, and disrespect, which are often motivated by task avoidance or to gain attention. Students who engage in behaviors to avoid tasks may do so because of a lagging academic skill. By boosting academic skills, students may be less likely to engage in disruptive behaviors and receive referrals. Social interaction and connection are another common motivator for students, which may lead to classroom disruptions and referrals

(Meadan et al., 2016). Given that neither academic skills nor social needs are built or supported through exclusionary discipline, what discipline practices are used in response to challenging behavior, and can we find strategies that build skills to prevent future challenging behaviors?

Limited Research on Minor Behavior Referrals

Minor referrals have not received the same attention in the literature, which primarily focuses on the impact of suspensions, likely due to the ease of measurement and the severity of the punishment. While OSS and expulsion are the most severe consequences of referrals, they are less common compared to other actions taken. More common consequences include detention, conferences with behavior staff, parent contact, community service, loss of privileges, time in the office, or Saturday school. While not all consequences are punitive, some consequences such as detention, time in the office, and ISS result in lost instructional time and exclusion from the learning environment. In addition to severity, these actions range in time out of class, lost instructional time, and skill building opportunity as well as frequency. The breadth and severity of potential responses may also lead to disproportionality in consequences based on a student's identity.

The Impact of Teacher Implicit Bias in Minor Referrals

Bias of educators has been shown to impact a teacher's perception of a student's behavior (Gregory & Roberts, 2017; Okonofua & Eberhardt, 2015). Girvan and colleagues (2017) conducted a study to determine the source of racial disproportionality in behavior referrals and found subjective referrals explained the majority of the disproportionality present (Girvan et al., 2017). The results suggest that teacher implicit bias is contributing to the disproportionality of ODRs. These implicit biases negatively impact minoritized students. An additional layer arises when teachers and students come from a different racial background. One study suggested that

Black teachers, while providing more severe punishments overall, were more likely to hold Black students to higher standards than White teachers (Samuels et al., 2016). In a study examining future expectations for students, the authors found teachers who held lower expectations were associated with higher ODR rates for Black students, but as teacher expectations increased, the racial disparity gap decreased (Santiago-Rosario et al., 2021). The results suggest that increasing teacher expectations may be an effective strategy to reduce referral disparities.

In addition to navigating the subjective nature of behavior referrals—where students from minoritized backgrounds are referred at higher rates than White, heteronormative students—the consequences and coding of referrals can be subjective and lead to difficulty interpreting data. For example, a referral can result in a conference, a detention, a parent phone call, or other responses varying in severity. Two students may have each received a behavior referral for “disruption”, but one student has a minoritized racial or ethnic identity with differing severity of responses. This may lead to identical ODR rates, but differences in the severity of consequences and potentially differences in the amount of lost instructional time. Another confounding variable is the potential for subjective coding of the challenging behavior. Teachers may code behaviors differently from each other or a teacher may code the same observed behavior differently for students of differing racial identities. These discrepancies can complicate the study of referrals based on the listed challenging behavior. Furthermore, an ODR captures a moment in time, yet when administrators determine consequences for ODRs they may consider past actions taken or previous student behavior. Therefore, a referral that appears identical in behavior may lead to differences in consequences due to the student’s history of challenging behavior. Teacher input on a referral similarly considers the teacher’s interactions with the

student prior to a particular referral and the classroom strategies previously attempted to manage the behavior. Those events that precede the referral are not captured in the data. Unfortunately, without direct classroom observations of student behavior, we cannot know the extent of the discrepancies in teacher coding of student behavior referrals. This limitation of the data presents a challenge to understanding the disproportionate nature of minor and subjective behavior referral.

Purpose of Study

It has been well documented in the literature base, exclusionary discipline negatively impacts students by disrupting their instructional time, increasing their risk of academic failure, and perpetuating cycles of punitive responses. This study will descriptively examine the types of student behaviors that result in minor behavior referrals and the subsequent disciplinary actions taken as well as differences in rates of referral between major and minor referrals by student characteristics. Furthermore, data on minor referrals and exclusion will be disaggregated by race, ethnicity, gender, and disability status, to assess disproportionality in minor referral outcomes. The study will examine differences in exclusionary discipline by student characteristics between major and minor referrals to ascertain if referral severity differs based on student characteristics in order to determine if certain student groups are overrepresented in the receipt of major referrals. Finally, analysis will examine the moderating effect of PBIS implementation on exclusionary discipline. By analyzing the patterns and consequences for minor behavior referrals, this study aims to highlight the rapid exclusionary responses that can result from minor behaviors and remove students from the instructional environment. The study also aims to demonstrate the need for equitable and instructional responses to student behaviors in lieu of punitive, exclusionary discipline.

This study expands upon the current knowledge base by examining minor referrals to understand the impact on students from different groups and the presence of disproportionate use of exclusionary discipline in minor referrals. Literature on the impact of exclusionary discipline on students has predominantly focused on suspensions and ODRs as the unit of measurement, yet the majority of referrals do not result in suspensions, leaving a large portion of data unexamined. Although the negative impact of minor referrals on educational outcomes may be less severe than suspensions, minor referrals are used more frequently in schools, resulting in increases in missed instruction and limiting opportunities for social development over time. Additionally, this study expands on the understanding of disproportionality in discipline practices by examining differences based on student race, ethnicity, gender, and disability status. To achieve this, the follow research questions will be examined:

Research Questions

1. What types of student behaviors result in minor behavior referrals in middle schools?

Based on historical context and exclusionary discipline data examining suspensions, I hypothesize minor referrals will be most commonly written for non-violent behaviors, but that a difference in rates of exclusion exists for non-violent behaviors based on student characteristics (RQ1).

2. What actions are most frequently taken in response to minor behavior referrals?

Additionally, I hypothesize that minor referrals are less likely to lead to exclusionary discipline compared to major referrals with conferences with students and phone calls home being the most common response. Detention and time in the office are likely to be the most common forms of exclusionary discipline for minor referrals (RQ2).

3. How does classification of referrals into major and minor infractions differ by race, ethnicity, gender, and IEP status for challenging behaviors?

When examining differences in classification of referrals into major and minor based on similar behaviors, I expect students of color to be more likely to receive a major referral compared to White students though I do not predict differences in referral categorization by student gender or IEP status (RQ3).

4. Are there differences in rates of exclusionary discipline for minor behavior referrals by student race, ethnicity, gender, and special education status?

I hypothesize students who identify as Black receive exclusionary discipline at a higher rate than other students for the same behavior. Furthermore, I expect students with disabilities are excluded at a lower rate due to protections afforded to them by their IEPs. I believe males will receive higher rates of exclusionary discipline as it is consistent with suspension data demonstrating their likelihood to be removed from class at higher rates than female students for the same behaviors (RQ4).

5. Does implementation of Positive Behavioral Interventions and Supports (PBIS) moderate the rates of exclusionary discipline?

Finally, I hypothesize schools with higher fidelity of Tier 1 PBIS implementation moderate the impact of student characteristics leading to a diminished effect of student minority status on likelihood of exclusion. I expect to see a similar pattern with Tier 2 fidelity associated with a reduced exclusionary discipline rate (RQ5).

Chapter II: Methodology

Settings

The study examined student behavior referrals from middle schools in four states in order to answer the research questions on disproportionality and patterns in minor referrals. The primary source of data was obtained through the National Technical Assistance Center on PBIS located in the Educational and Community Supports (ECS) office at the University of Oregon. The data is comprised of schools that use the School-Wide Information System, SWIS (May et al., 2005) for tracking student behavior referral data and contains only students who received at least one behavior referral during the two years of behavior data obtained. All schools using SWIS are required to implement PBIS, therefore all schools in the sample are considered PBIS schools. To assess the fidelity of PBIS implementation the dataset contains a total fidelity score by PBIS tier. To examine the moderating effect of PBIS on disproportionality, school implementation was examined at the Tier 1 and Tier 2 level, which corresponds to universal prevention practices in Tier 1 and interventions targeted to support students struggling with behavior challenges at Tier 2. Although PBIS contains 3 tiers, the Tier 3 data was not used in the analysis due to low number of responses and the focus of Tier 1 and Tier 2 on low level behaviors. Previous research has grouped schools into high PBIS and low PBIS schools based on the fidelity of implementation scores—typically around 70-80% implementation fidelity cut score—but this analysis used centered non-zero numeric fidelity scores on a continuous scale with the goal of examining at what level of fidelity is a change in exclusion observed.

Sample Selection

The data consists of behavior referrals from the 2022-2023 and 2023-2024 school years in 4 states: Minnesota, Massachusetts, Georgia, and Oregon. The years were selected to capture

the patterns in behavior discipline from the post-COVID return to in-person instruction; and the states were selected to represent different geographic regions in the United States. Specifically, Minnesota was picked to represent the Midwest, Oregon the West Coast, Massachusetts New England, and Georgia the South. In addition, state selection was influenced by the researcher's background as a student at the University of Oregon and a former Minneapolis Public Schools teacher and behavior specialist in the district.

The sample was limited to middle schools, due to evidence suggesting higher referral rates in middle school compared to elementary and high school as well as the researcher's experience as a middle school teacher. In general, middle school is defined as those serving 6th through 8th grade students, though a handful of schools in the sample included either 5th or 9th grade students. Only students who received a behavior referral were available therefore the data did not represent the full school population. As a result, population level risk of exclusion was not available for analysis. All referral data was de-identified from ECS prior to providing the researcher access. For analysis, the schools, students, referrals, and districts were assigned an identification number by ECS. Prior to examining the data, the Institutional Review Board (IRB) determined the study did not qualify as human subjects' research as the data was de-identified allowing it to be exempt from IRB.

Sample Preparation

The dataset provided by ECS contained variables for student and school level characteristics, in addition to the referral level details. For student level characteristics, the data consisted of student race and ethnicity, two separate columns that were transformed into one combined column, IEP status, grade, and gender. One limitation of the data was the lack of a student level socio-economic status (SES) variable such as Free and Reduced-Price Lunch

(FRPL), which limited the ability to control for SES in analysis of disproportionality. Because the study was focused on examining disproportionality based on student characteristics, referral data that did not include student race/ethnicity, gender, and disability status were excluded from the dataset using model determined list-wise deletion.

School level characteristics included state, total enrollment, total minority enrollment, enrollment by grade, urban location, enrollment by gender, enrollment by free or reduced-price lunch status, and PBIS Tier 1 fidelity scores. Although Free and Reduced-Price Lunch (FRPL) enrollment was present at the school level as a control variable it was not present at the individual student level for direct analysis of the differences in exclusion based on student FRLP status. Additionally, Massachusetts schools did not report FRPL enrollment data, leading to the variable being dropped in school wide analysis as well. The school level characteristic data were collected from the National Center for Education Statistics (NCES) data. The enrollment information was merged with the schools' discipline data to create school level variables by ECS. Some school level variables were calculated from the existing data such as percent free and reduced-price lunch (FRPL) enrollment, percent enrollment by minority group, and percent enrollment by gender, which were controls for differences in school population demographics when examining exclusionary discipline rates. The final model used Black student enrollment at the school level as a control due to the over-representation of Black students in ODR research. Other school level characteristics such as urbanicity, which indicate school locale type: urban, suburb, rural, or town were collapsed into larger groups due to low observations. In this study town and rural were combined to create a variable with three categories: city, suburb, and rural.

The school variables also contained pre-calculated total ODRs as well as total referrals by gender, number of students receiving an ODR, and rate of ODR per student. A summation of the

total ODR variable suggested a total of 266,264 referrals for 56,280 students with an average rate of 2.08 referrals per student. The difference between the estimates based on these variables and the data descriptives suggested the dataset may overestimate the total number of referrals written and underestimate students at the schools. Due to the discrepancy, the data analysis excluded the pre-calculated ODR level variables, instead conducting analysis on the referrals present in the dataset. Total enrollment by student characteristic group such as race, IEP status, and gender were not included in the multi-level model but were used for calculating RDR and RR.

Participants

The sample contained a total of 271,979 behavior referrals across the 2022-2023 and 2023-2024 school year with 161,115 (59%) coded major behavior referrals and 110,864 (41%) labeled minor behavior referrals. The unfiltered data came from 49,203 individual students at 159 unique schools represented across the four states. Georgia contributed 85 schools, 53% of the full sample. Male students composed 68% of the total referrals in the dataset with 32% female students and less than 0.05% of the referrals for students listed as other gender. The ratio of male to female students remained consistent across major and minor referrals and states. Other gender students were dropped from the analytic sample due to low sample size. Students with disabilities—as determined through IEP status—composed 25% of the behavior referrals and an average of 17% of the total students in the sample, though that percentage varied by school and state. In Massachusetts students with IEPs were a higher than average, 20% of the students.

Grade level was distributed evenly across middle school grades with an increase in referrals from grade 6 (28%) to grade 7 (36%) and a slight decrease in grade 8 (34.5%). Student race and ethnicity proportions of referrals were as follows: Asian (1%), Black (29%), Latiné (24.5%), multi-racial (6%), Native (<1%), Pacific Islander (2%), White (37%). However, again

there were large differences in racial composition by state where Georgia had a higher percentage of Black students than other states. Major referrals for Black students (21%) constituted a larger proportion of total referrals compared to minor referrals (7.5%), while other race groups were relatively evenly split between major and minor referrals, suggesting a higher proportion of major referrals for Black students.

Student level demographics raw counts were as follows; Asian (798), Black (14807), Latino (7661), multi-racial (2254), Native (234), Pacific Islander (1237), and White (14968), students with IEPs (7672), males (26250), females (15690), and other (19) with only students who received at least one behavior referral present in the dataset. A full demographic breakdown can be found in Table 1.

In the analytic sample for the multi-level model students above the 99th percentile for minor referrals received in a given school year were excluded as they were deemed to be high referral outliers, this resulted in a cut off of 29 referrals for 2022-23 and 31 referrals for 2023-24 and a total of 287 students excluded though they contributed about 14% of the minor referral observations. Roughly half of the high referral students came from two schools, one in Massachusetts and the other in Minnesota. An examination of the high referral student demographics showed 42.9% were Latiné and 35.5% were White, Black and Other race made up a smaller portion, 10% and 8.4% respectively. The high referrals were also more likely to be males, 75.6% of the students, versus 66% of the referral sample were males. Students with IEPs made up 35.5% of high referrals students despite averaging 17% of the referral sample.

In addition, 5 schools with fewer than 20 referrals, a total enrollment of less than 10 students or fewer than 10 unique students referred were dropped to improve model stability. Schools with 100 percent exclusionary discipline as well as those with zero exclusion were also

dropped from the analytic sample to improve model estimates. In total, 26 schools were dropped due to these criteria, 23 were full exclusion and 3 were no exclusion. Of those, 22 were from Georgia and one from Oregon. Finally, observations with missing exclusionary discipline data were filtered out. The filtered sample of minor referrals consisted of 100 schools, 18,977 students, and 74,059 referrals. This subsample of only minor referrals was used to analyze RQ4.

Table 1
Sample demographic characteristics by referral, student, and school

	Referrals <i>n</i> (%)	Major Referrals <i>n</i> (%)	Minor Referrals <i>n</i> (%)	Students <i>n</i> (%)
Race / Ethnicity				
Asian	2698 (0.99)	1677 (1.04)	1021 (0.92)	869 (1.76)
Black	76429 (28.10)	56396 (35.00)	20033 (18.07)	16323 (33.09)
White	91151 (33.51)	47999 (29.79)	43152 (38.92)	15717 (31.86)
Latíné	65020 (23.91)	32995 (20.48)	32025 (28.89)	9115 (18.48)
Native American	2122 (0.78)	1104 (0.69)	1018 (0.92)	253 (0.51)
Pacific Islander	5584 (2.05)	3239 (2.01)	2345 (2.12)	1519 (3.08)
Multi-racial	16165 (5.94)	10254 (6.36)	5911 (5.33)	2424 (4.91)
Missing Race	12810 (4.71)	7451 (4.62)	5359 (4.83)	3110 (6.30)
IEP Status				
IEP	67119 (24.68)	41087 (25.50)	26032 (23.48)	8756 (17.65)
Non-IEP	204860 (75.32)	120028 (74.50)	84832 (76.52)	40860 (82.35)
Gender				
Male	184198 (67.73)	106203 (65.92)	77995 (70.35)	30648 (62.27)
Female	87671 (32.23)	54823 (34.03)	32848 (29.63)	18545 (37.68)
Other	110 (0.04)	89 (0.06)	21 (0.02)	22 (0.04)
Grade				
Grade 5	2692 (0.99)	1522 (0.94)	1170 (1.06)	481 (0.83)
Grade 6	76265 (28.04)	47981 (29.78)	28284 (25.51)	16656 (28.78)
Grade 7	98453 (36.20)	57183 (35.49)	41270 (37.23)	20314 (35.10)
Grade 8	93487 (34.37)	53461 (33.18)	40026 (36.10)	20026 (34.60)
Grade 9	1008 (0.37)	961 (0.60)	47 (0.04)	378 (0.65)

Note: The descriptive statistics are from unfiltered data and does not represent the final analytic sample. Student race and ethnicity data was subsequently collapsed into Black, White, Latíné, and Other with the other group consisting of Asian, Pacific Islander, Multi-racial, and Native students. Additionally, students identifying as other gender were dropped in the final sample.

When examining student composition by state, Georgia had about 51% of their student referrals for Black students, 18% for White students, and 11% for Latiné and Other students. By contrast Massachusetts had only 4.5% of minor referrals going to Black students and 57% for White students, 27% of referrals for Latiné students, and 4.5% of the other race group. Oregon followed a similar pattern to Massachusetts: 52% of minor referrals for White students, 2% for Black students, 34% for Latiné students, and 9% for other race students. Minnesota fell in between with more evenly split racial demographics, 38% of minor referrals for White students, 15% Black, 20% Latiné, and 12% other race.

Importantly, the analytic minor referral sample of students from Minnesota was much smaller than other states with only 1,830 unique students in 10 schools from the Minnesota minor referral sample. Oregon comprises the majority of students in the sample with 8,575 students across 41 schools, followed by Georgia with 5,782 in 34 schools, and Massachusetts contributed 2,790 students from 15 schools. Because of the differences in state composition of racial groups a dichotomous state variable with observations from Georgia compared to students from other states. In the pre-filtered minor referral sample prior to removing 100 percent exclusion schools, Georgia had an exclusion rate of 59% and Oregon 49% compared to 19% and 17% for Massachusetts and Minnesota respectively, but after filtering Georgia's exclusion rate dropped to 42% exclusionary discipline while other states remained the same.

PBIS Subsample for RQ5

The study examined the relationship between PBIS implementation and exclusionary administrative actions taken (RQ5), which required a separate restricted dataset that included only schools with PBIS implementation fidelity scores. The Tiered Fidelity Inventory (TFI) data was provided from ECS and matched to schools using the assigned school identification number.

Tier 1 and Tier 2 implementation fidelity were selected for analysis as it was theorized Tier 1 and Tier 2 interventions would have the greatest impact on low level student behaviors. Tier 3 PBIS interventions are often reserved for students with the highest behavioral needs therefore they may be more likely to be outliers in the number of minor behavior referrals received or receive more major referrals. The PBIS data was split into schools with Tier 1 fidelity scores and those with Tier 2 fidelity scores in order to run a multi-level model with the largest available sample for the fidelity variable. The Tier 1 model included a total of 52 schools after removing full exclusion, high referral outlier students, low-enrollment and low referral schools as well as those with missing or zero scores for PBIS fidelity. In contrast, the Tier 2 data sample contained 33 schools as a result of an increase in zero scores indicating a lack of Tier 2 evaluation. Tier 1 scores ranged from 0-30 with a max of 30 points representing complete implementation of the PBIS intervention. Tier 2 scores had a range of 0-26, though for both tiers zeros were removed as they indicated a lack of implementation. The mean score for Tier 1 schools was 23.74 and the mean for Tier 2 was 16. Tier 2 also had a greater spread in the distribution of scores while Tier 1 was skewed toward the higher range of fidelity scores. In the scoring guide for the TFI, total scores at or above 70% indicated a high level of fidelity, though for this analysis the scores were examined using a centered continuous scale.

Again, the sample of students cannot be considered representative of the total enrollment in schools in these states as only students who received at least one referral were contained in the dataset and only schools implementing PBIS were included in SWIS data, meaning students without behavior referrals and those from schools not using SWIS were absent from the analysis. As a result, Risk Ratio (RR) and Raw Differential Representation (RDR) represent disproportionality for students who have received a referral. In other words, what is the risk of

major referral or exclusion for a student receiving a referral based on student characteristics and does that risk vary based on a school's level of PBIS implementation?

Key Variable Processing

To address the research questions, this study implemented a multi-site, multi-year observational approach using administrative behavior referral data from middle schools across four states: Georgia, Oregon, Minnesota, and Massachusetts. The unit of analysis was the referral with student and school level covariates. The study aggregated data across two academic years to increase sample size—though limited the years to post-COVID and the return to in-person instruction—to examine disciplinary actions, exclusionary versus non-exclusionary by student demographic characteristic (race/ethnicity, gender, IEP status). Prior to analysis, the data underwent processing steps to ensure variables such as challenging behavior, disciplinary response, student race, PBIS fidelity, and exclusionary discipline were properly coded.

Student Race and Ethnicity

As part of data processing, a new race and ethnicity variable was created, combining the categorical race and ethnicity columns into one race/ethnicity variable to examine Black and Latiné students within the same variable. The new column was created using the original race data and binary Hispanic ethnicity column to check if Black Latiné students differed from White Latiné students in terms of disproportionality in the response to behavior referrals, as Latiné students are typically clustered together for analysis despite students potentially facing different forms of discrimination based on the race they present. However, due to a limited sample size the primary model included just a consolidated race and ethnicity variable with students categorized as Black, Latiné, White, or Other. The other race category comprised of students listed as Asian, Native American, Hawaiian/Pacific Islander, or Multi-racial.

Challenging Behavior Variables

The primary focus of the study was on disproportionality in referral outcomes, but descriptive analysis also examined the types of challenging behaviors listed in the referrals as well as patterns of challenging behaviors between minor and major referrals. According to the data dictionary for the SWIS referral data there are 42 options for challenging behaviors. To conduct the analysis, the challenging behavior columns were aggregated into a single column with a list of the challenging behaviors associated with each referral. The counts of challenging behaviors were counted for descriptive analysis. Importantly, the data dictionary listed some ambiguous options for challenging behaviors such as “minor- unknown” and “minor – other”, which were included in the counts despite the unclear definition.

Defining Exclusionary Discipline

Initially, behavior referrals were labeled with a primary behavior and options for up to four secondary behaviors. These were combined into a single variable for challenging behavior, as described above, that contained a list of all behaviors on the referral. Similarly, administrative actions taken in response to a referral were originally listed as a primary action with additional optional actions. The actions taken columns were transformed into a variable with the list of administrative actions taken in response to the referral. Subsequently, the cleaned actions taken variable was used to create an indicator variable for exclusionary discipline, which was the primary dependent variable. If any of the consequences attached to the referral were considered exclusionary, defined as removal from academic instructional time, the indicator variable was set to 1 to indicate exclusion. Otherwise, if the action taken was not considered exclusionary then the referral was deemed non-exclusionary. Exclusionary responses were comprised of the following disciplinary actions: time in the office, detention, in-school suspension, out-of-school

suspension, expulsion. Non-exclusionary responses consisted of loss of privileges, conference with the student, parent contact, individualized instruction, Saturday school, restitution, restorative justice, and community service. A subset of administrative actions was deemed to be neither exclusion nor non-exclusion and were removed from analysis. The excluded administrative responses were other administrative action, unknown administrative action, bus suspension, action pending, additional support requested, and alternative placement.

PBIS Fidelity Scores

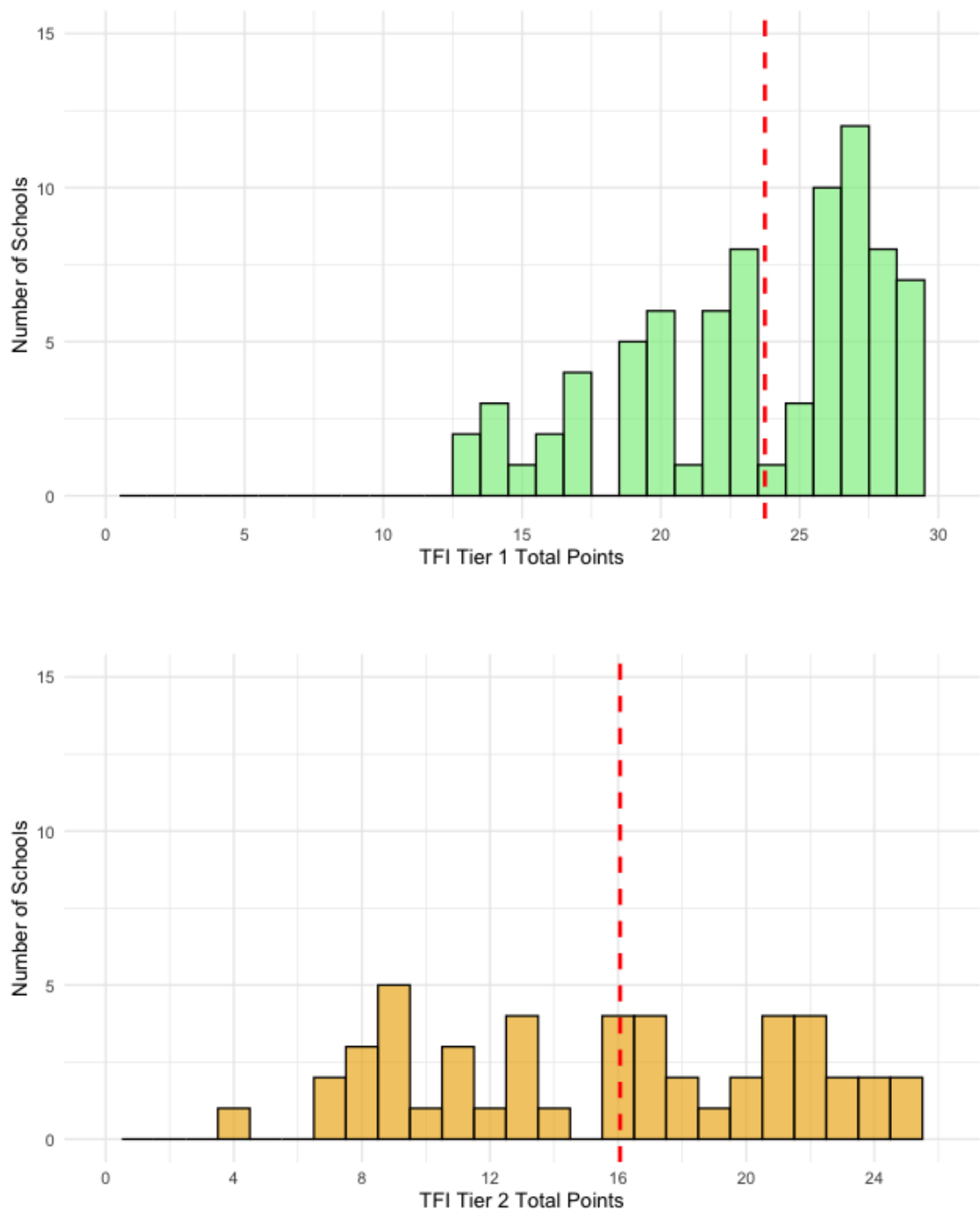
PBIS fidelity scores were used at Tier 1 and Tier 2 with scores ranging from 0 to 30 for Tier 1 and from 0 to 24 for Tier 2, however several schools were either missing fidelity scores entirely or had a score of zero. Sensitivity checks—including chi-squared and regression analyses— were conducted to understand the validity of zero scores. Schools with fidelity scores of zero were excluded after the checks indicated that zero scores resembled missing data rather than valid low-fidelity scores. Furthermore, missing data patterns indicated that PBIS fidelity was not reported consistently across years. In response, valid Tier 1 scores from 2022–2023 were carried forward to 2023–2024 if the latter year was missing, under the assumption that fidelity practices continued even if not formally reported. Sensitivity checks also examined potential selection effects. Students attending schools without PBIS fidelity data received, on average, one more referral than their peers in schools with PBIS scores.

A Little’s MCAR test indicated that data were not missing completely at random, likely due to patterns clustered by state and school reporting practices. After carrying forward the PBIS fidelity scores, missing PBIS fidelity data was examined again and found to contain a large amount of missing PBIS data for Minnesota, over 90%, Oregon about 60%, and Massachusetts, 40% at the referral level. Georgia had the most complete data on fidelity scores with only 2% of

observations missing PBIS fidelity scores. As a result, Georgia contributed the majority of the analytic sample for PBIS moderation analyses: 31 of the 52 schools (60%) in the Tier 1 model and 21 of 33 schools (64%) in the Tier 2 model were from Georgia. This translated to approximately 55.9% of students with Tier 1 data and 57% of students with Tier 2 data originating from Georgia. A summary of school-level distribution by year of PBIS fidelity scores for Tier 1 and Tier 2 aggregated across states can be found in Figure 2 and the distribution of PBIS fidelity data availability disaggregated by state is presented in Figure 3.

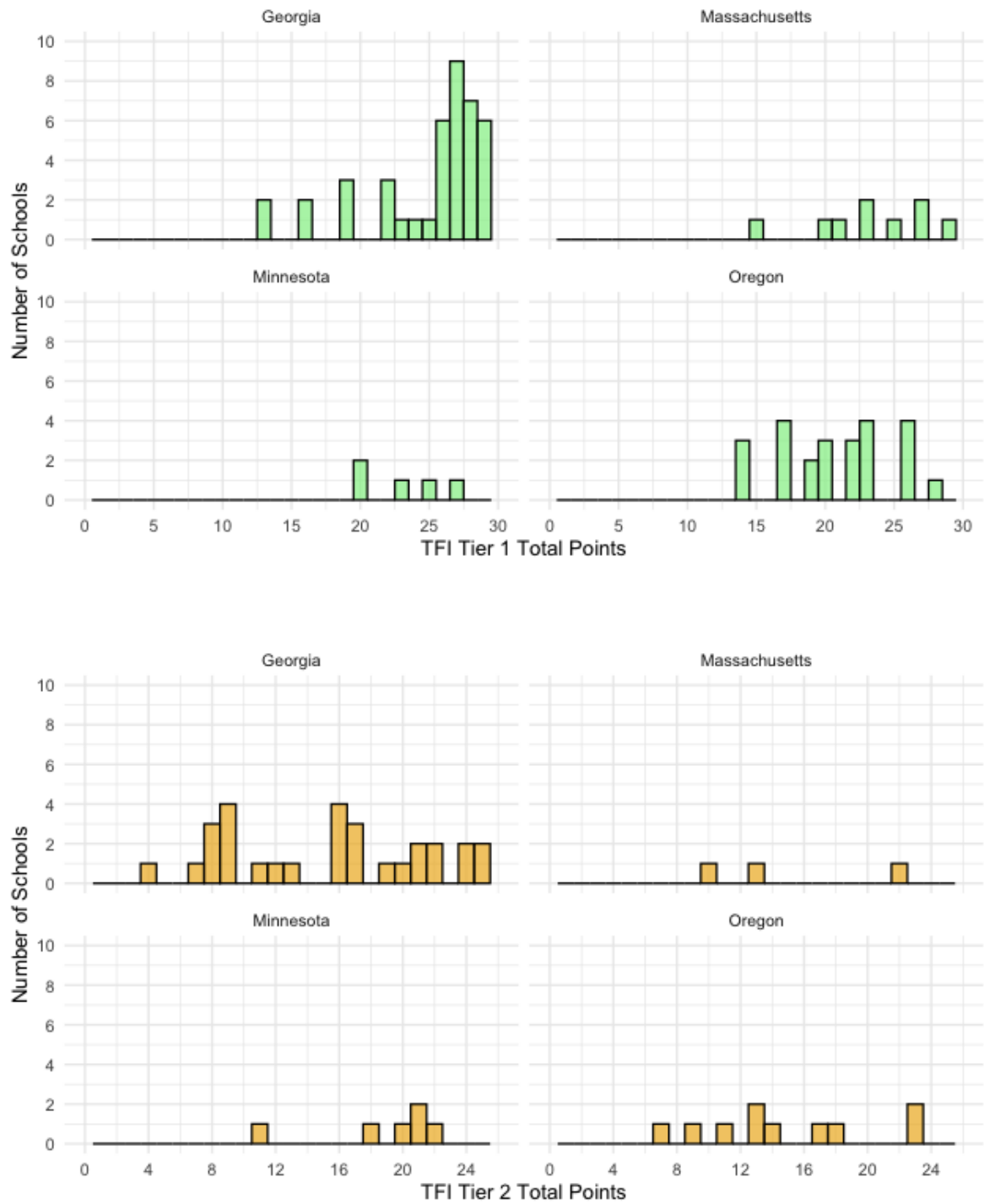
To further understand the data, checks were conducted to compare exclusion rates across schools with valid TFI scores, schools with a TFI score of zero, and schools missing TFI data in the pre-filtered sample. Zero scores were rare for Tier 1 ($n = 4$) but more common for Tier 2 ($n = 38$), suggesting Tier 2 implementation was less established or less consistently measured. State-level patterns in missingness were consistent across tiers, with Massachusetts and Oregon missing Tier 1 data for about 65% of schools, compared to roughly 50% in Minnesota and just 10% in Georgia. However, when measured by referral-level observations, these missing data represented only 2% of referrals in Georgia, compared to 25% in Massachusetts, nearly 50% in Oregon, and approximately 97% in Minnesota. Chi-squared tests revealed statistically significant differences in exclusion rates across TFI score categories (valid, zero, missing) for both Tier 1 and Tier 2, with zero-score schools showing patterns more similar to schools with missing data than those with valid fidelity scores. These results further justified the analytic decision to exclude schools with zero PBIS fidelity scores from moderation models.

Figure 2: School-level Distribution of Tier 1 and Tier 2 Fidelity scores



Note: The range of scores for Tier 1 fidelity is between 0-30 and the range for Tier 2 fidelity is 0-26. Fewer schools reported Tier 2 and the average score was lower as well. School-level scores by year are shown above but the multi-level model uses referral level scores, so some schools are given a heavier weight in the model.

Figure 3: School-level Distribution of PBIS schools by State and Tier



Note: Tier 2 had fewer schools reporting scores and the majority of schools that reported scores were from Georgia. The x-axis differs due to different ranges of scores for PBIS fidelity by tier.

Missing Data

As previously described, the student characteristic data was processed to initially create two student race and ethnicity variables, one that allowed for the combination of race and Latiné status and another that merged the race and ethnicity columns. In the dataset, if a student was listed as Latiné using the binary variable, “Is Hispanic?”, then their race was combined with Latiné to assign a label such as Black Latiné or White Latiné. This variable was used to check for disproportionality in race and ethnicity. Ultimately, the variable using a combination of race and ethnicity was not used in analysis due to the small group sizes which inhibited the analytic model from converging on statistical estimates. The primary race and ethnicity variable in the multi-level model was created by collapsing groups into four categories, Black, White, Latiné, and Other. The race and ethnicity variable, as well as other key variables on gender, IEP status, and exclusionary discipline were then examined for missing data patterns.

An analysis of missing data at the referral level revealed that in the full dataset, 4.7% of referrals were missing student race information, 4.5% were missing urbanicity data, 2.4% were missing school Black enrollment data, and student gender and IEP status were complete. At the school level, urbanicity was missing for 2.6% of referrals, percent Black enrollment was missing for 1.6%, and percent FRPL enrollment was missing for 19.6%—primarily from Massachusetts. In the minor referral sample specifically, 15% of referrals were excluded due to missing or ambiguous exclusion status. FRPL data was missing for 24% of the sample, again largely due to incomplete reporting in Massachusetts and, to a lesser extent, Minnesota. Oregon had the most complete data. A Little’s MCAR test indicated that data were not missing completely at random, driven in part by these state-level reporting differences. Due to this pattern and the extent of

missingness, the FRPL variable was excluded from the final model. All remaining variables demonstrated low multicollinearity and were retained for analysis.

The full minor referral sample consisted of 25,319 unique students with 9,320 students identified as White (36.8%), 6,263 Black students (24.7%), 5,586 Latiné students (22.1%), and 2,516 Other race (9.9%), the remaining students had missing race data, roughly 6.7% of students. From this, an analysis of distributions by state did not show patterns of missingness for student race by state. However, the percentage of missing race data varied by state with Minnesota having the largest percentage of missing race at almost 16% while Oregon was only missing 2.5% of referral race data. The variation in race data meant Georgia and Oregon had an oversized influence on the analysis and Georgia's racial composition was different in comparison to other states. In light of these findings, the main model examined exclusion for Georgia and other states separately. Missing data was also evaluated for the primary dependent variable in the model. The exclusion indicator variable had 16% of observations listed as neither exclusionary nor non-exclusionary, resulting in those observations being removed from the analysis.

After examining the distributions of referrals, students, and schools in the four states, a state variable was created to produce two groupings: Georgia and Other states, which included Minnesota, Massachusetts, and Oregon. The data from Georgia was significantly different than the other states in terms of racial composition and rate of exclusion, providing much higher odds of exclusionary discipline for students in Georgia. In fact, Georgia had 22 schools with 100 percent use of exclusionary discipline for minor referrals. These schools were subsequently removed as they would impact the predictive model. The distribution of exclusion rate by schools and states in the unfiltered minor referral sample can be found in Figure 4.

Finally, in the subset of data used to examine the impact of PBIS fidelity on exclusionary discipline 17% of referrals were excluded due to having responses that were neither exclusionary nor non-exclusionary, remaining covariate data were mostly complete. Specifically, 3.3% of referrals were missing student race data, 1.3% were missing urbanicity, and 1.3% were missing percent Black enrollment. Specific handling of the PBIS subset can be found in the following section.

Figure 4: Distribution of Exclusionary Discipline rates by School and State



Note: Overall, 23 schools exclusively used exclusionary discipline and of those, 22 were in Georgia, 1 in Oregon. In addition, 2 schools from Georgia used 0 exclusionary discipline and 1 school in Oregon.

Analytic Software

After processing the data, the analysis involved descriptive statistics and calculation of disproportionality measures, before fitting a multi-level logistic regression to examine differences in exclusionary discipline by student characteristics with covariates for school level factors. All analyses were performed using R Statistical Software (v4.4.1; R Core Team, 2024). The descriptive analysis used counts for exclusionary discipline by student characteristic to

calculate RR and RDR. Descriptive data examined the variation in rates of actions taken for major and minor referrals before aggregating the outcome data into exclusionary discipline and non-exclusionary. Multi-level logistic regression models were fit using the `glmer` function from the `lme4` R package (v1.1.35.5; Bates et al., 2015). Subsequently the `tab_model` function from the R package, `sjPlot`, was used to create regression output tables (v2.8.16; Lüdtke, 2024). Model fit was analyzed through a variety of functions including the `Anova` function from R software (v4.4.1; R Core Team, 2024) and the `car` package to test multicollinearity (v3.1.3; Fox & Weisberg, 2019). A full list of the packages used can be found in the supplemental methods.

The binary exclusionary discipline outcome variable indicated if the referral resulted in an exclusionary or non-exclusionary disciplinary response. The model included random intercepts for both student and school to account for repeated referrals within students and for clustering of students within schools. Fixed effects for state were included to capture contextual variation across Georgia, and the grouped data from Massachusetts, Minnesota, and Oregon. Student-level predictors included race, gender, IEP status, and whether the student had a prior referral in that school year. School-level predictors included percent of students enrolled who were Black, grand mean centered, and a categorical variable for school urbanicity. Due to substantial missingness, percent of students eligible for free or reduced-price lunch (FRPL) was excluded from the final model. To ensure adequate representation within each school, the analytic sample was restricted to schools with at least 10 students enrolled, at least 20 total referrals, and at least 10 students with referrals. In addition, students above the 99th percentile for the number of referrals received in a school year were removed as outliers. High referral students were more likely to receive exclusionary discipline and skew the data on the risk of exclusion for minor referrals. The multi-level models also incorporated a list-wise deletion to remove

observations with missing data. Patterns of missingness was examined prior to analysis using Little's MCAR test from the naniar R package (v1.1.0; Tierney & Cook, 2023).

Finally, another multi-level logistic regression model was fit to test for exclusion versus non-exclusionary responses as well as the moderating effect of PBIS implementation. To examine the potential moderating role of PBIS implementation fidelity on exclusionary discipline, separate multilevel logistic regression models were fit for Tier 1 and Tier 2 fidelity scores. Each fidelity tier was examined using a main model as well as an interaction model between state and fidelity score. An iterative model-building process was used, with comparisons of AIC, BIC, and nested model chi-squared tests guiding final model selection. The logistic regression provided odds ratios for exclusion for a referral based on student race, ethnicity, gender, and IEP status as well as for school level factors.

PBIS Subsample for Analysis

After applying exclusion rules for missing or invalid fidelity scores and covariates, the final analytic sample for the Tier 1 PBIS model included 52 schools across the four participating states, representing 9,988 students and 36,174 minor referrals. Georgia contributed the majority of the analytic sample with 31 schools, 5,581 students, and 14,703 minor referrals—over half of the Tier 1 dataset. Due to this disproportionate representation and the unique disciplinary patterns observed in Georgia, schools were grouped into two clusters: Georgia and a combined “Other States” group consisting of Minnesota, Oregon, and Massachusetts. This clustering approach allowed for the examination of PBIS implementation effects while accounting for state-level variation in discipline practices and data availability. For the Tier 2 PBIS model, the final analytic sample included 33 schools, 5,877 students, and 22,326 minor referrals. Of these, 21 schools were from Georgia, which also accounted for approximately 57% of all Tier 2 student

observations. Consistent with the Tier 1 approach, schools were clustered by state to facilitate interpretation of the Tier 2 PBIS moderation effects and control for differences in missingness and disciplinary practices across states.

Measurement

Behavior referrals contained information on the type of challenging behavior, the date and location of the incident, the participation of others in the referral, the perceived motivation for the behavior, and the administrative actions taken as a result of the referral. The date and student identification variables were used to create an indicator variable for prior referrals to control for students receiving more exclusionary consequences as a result of multiple referrals. Additional student and state level variables were also included in the data as described previously.

Minor Referral Identification

A binary variable to indicate if the referral was considered minor was present in the dataset. A sensitivity check was conducted check the validity of the indicator variable. The behavior labels, which also contained tags for minor behaviors, were compared with the indicator variable and were found to have a 1-to-1 correspondence, which affirmed the validity of the “Is minor?” indicator variable.

In the SWIS referral data there were 42 options for challenging behaviors in the data dictionary with multiple behaviors possible for a single referral. Behaviors were labeled as major or minor according to the SWIS data dictionary where minor behaviors had a prefix “minor” plus the behavior type, resulting in a single challenging behavior variable composed of a list of all challenging behaviors in the referral. If a referral contained exclusively minor behaviors, it was labeled as minor, otherwise the referral was considered major.

Exclusionary Discipline Variable

The administrative response to behavior was processed in a similar manner by creating a singular variable with multiple consequences possible per referral. The SWIS data dictionary contained 18 options for administrative action, which were categorized into three groups: exclusionary, non-exclusionary, and other based on the loss of academic instructional time. An additional variable to indicate an exclusionary discipline response was created using a similar rule to the challenging behavior variable. If any of the consequences attached to the referral were exclusionary the variable was set to 1 to indicate exclusion. If there were a mix of non-exclusionary and other responses, the referral was listed as non-exclusionary. Only if the actions taken were in neither category was the referral labeled as neither. The subset of administrative actions deemed neither exclusionary nor non-exclusionary was removed from analysis. The removed administrative responses included other administrative action, unknown administrative action, bus suspension, action pending, additional support requested, and alternative placement. Exclusionary responses were comprised of time in the office, detention, in-school suspension, out-of-school suspension, expulsion. Non-exclusionary responses consisted of loss of privileges, conference with the student, parent contact, individualized instruction, Saturday school, restitution, restorative justice, and community service. The breakdown of administrative responses and exclusionary discipline can be found in Table 5.

Dependent Variables by Research Question

The final sample contained referral, student, and school level variables as well as ECS assigned identification variables. Each observation contained the following identifiers: a school year, referral id, school id, school district id, and a student id number. The student level details included grade, race, ethnicity, gender, IEP status, and a binary indicator for Hispanic students.

At the referral level each observation contained date and time, incident location, referring teacher id, the transformed list of challenging behaviors, the list of administrative actions taken, the binary indicator for minor referrals, an indicator variable for the presence of a prior referral for the student, a binary indicator variable for an exclusionary discipline response, and for referrals that result in suspension, the number of days of ISS, OSS, or expulsion. Finally at the school level the data included state name, NCES total school enrollment and enrollment by race/ethnic group, gender, and grade, a binary indicator for charter schools, urbanicity, NCES FRPL enrollment, as well as an ODR count, number of students with ODRs, and ODR rate for the school. These final numbers did not align with calculated ODR rate based on the referral data. In addition, at the state level, a variable was added to group states to compare Georgia to other states in the sample.

The dependent variables for the study changed by research question. RQ1 and RQ2 descriptively examined the trends in minor referrals by behavior type and the actions taken in response to the referral. The frequency of different behavior types for minor referrals is reported for RQ1. The primary variable in RQ2 was the frequency of actions taken. While in RQ3, Risk Ratio (RR) and Raw Differential Representation (RDR) were calculated for major referrals using student demographics, specifically race, ethnicity, gender, and IEP status to examine disproportionality in the receipt of a major referral. For example, RQ3 explored whether Black students more likely to receive a major or minor referral compared to White students. After descriptive analysis, RQ4 and RQ5 incorporated statistical analysis to test for disproportionality in exclusionary discipline for minor behavior referrals. The multi-level logistic regression model provided log odds, which were used to calculate odds ratios, to quantify disproportionality in the data for students belonging to different groups.

Data Analytic Plan

Descriptive data consisted of counts and frequency of challenging behaviors, administrative action taken by student demographic group and exclusionary discipline combined with RR and RDR to examine disproportionality. Logistic multi-level regression was used to test for disproportionality in exclusionary discipline while accounting for the clustering of referrals within student and schools. The resulting three level model provided odds of exclusionary discipline for minor referral and allowed for examination of differences by state. The final model incorporated a PBIS fidelity variable to test for a moderating effect of high quality PBIS implementation.

Overview of Data Analysis

Step 1 of the analytic plan was to examine the patterns in the referral data to understand common challenging behaviors, whether certain groups of students received more referrals for specific behaviors, and the common actions taken as a result of minor referrals. Although the study aimed to examine disproportionality for minor referrals, an important step in the analysis was to analyze the rates of referrals broadly, including major referrals, to ascertain if rates of major referrals for certain groups differed from rates of minor referrals for those students. This would indicate that students who may be at increased risk of receiving a major referral were less likely to be included in the minor referral data and mask the potential for disproportionality in minor referrals.

In step 2 after examining the initial descriptive data in RQ1, RQ2, RQ3 both Risk Ratio (RR) and Raw Differential Representation (RDR) were calculated to explore disproportionality based on race, ethnicity, gender, and IEP status (RQ4). In addition to RDR, RDR per 100 students (RDR_{100}) was reported to generalize the metric across contexts. These metrics provided

information on disproportionality for minor referrals by student characteristics, without controlling for other student and school level factors. Although RR and RDR are useful metrics they do not control for covariates in the data. To analyze disproportionality while controlling for student and school level characteristics, a multi-level logistic regression analysis was conducted with student and school level random effects as well as state level fixed effect to account for clustering of referrals in schools. In the final research question, Tier 1 and Tier 2 TFI scores were added to the model to examine the moderating effect of PBIS implementation on disproportionality (RQ5).

Risk Ratio Calculation

For analysis of disproportionality, RR measured the likelihood of receiving an exclusionary response for different student groups relative to a reference group where a ratio greater than 1 indicated the target group was disproportionately disciplined compared to the reference group. The risk ratio was calculated from the number of target students excluded (n_t), the total number of target students in the population (N_t), the number of reference students excluded (n_{ref}) and the total number of reference students in the population (N_{ref}). For this data the population totals are based on the students in the referral sample meaning the risk should be interpreted as for a student receiving a referral, “what is the chance it is a major referral?” or for a student receiving a referral, “what is the risk of exclusion?”

$$RR = \frac{\frac{n_t}{N_t}}{\frac{n_{ref}}{N_{ref}}}$$

Raw Differential Representation Calculation

Raw differential representation (RDR) measures the differences in the proportion of exclusionary responses across student groups to provide an estimate for the expected number of

students receiving exclusion based on an equitable system. RDR provides a quantifiable number in terms of impact on students, but to contextualize across schools RDR per 100 students (RDR_{100}) is a preferred metric. RDR can be found using the following formula with the risk ratio for the reference group and the target group:

RDR:

$$RDR = n_t - \left(\frac{n_t}{RR_T - RR_{ref}} \right)$$

RDR_{100} :

$$RDR_{100} = \frac{RDR}{n_t} \times 100$$

RQ4 Multi-level Logistic Regression Model

For the statistical models in RQ4 and RQ5 the binary outcome variable of exclusionary or non-exclusionary response to behavior was predicted by student demographic variables. For the purpose of addressing RQ4, which examined disproportionality in exclusionary discipline for minor referrals, an iterative process of building the multi-level logistic regression was used to determine the best model. The final model included student race, ethnicity, gender, IEP status, percent Black enrollment and urbanicity at the school level as well as state fixed effect. The model equation is shown below and was estimated using the `glmer()` function in the `lme4` package (Bates et al., 2015) in R.

Main Minor Referral Model

The model estimates the log odds of exclusionary discipline, y , conditional on the receipt of a minor behavior referral, for each referral, i , clustered within student, j , and nested within school,

k. The exponentiated estimate coefficient for exclusion can be interpreted as the odds of a referral given specific student and school characteristics leading to exclusionary discipline.

Level 1 (Referral)

$$\text{logit} \left(p(\text{Exclusion}_{ijk} = 1) \right) = \beta_{0ijk} + \varepsilon_{ijk}$$

The equation above represents level one, the individual observations, denoted in the subscript as *i* for referral observations. The β represent the regression coefficient of the intercept and ε represents the error present at level 1. The model did not contain level one covariate predictors at the referral level.

Level 2 (Student)

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k}\text{Race}_{jk} + \gamma_{02k}\text{Gender}_{jk} + \gamma_{03k}\text{IEP}_{jk} + \gamma_{04k}\text{Prior Referral}_{jk} + \tau_{0jk}$$

The second level of the model contains the coefficient estimates for student level predictors in the model, denoted as *j*, and the coefficients represented as gamma, γ . Race is dummy coded with White as the reference group compared to Black, Latiné, and other race. Gender is similarly dummy coded with male as the reference compared to female. IEP status and prior referral are also dummy coded with 0 representing no IEP and no prior referral as the reference groups. The student level random intercept is represented at this level, τ_{0jk} .

Level 3 (School)

$$\gamma_{00k} = \delta_{000} + \delta_{001}\text{Black Enroll}_k + \delta_{002}\text{Urbanicity}_k + \delta_{003}\text{State}_k + \mu_{00k}$$

Fixed slopes across schools:

$$\gamma_{01k} = \delta_{010}\text{Race}$$

$$\gamma_{02k} = \delta_{020}\text{Gender}$$

$$\gamma_{03k} = \delta_{030}\text{IEP}$$

$$\gamma_{04k} = \delta_{040}\text{Prior Referral}$$

The third level of the equation contains school level coefficients, δ , denoted by subscript k , as well as the fixed effects from the student level variables. The school level random intercept is estimated at level 3 and represented as μ_{00k} . Percent Black enrollment is grand mean centered while urbanicity and state are dummy coded with the reference groups as urban and Georgia respectively. In addition, the model assumes fixed slopes for race, gender, IEP status, and prior referral across schools.

RQ5 PBIS Moderation Multi-level Logistic Regression Model

In RQ5 a series of multilevel logistic regression models were specified to evaluate whether school-level PBIS implementation moderated the relationship between student demographics and the likelihood of exclusionary discipline for minor referrals. Due to the limited number of clusters at the district and state levels—many of which fell below the threshold of 30 units, which can lead to unstable variance estimates—state was modeled as a fixed effect rather than as a random intercept. Model comparison using fit statistics indicated that a dichotomous grouping of states, Georgia vs. all other states combined, best captured variation in exclusionary practices across the sample. Consequently, an interaction term was included between state (grouped) and PBIS fidelity scores at both Tier 1 and Tier 2 levels to examine potential moderation effects.

The models included fixed effects for student-level predictors (race, gender, IEP status, and prior referral), as well as school-level contextual variables (urbanicity and percentage of Black student enrollment). PBIS fidelity scores from the Tiered Fidelity Inventory (TFI) were entered as continuous predictors and were grand mean centered at the school level PBIS fidelity score. The reference group again was White male students without IEPs who had not previously had a minor referral and who were in urban schools in Georgia with a mean PBIS fidelity score.

Random intercepts were specified at both the student and school levels to account for the nested structure of the data, where multiple referrals were clustered within students and students were clustered within schools. All models were estimated using the `glmer()` function from the `lme4` package in R (Bates et al., 2015). Equations for the fixed and interaction models are provided below informed by *Multilevel Modeling Using R* (Finch et al., 2019).

PBIS Minor Referral Model with and without interaction term

Similar to the main model, the PBIS model adds a covariate at the school level for PBIS fidelity of implementation, TFI. An additional interaction term was added as well to examine the interaction between state (Georgia versus other states) and TFI score. The model is identical to the main model at level 1 and level 2 and is again estimating the log odds of a minor referral resulting in exclusionary discipline, conditional on the receipt of a minor referral.

Level 1 (Referral)

$$\text{logit} \left(p(\text{Exclusion}_{ijk} = 1) \right) = \beta_{0ijk} + \varepsilon_{ijk}$$

In the equation above the subscript i denotes individual referral observations, j represents the students who may have received multiple referrals, and k denotes the schools these students are located within. The β represents the regression coefficient of the intercept.

Level 2 (Student)

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} \text{Race}_{jk} + \gamma_{02k} \text{Gender}_{jk} + \gamma_{03k} \text{IEP}_{jk} + \gamma_{04k} \text{Prior Referral}_{jk} + \tau_{0jk}$$

The second level of the model contains the coefficient estimates for student level predictors, j , and the coefficients are represented by γ . Additionally, level 2 contains the student level random intercept estimates, τ_{0jk} . Again Race, Gender, IEP status, and Prior referral are dummy coded with the reference groups as White, male, no IEP, no prior referral respectively.

Level 3 (School)

$$\gamma_{00k} = \delta_{000} + \delta_{001}\text{Black Enroll}_k + \delta_{002}\text{Urbanicity}_k + \delta_{003}\text{State}_k + \delta_{004}\text{TFI}_k \\ + \delta_{005}(\text{State}_k \times \text{TFI}_k) + \mu_{00k}$$

Fixed slopes across schools:

$$\gamma_{01k} = \delta_{010}\text{Race}$$

$$\gamma_{02k} = \delta_{020}\text{Gender}$$

$$\gamma_{03k} = \delta_{030}\text{IEP}$$

$$\gamma_{04k} = \delta_{040}\text{Prior Referral}$$

The third level of the equation contains the school level coefficients, k , as well as the fixed effects from the student level variables and the school level random intercept estimates, μ_{00k} . In level three, the coefficients are represented by δ . Similar to the previous model, percent Black student enrollment is grand mean centered and urbanicity and state are dummy coded, again with the reference groups as urban and Georgia. The TFI score was also grand mean centered at the school level resulting in a centered TFI variable that represents each school's fidelity score relative to the overall mean of schools in the sample. In addition, this model again uses fixed slopes for level two covariates, race, gender, IEP status, and prior referral, across schools.

In these models, the outcome variables were the log-odds of receiving exclusionary discipline for a minor referral. The primary focus was on the interaction term, denoted as $\beta_{3k}(\text{State}_k * \text{TFI}_k)$ representing the interaction between student-level characteristics and school-level PBIS Tier 1 or Tier 2 implementation fidelity, moderated by state. The interaction allowed for the examination of the relationship between PBIS implementation and exclusionary discipline by state context. The models also included control variables at both the student and school levels. School-level covariates $\beta_k X_k$ included percent Black enrollment and urbanicity,

while student-level covariates $\beta_{jk}X_{jk}$ included prior referral history. Random intercepts were specified at the school and student levels to account for the nested data structure.

Intraclass correlation (ICC) was used to examine the effect of clustering. The ICC is a measure between 0 and 1 to quantify the proportion of the outcome that can be attributed to shared contextual characteristics with a 0 indicating no-information is shared and observations are independent. This is important as independence of observation is a key assumption in regression analysis. The ICC was used as a metric to identify the extent to which clustering impacts the results and the method to account for the clustering in the model. A high ICC indicated the need to account for clustering through methods such as multi-level modeling and random effects. Additionally, marginal and conditional R^2 were reported to analyze the share of variance explained by the model coefficients and the student and school clustering. Marginal R^2 shows the variance explained by the model coefficient whereas Conditional R^2 shows the combined variance explained by coefficients and clustering.

Chapter III: Results

The study examined a multistate dataset of middle school behavior referrals from schools implementing Positive Behavioral Interventions and Supports (PBIS) during the 2022–2023 and 2023–2024 academic years. The data were collected from the School-Wide Information System (SWIS), a discipline tracking system used by schools implementing PBIS. The analysis focused on assessing patterns of exclusionary discipline in response to minor behavior referrals, with a predictive three-level logistic regression models to examine the odds of a referral resulting in exclusionary discipline by student demographics and the moderating role of school-level PBIS implementation fidelity.

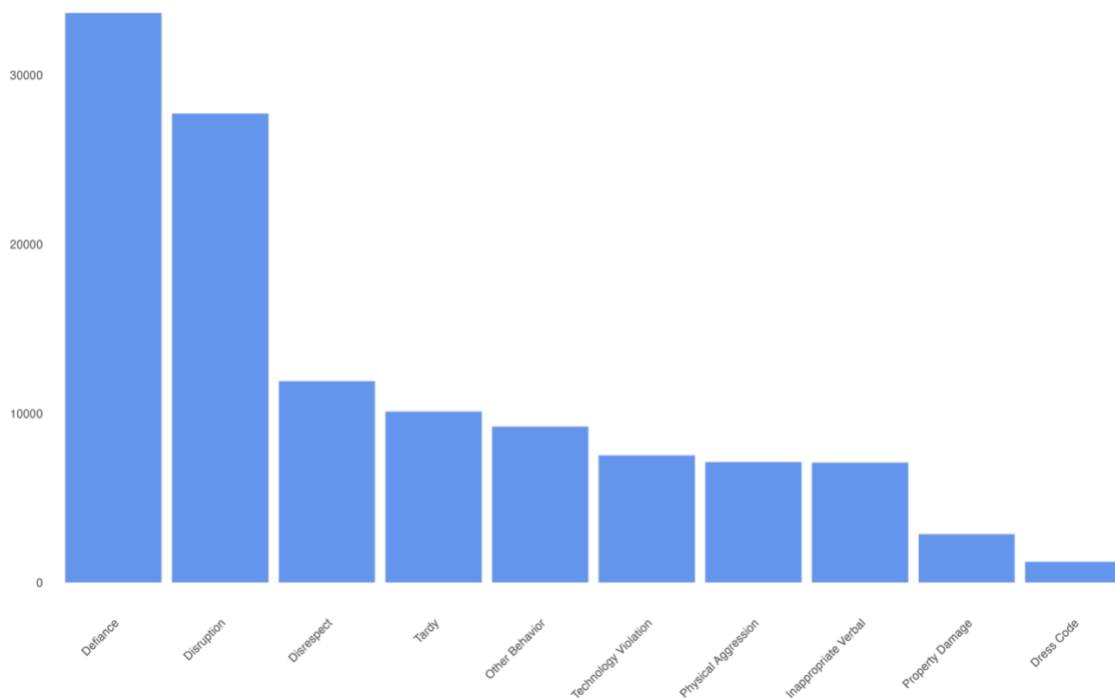
The unfiltered dataset contained total of 271,979 referrals, representing 49,203 unique students across four states: Georgia, Minnesota, Oregon, and Massachusetts. Although both major and minor referrals were included in the dataset, the primary research question focused on exclusionary discipline in minor referrals, leading to a sample of 110,864 minor referrals prior to filtering. The risk ratios were calculated at the student level based on the unfiltered data to provide insight into the risk of receipt of major referral or exclusion ever occurring. The minor referral data was subsequently filtered to a subset optimized for model fit. First, the data were filtered by removing students above the 99th percentile in referrals received per school year then removing schools with fewer than 10 students enrolled, 10 unique students referred, or fewer than 20 total referrals. Finally, referrals with missing exclusionary discipline data, as well as a listwise deletion of observations missing the primary model covariates were filtered. The PBIS fidelity data was also filtered to remove missing and zero scores after analysis suggested zero scores were more similar to missing PBIS fidelity scores than non-zero scores, resulting in varying analytic samples based on the research question.

RQ1: Patterns of Challenging Behaviors in Minor Referrals

An analysis of the patterns of challenging behaviors examined counts across minor referrals. Although most referrals listed a single behavior—the average number of behaviors per referral was 1.06—some referrals documented multiple behaviors. As a result, the total number of behaviors exceeded the number of referrals. Figure 5 displays the overall frequency of challenging behaviors and Table 2 shows the frequency of each challenging behavior label for minor referrals by state.

Descriptive analyses showed consistent behavior trends across referral types. For minor referrals the most common challenging behaviors aggregated across states were defiance (29%), disruption (23%), and disrespect (10%). Followed by tardiness, other behavior, technology violations, physical aggression, and inappropriate language all of which were between 5-9% of the referrals. A similar pattern was observed for major referrals with defiance and disruption as the most common challenging behaviors listed. A difference for major referrals was the prevalence of physical aggression (9%) compared to 6% in minor referrals. For major referrals, defiance (17%), disruption (16%), and disrespect (9%) continued to be the most common challenging behaviors followed by physical aggression (9%). Major referrals had more types of challenging behaviors and a higher percentage of objective behaviors with skipping class, inappropriate language, and fighting consisting of the remaining challenging behaviors between 5-9%. The distribution suggests that most of the referrals were for subjective behaviors, particularly prevalent for minor behavior referrals. The multi-level analysis did not include behavior category as a predictor in multilevel models, due to the wide range of behavior types and limited sample sizes within each category, which prevented the multi-level models from converging on estimates for odds of exclusion for different behaviors.

Figure 5 Frequency of Challenging Behaviors for Minor Referrals



Note: Some referrals contained multiple behavior labels.

Table 2
Frequency of challenging behaviors for minor referrals by State

	Total n (%)	Georgia n (%)	Massachusetts n (%)	Minnesota n (%)	Oregon n (%)
Defiance	31814 (28.7)	9452 (33.8)	8152 (33.9)	5030 (34.7)	9180 (20.7)
Disruption	25501 (23.0)	5905 (21.1)	6598 (27.5)	3572 (24.6)	9426 (21.3)
Disrespect	10701 (9.7)	2018 (7.2)	1691 (7.0)	1311 (9.0)	5681 (12.8)
Tardy	100007 (9.0)	896 (3.2)	1306 (5.4)	772 (5.3)	7033 (15.9)
Other Behavior	8688 (7.8)	3982 (14.2)	1407 (5.9)	685 (4.7)	2614 (5.9)
Technology Violation	7226 (6.5)	1231 (4.4)	1809 (7.5)	924 (6.4)	3262 (7.4)
Physical Aggression	6881 (6.2)	1971 (7.0)	1074 (4.5)	828 (5.7)	3008 (6.8)
Inappropriate Verbal	6226 (5.6)	1402 (5.0)	1315 (5.5)	958 (6.6)	2551 (5.8)
Property Misuse	2633 (2.4)	353 (1.3)	441 (1.8)	362 (2.5)	1477 (3.3)
Dress Code	1187 (1.1)	798 (2.9)	240 (1.0)	58 (0.4)	91 (0.2)

Note: Percent is calculated within each state for the challenging behavior while total percentage is calculated from all minor behaviors, but due to rounding the totals may not equal 100. Percentages were calculated with the cleaned sample of minor referral data.

Behavior patterns for minor referrals also varied by state. Defiance was the most frequently listed behavior in Georgia (33.8%), Massachusetts (33.9%), and Minnesota (34.7%), while disruption was the most common behavior in Oregon (20.7%). Tardiness was more prevalent in Oregon (15.9%) than in other states, where it represented about 5% of challenging behaviors. Similarly, other behavior label was cited more often in Georgia (14.2%) compared to Massachusetts (5.9%), Minnesota (4.7%), and Oregon (5.9%). Technology violations were more frequently listed in Massachusetts (7.5%) and Oregon (7.4%) than in Georgia (4.4%) or Minnesota (6.4%).

Differences in labeling of challenging behaviors also emerged when comparing major referrals across states. In particular, Minnesota was more likely to label a behavior as disruption (28%) compared to defiance (11%) while in other states defiance was the most common. In addition, although Georgia used the other behavior label more for minor referrals, Minnesota was over twice as likely to use the other behavior label for major referrals compared to other states. Other differences in major referrals existed in labels such as elevated use of skipping and truancy in Oregon and Massachusetts as well as a higher share of referrals for technology violation in Minnesota. A full comparison of behavior labels for minor and major referrals by state can be found on Table 3. These results suggest that challenging behavior referral patterns vary by state and are influenced by school or district-level practices in documenting student behavior.

Table 3

Challenging Behaviors for Major and Minor Referrals by State

Subjective	Minor Referrals					Major Referrals				
	Total (%)	Georgia (%)	Massachusetts (%)	Minnesota (%)	Oregon (%)	Total (%)	Georgia (%)	Massachusetts (%)	Minnesota (%)	Oregon (%)
Defiance/Insubordination/Non-compliance	31,814 (28.7)	9,452 (33.8)	8,152 (33.9)	5,030 (34.7)	9,180 (20.7)	26,943 (16.7)	11,133 (17.6)	5,382 (21.1)	2,901 (11.0)	7,527 (16.3)
Disruption	25,501 (23.0)	5,905 (21.1)	6,598 (27.5)	3,572 (24.6)	9,426 (21.3)	26,196 (16.3)	10,227 (16.2)	3,717 (14.6)	7,272 (27.6)	4,980 (10.8)
Disrespect	10,701 (9.7)	2,018 (7.2)	1,691 (7)	1,311 (9)	5,681 (12.8)	13,824 (8.6)	8,729 (12.8)	2,579 (10.1)	399 (1.5)	2,117 (4.6)
Inappropriate/Abusive Language	6,226 (5.6)	1,402 (5)	1,215 (5.5)	958 (6.6)	2,551 (5.8)	12,715 (7.9)	4,635 (7.3)	2,469 (9.7)	1,698 (6.5)	3,913 (8.5)
Dress Code	1,187 (1.1)	798 (2.9)	240 (1)	58 (0.4)	91 (0.2)	458 (0.3)	352 (0.6)	43 (0.2)	4 (0.0)	59 (0.1)
Other Behavior	8,688 (7.8)	3,982 (14.2)	1,407 (5.9)	685 (4.7)	2,614 (5.9)	6,553 (4.1)	2,034 (3.2)	522 (2.0)	2,323 (8.8)	1,674 (3.6)
Objective										
Physical Aggression	6,881 (6.2)	1,971 (7)	1,074 (4.5)	828 (5.7)	3,008 (6.8)	14,865 (9.2)	4,413 (7.0)	2,392 (9.4)	2,515 (9.6)	5,545 (12.0)
Skipping Class / Truancy	NA					14,540 (9.0)	3,051 (4.8)	3,547 (13.9)	1,543 (5.9)	6,399 (13.9)
Fighting	NA					11,739 (7.3)	8,056 (12.8)	244 (1.0)	1,317 (5.0)	2,122 (4.6)
Inappropriate Location	NA					5,998 (3.7)	1,522 (2.4)	1,754 (6.9)	2,448 (9.3)	274 (0.6)
Technology Violation	7,226 (6.5)	1,231 (4.4)	1,809 (7.5)	924 (6.4)	3,262 (7.4)	5,753 (3.6)	1,933 (3.1)	838 (3.3)	1,610 (6.1)	1,372 (3.0)
Harassment	NA					4,922 (3.1)	640 (1.0)	424 (1.7)	478 (1.8)	3,380 (7.3)
Possession Tobacco Drugs Alcohol	NA					4,629 (2.9)	2,207 (3.5)	172 (0.7)	576 (2.2)	1,674 (3.6)

Table 3 continued

Tardy	10,007 (9.0)	896 (3.2)	1,306 (5.4)	772 (5.3)	7,033 (15.9)	2,847 (1.8)	911 (1.4)	574 (2.3)	255 (1.0)	1,107 (2.4)
Bullying	NA					2,777 (1.7)	1,097 (1.7)	132 (0.5)	369 (1.4)	1,179 (2.6)
Property Damage Misuse	2,633 (2.4)	353 (1.3)	441 (1.8)	362 (2.5)	1,477 (3.3)	2,172 (1.4)	435 (0.7)	428 (1.7)	288 (1.1)	1,021 (2.2)
Cheating / Plagiarism	NA					2430 (1.5)	847 (1.3)	217 (0.9)	196 (0.7)	1,170 (2.5)
Possession Weapon / Combustible	NA					779 (0.5)	292 (0.5)	18 (0.1)	96 (0.4)	373 (0.8)
Display of Affection	NA					739 (0.5)	537 (0.9)	66 (0.3)	43 (0.2)	93 (0.2)
Bomb Threat	NA					88 (0.1)	34 (0.1)	14 (0.1)	2 (0.0)	38 (0.1)
Gang Affiliation	NA					125 (0.1)	52 (0.1)	5 (0.0)	2 (0.0)	66 (0.1)
Arson	NA					23 (0.0)	7 (0.0)	1 (0.0)	3 (0.0)	12 (0.0)

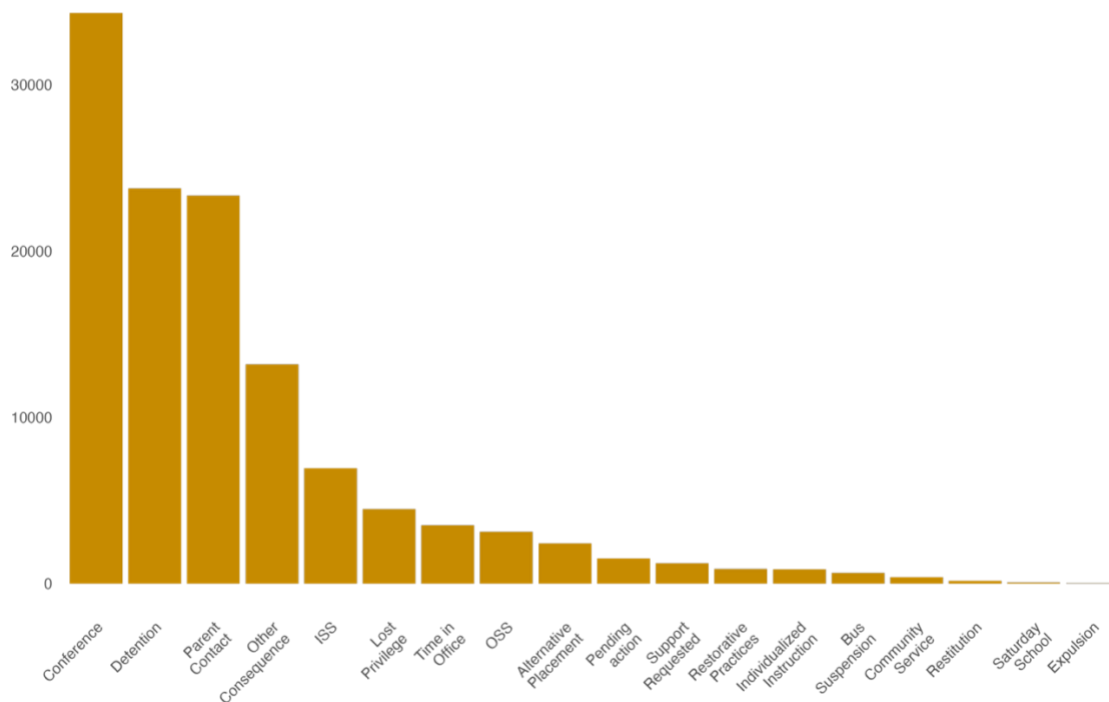
Note: Behaviors are grouped by exclusion and referral type. Totals reflect number of times each action was assigned across all states with the following behavior categories combined due to discrepancies in naming between major and minor referrals as well as grouping of similar behavior: property violation (property damage & property misuse), possession of illegal substances (alcohol, tobacco, drugs), possession of a weapon (weapon & combustible), academic dishonesty (Cheating & plagiarism), and truancy (skipping class & truancy). Percentages are within-state proportions separated by referral type (major or minor) and due to rounding may not sum to 100. Percentages of total contribution of a behavior label to a category is calculated across referral type.

RQ2: Patterns of Administrative Actions in Minor Referrals

Administrative responses to minor referrals were recorded as a primary action, with optional secondary actions, resulting in an average of 1.09 actions per referral. For analytic purposes, responses were categorized into three groups: exclusionary, non-exclusionary, and neither. Exclusionary responses included actions that removed the student from the instructional environment, such as time in the office, detention, in-school suspension (ISS), out-of-school suspension (OSS), and expulsion. Non-exclusionary responses included actions such as staff conferences, parent contact, loss of privileges, and restorative or instructional supports. A third category, labeled “neither” included responses that could not be clearly classified due to their administrative nature or missing outcome information (e.g., action pending, alternative placement, bus suspension, and requests for additional support). Actions in this third category were excluded from subsequent multilevel models.

Table 4 displays the frequency of administrative responses across all minor referrals and Figure 6 visualizes the frequency of administrative actions to minor referrals. The most frequent actions taken as a result of minor behavior referrals included conference with staff (28.4%), detention (19.7%), parent contact (19.3%), other consequence (10.9%), and ISS (5.7%). The remaining responses each contributed less than 5% of the actions taken: loss of privileges (3.7%), time in the office (2.9%), OSS (2.6%), alternative placement (2%), action pending (1.2%), and additional support (1%). Rarely used responses, each accounting for less than 1% of actions, included restorative practices, individualized instruction, bus suspension, community service, restitution, Saturday school, and expulsion.

Figure 6 Frequency of Administrative Actions for Minor Referrals



Note: Each referral may contain multiple administrative action.

Table 4
Frequency Administrative Responses to Minor Referrals

Count (%)	
Exclusion	
Time in Office	3,503 (2.9)
Detention	23,756 (19.7)
ISS	6,931 (5.7)
OSS	3,117 (2.6)
Expulsion	6 (0)
Non-Exclusion	
Student Conference	34,316 (28.4)
Parent Contact	23,337 (19.3)
Privileges Removed	4,463 (3.7)
Restorative Practices	867 (0.7)
Individualized Instruction	855 (0.7)
Community Service	374 (0.3)
Restitution	150 (0.1)
Saturday School	51 (0)
Other	
Pending	1,498 (1.2)
Additional Support	1,207 (1)
Other	13,188 (10.9)
Alternative Placement	2,420 (2)
Bus Suspension	634 (0.5)

Note: The count is aggregated across states but is restricted to only minor referrals

An examination of administrative actions taken by state highlighted stark differences in the approach to behavior management and recording of behavior referrals by state. Georgia issued ISS and OSS more frequently than other states. Parent contact was more prevalent in Massachusetts and Oregon, while Massachusetts reported a higher use of loss of privileges than other states. Although Georgia had a total of 26,049 minor behavior referrals with race data compared to 23,168 referrals in Massachusetts, Georgia issued ISS as a response for 6,566 referrals compared to only 31 times in Massachusetts. Other differences in response to discipline appeared as well. Oregon had a much higher rate of detention, 37% of the referrals, compared to 5% in Minnesota, 9% in Massachusetts, and 10% in Georgia. These differences demonstrate the variety of approaches to challenging behavior that exist between states and schools. Though it is also possible that schools implement these administrative actions differently from each other whereby detention in Oregon may look different from detention in Georgia. A full breakdown of administrative responses by state is presented in Table 5.

Across all states, 39% of minor referrals resulted in an exclusionary discipline. However, this rate varied substantially by state: 58% of minor referrals in Georgia resulted in exclusion, compared to 48% in Oregon, 16% in Massachusetts, and 15% in Minnesota. Nevertheless, the high percentage of exclusionary discipline suggests students miss instructional time due to minor behaviors and receipt of minor behavior referrals. Table 6 contains the frequency of exclusionary and non-exclusionary discipline and Table 7 displays exclusionary discipline patterns by student characteristics.

Table 5**Administrative Responses to Major and Minor Referrals by State**

	Minor Referrals					Major Referrals				
	Total	Georgia	Massachusetts	Minnesota	Oregon	Total	Georgia	Massachusetts	Minnesota	Oregon
Exclusion										
Detention	23,756	3,146	2,346	786	17,478	20,437	1,916	3,521	1,142	13,858
ISS	6,931	6,565	31	106	229	32,975	24,810	446	637	7,082
Time-in Office	3,503	571	1,095	756	1,081	9,426	859	6,128	608	1,831
OSS	3,117	2,843	45	129	100	32,963	20,351	494	4,723	7,395
Expulsion	6	2	1	1	2	130	56	1	0	73
Non-Exclusion										
Student Conference	34,316	5,679	7,943	8,066	12,628	31,157	3,689	7,101	8,405	11,962
Parent Contact	23,337	3,800	9,357	1,116	9,064	20,754	3,677	4,947	4,337	7,793
Privileges Removed	4,463	1,207	2,105	394	757	5,303	855	1,555	1,421	1,472
Restorative Practices	867	196	81	435	155	512	154	139	100	119
Individual Instruction	855	44	413	140	258	2,791	87	227	1,914	563
Community Service	374	179	11	5	179	323	54	13	17	239
Restitution	150	15	16	27	92	1,399	59	46	560	734
Saturday	51	46	4	0	1	631	596	15	1	19
Other										
Other	13,188	4,284	2,180	1,803	4,921	12,122	4,119	2,268	2,278	3,457
Alternative Placement	2,420	924	395	601	500	2,923	1,104	648	580	591
Pending	1,498	1,047	228	115	108	7,557	5,986	510	940	121
Additional Support	1,207	271	681	192	63	3,830	230	3,286	194	120
Bus Suspension	634	596	1	0	37	2,389	2,177	27	75	110

Note: Administrative actions are grouped by exclusion and referral type. Totals reflect number of times each action was assigned across all states

Table 6**Exclusionary Discipline from Minor Referrals by State**

	Non-Exclusion	Exclusion	Neither	Total Referrals	Exclusion (%)
Georgia	9,148	12,820	6,040	21,968	58.36
Massachusetts	17,736	3,426	2,871	21,162	16.19
Minnesota	10,068	1,774	2,658	11,842	14.98
Oregon	20,327	18,801	5,195	39,128	48.05

Note: The percent exclusionary discipline is at the incident level meaning the percentage represents the number of referrals leading to exclusionary discipline, which is different from percent of exclusionary discipline by student as many students had multiple referrals.

Differences in exclusionary discipline outcomes must be interpreted in the context of variation in student demographics across states. For example, in Georgia, Black students comprised over 50% of students in the sample, while Latiné students represented 11%. In contrast, Massachusetts, Minnesota, and Oregon had lower percentages of Black students—4.5%, 14.9%, and 2.2%, respectively—and White students were the majority in all states except Georgia. In addition, differences in rates of exclusion existed between states. Georgia and Oregon had a higher percentage of referrals resulting in exclusionary discipline across groups.

Differences in exclusion were most pronounced for race, but student IEP status also showed differences. Massachusetts had higher percentage of students with IEPs, 28.8% compared to an average of 16% in the other states. Exclusionary discipline by gender was uniform across states with 35% minor referrals for females and 65% for males. Because of the small number of students in the other gender category, 12 students across states, a dichotomous gender variable was used in the analysis.

Table 7
Descriptive Statistics of Student Characteristics by State

	Students <i>n</i> (%)	Referral <i>n</i> (%)	Exclusion <i>n</i> (%)
Georgia			
Race			
Black	5,620 (51.3)	15,489 (55.3)	7,167 (46.3)
Latine	1,212 (11.1)	2,902 (10.4)	1,340 (46.2)
Other	1,268 (11.6)	2,946 (10.5)	1,557 (52.9)
White	1,941 (17.7)	4,712 (16.8)	2,252 (47.8)
Gender			
Female	3,892 (35.7)	9,218 (32.9)	4,017 (43.6)
Male	7,004 (64.3)	18,790 (67.1)	8,803 (46.9)
IEP Status			
No IEP	9,220 (84.4)	22,633 (80.8)	10,397 (45.9)
IEP	1,704 (15.6)	5,375 (19.2)	2,423 (45.1)
Massachusetts			
Race			
Black	138 (4.5)	1,278 (5.3)	242 (18.9)
Latine	815 (26.7)	10,149 (42.2)	900 (8.9)
Other	140 (4.6)	1,287 (5.4)	246 (19.1)
White	1,739 (56.9)	10,454 (43.5)	1,840 (17.6)
Gender			
Female	1,089 (35.7)	7,370 (30.7)	929 (12.6)
Male	1,966 (64.4)	16,660 (69.3)	2,495 (15.0)
IEP Status			
No IEP	2,231 (71.2)	15,218 (63.3)	2,016 (13.3)
IEP	901 (28.8)	8,815 (36.7)	1,410 (16.0)
Minnesota			
Race			
Black	300 (14.9)	2,341 (16.1)	259 (11.1)
Latine	402 (19.9)	2,880 (19.9)	201 (7.0)
Other	235 (11.7)	1,796 (12.4)	127 (7.1)
White	757 (37.6)	5,636 (38.9)	980 (17.4)
Gender			
Female	724 (36.0)	4,506 (31.1)	515 (11.4)
Male	1,289 (64.0)	9,994 (68.9)	1,259 (12.6)
IEP Status			
No IEP	1,701 (83.3)	12,063 (83.2)	1,356 (11.2)
IEP	340 (16.7)	2,437 (16.8)	418 (17.2)
Oregon			
Race			
Black	205 (2.2)	925 (2.1)	484 (52.3)
Latine	3,157 (33.7)	16,094 (36.3)	6,920 (43.0)
Other	873 (9.3)	4,266 (9.6)	1,696 (39.8)
White	4,883 (52.2)	22,350 (50.4)	9,196 (41.2)
Gender			
Female	3,308 (35.4)	11,754 (26.5)	5,491 (46.7)
Male	6,037 (64.6)	32,551 (73.5)	13,303 (40.9)
IEP Status			
No IEP	7,744 (82.2)	34,918 (78.8)	15,165 (43.4)
IEP	1,680 (17.8)	9,405 (21.2)	3,636 (38.7)

Note: Counts are shown at the student and referral level by demographic. Exclusion, counted at the referral level, refers to the number of referrals that resulted in exclusionary discipline.

The student intersectional student groups consisted of a mix of student race, gender, and IEP status demographics, such as Black male students with IEPs or Latiné females without IEPs. The testing of disproportionality in exclusionary discipline for the intersection of student demographics using the multi-level regression model was not possible due to low sample size across groups. Instead, the intersectional student groups were examined for disproportionality using RR and RDR₁₀₀ for both receipt of a major referral and exclusion from minor referrals. Table 8 demonstrates exclusionary discipline patterns of minor referrals for the combined student characteristics by state.

The results again found large differences in the use of exclusionary discipline by student group by state. In Oregon, 58.9% of the minor referrals received by Black female students resulted in exclusionary discipline and that percentage was only slightly lower for Black female students with an IEP, 47.4%. Black male students in Oregon also received a high percentage of exclusionary discipline for minor referrals, 55.5% of referrals received by Black male students resulted in exclusion. Georgia had higher rates of exclusion for Latiné male and females with IEPs who received a minor referral, 57.7% and 65.6% respectively. One constant across the data was the lower percentage of minor referrals leading to exclusion in Minnesota and Massachusetts compared to Oregon and Georgia.

Table 8

Examining exclusion rate for the interaction of student characteristics for minor referrals

	Georgia			Massachusetts			Minnesota			Oregon		
	Students (%)	<i>n</i> Referrals (%)	<i>n</i> Exclusions (%)	Students (%)	<i>n</i> Referrals (%)	<i>n</i> Exclusion (%)	Students (%)	<i>n</i> Referrals (%)	<i>n</i> Exclusion (%)	Students (%)	<i>n</i> Referrals (%)	<i>n</i> Exclusion (%)
Black Female IEP	240 (2.4)	849 (3.3)	323 (38.0)	22 (0.8)	159 (0.7)	50 (31.5)	16 (0.9)	66 (0.5)	14 (21.2)	11 (0.1)	38 (0.1)	18 (47.4)
Black Female None	2042 (20.3)	5121 (19.7)	2295 (44.8)	23 (0.8)	93 (0.4)	13 (14.0)	102 (5.9)	649 (5.1)	66 (10.2)	64 (0.7)	219 (0.5)	129 (58.9)
Black Male IEP	658 (6.5)	2252 (8.7)	927 (41.2)	48 (1.7)	640 (2.8)	118 (18.4)	60 (3.5)	549 (4.3)	78 (14.2)	33 (0.4)	212 (0.5)	84 (39.6)
Black Male None	2695 (26.8)	7267 (27.9)	3622 (49.8)	51 (1.8)	386 (1.7)	61 (15.8)	129 (7.5)	1077 (8.5)	101 (9.4)	99 (1.1)	456 (1.1)	253 (55.5)
Latine Female IEP	41 (0.4)	90 (0.4)	59 (65.6)	97 (3.3)	1219 (5.3)	189 (15.5)	27 (1.6)	106 (0.8)	10 (9.4)	146 (1.6)	555 (1.3)	243 (43.8)
Latine Female None	351 (3.5)	716 (2.8)	301 (42.0)	251 (8.6)	2563 (11.1)	136 (5.3)	154 (8.9)	1189 (9.4)	72 (6.1)	1062 (11.6)	3893 (8.9)	1954 (50.2)
Latine Male IEP	133 (1.3)	352 (1.4)	203 (57.7)	202 (7.0)	2607 (11.3)	334 (12.8)	38 (2.2)	225 (1.8)	21 (9.3)	372 (4.1)	2492 (5.7)	898 (36.0)
Latine Male None	689 (6.8)	1744 (6.7)	777 (44.6)	289 (9.9)	3760 (16.2)	241 (6.4)	188 (10.9)	1360 (10.8)	98 (7.2)	1604 (17.5)	9151 (21.0)	3824 (41.8)
Other Female IEP	53 (0.5)	107 (0.4)	78 (72.9)	13 (0.5)	91 (0.4)	11 (12.1)	11 (0.6)	88 (0.7)	9 (10.2)	46 (0.5)	190 (0.4)	42 (22.1)
Other Female None	345 (3.4)	713 (2.7)	356 (49.9)	40 (1.4)	246 (1.1)	26 (10.6)	73 (4.2)	331 (2.6)	16 (4.8)	282 (3.1)	1101 (2.5)	545 (49.5)
Other Male IEP	182 (1.8)	553 (2.1)	293 (53.0)	26 (0.9)	332 (1.4)	71 (21.4)	35 (2.0)	249 (2.0)	48 (19.3)	107 (1.2)	645 (1.5)	277 (43.0)
Other Male None	693 (6.9)	1573 (6.0)	830 (52.8)	65 (2.2)	618 (2.7)	138 (22.3)	122 (7.1)	1128 (8.9)	54 (4.8)	443 (4.8)	2330 (5.3)	832 (35.7)
White Female IEP	80 (0.8)	190 (0.7)	81 (42.6)	128 (4.4)	761 (3.3)	159 (20.9)	33 (1.9)	119 (0.9)	19 (16.0)	223 (2.4)	885 (2.0)	328 (37.1)
White Female None	450 (4.5)	799 (3.1)	362 (45.3)	463 (15.9)	2066 (8.9)	312 (15.1)	216 (12.5)	1489 (11.8)	231 (15.5)	1411 (15.4)	4691 (10.8)	2089 (44.5)
White Male IEP	274 (2.7)	838 (3.2)	428 (51.1)	319 (11.0)	2800 (12.1)	467 (16.7)	120 (7.0)	1035 (8.2)	219 (21.2)	709 (7.7)	4314 (9.9)	1693 (39.2)
White Male None	1141 (11.3)	2885 (11.1)	1381 (47.9)	870 (30.0)	4824 (20.8)	900 (18.7)	398 (23.1)	2993 (23.7)	511 (17.1)	2562 (27.9)	12445 (28.5)	5080 (40.8)

Note: Each row represents a unique Race × Gender × IEP group within a state. Exclusion (%) is based on referral-level exclusion rates. The percentages are calculated from the data after filtering missing student demographic data. Percent of student and referrals refers to percent of minor referral students and referrals, and percent exclusion represents the percent of minor referrals for that group leading to exclusionary discipline. Comparing across groups shows which student groups may receive exclusionary discipline at a higher rate. Percentages may not sum to 100 due to rounding.

RQ3: Differences in Major and Minor Referral Patterns

Although the primary focus of the study was minor referrals, comparisons across referral types revealed meaningful differences in referral categorization by student demographics. In an examination of challenging behavior for major and minor referrals, defiance and disruption were found to be the most prevalent labels across referral type. Though defiance and disruption were both a greater share of the challenging behaviors for minor referrals, 29% and 23% respectively than for major referrals, 17% and 16% each. These differences in frequency correspond with a longer list of challenging behaviors for major referrals compared to minor referrals and an increased use of physical aggression and fighting for major referrals compared to minor referrals. Minor referrals had a higher percentage of referrals with subjective behaviors compared to major referrals. A full breakdown of the frequency and percentage of behavior labels across major and minor referrals and by state can be found in Table 3.

Overview of Differences in Exclusionary Discipline

A review of exclusionary discipline practices between major and minor referrals data suggested differences in rate of exclusion by race and IEP status. Broadly, Black students received exclusionary discipline for 58% of their major referrals and 40% of their minor referrals compared to 56% of major referrals for White students and 33% of minor referrals for White students. Latiné and Other students also had lower rates of exclusionary discipline for minor referrals compared to Black students. Students with IEPs had lower rates of exclusionary discipline compared to students without IEPs for both major and minor referrals. While there were differences in rate of exclusion by race and IEP status, there was no difference between male and female students. Table 9 contains the counts and percentages of exclusionary discipline referrals by student demographics.

Table 9

Exclusionary Responses by Student Group for Major and Minor Referrals

	Major Referrals			Minor Referrals		
	Non-Exclusion	Exclusion	Neither	Non-Exclusion	Exclusion	Neither
Race/Ethnicity						
Black	13040 (23.1%)	32778 (58.1%)	10578 (18.8%)	8305 (41.5%)	8152 (40.7%)	3576 (17.9%)
Latine	9719 (29.5%)	19761 (59.9%)	3515 (10.7%)	18285 (57.1%)	9361 (29.2%)	4379 (13.7%)
Other	5221 (32.1%)	8389 (51.5%)	2664 (16.4%)	4794 (46.6%)	3626 (35.2%)	1875 (18.2%)
White	14981 (31.2%)	27143 (56.5%)	5875 (12.2%)	23275 (53.9%)	14268 (33.1%)	5609 (13%)
N/A	1324 (17.8%)	5015 (67.3%)	1112 (14.9%)	2620 (48.9%)	1414 (26.4%)	1325 (24.7%)
Gender						
Female	15223 (27.8%)	31297 (57.1%)	8303 (15.1%)	16898 (51.4%)	10952 (33.3%)	4998 (15.2%)
Male	29024 (27.3%)	61745 (58.1%)	15434 (14.5%)	40369 (51.8%)	25860 (33.2%)	11766 (15.1%)
Other	38 (42.7%)	44 (49.4%)	7 (7.9%)	12 (57.1%)	9 (42.9%)	0 (0%)
IEP Status						
No IEP	30736 (25.6%)	71873 (59.9%)	17419 (14.5%)	43115 (50.8%)	28934 (34.1%)	12783 (15.1%)
IEP	13549 (33%)	21213 (51.6%)	6325 (15.4%)	14164 (54.4%)	7887 (30.3%)	3981 (15.3%)

Note: Percentages were calculated based on student category, i.e. race/ethnicity, gender, IEP status by referral type.

State Differences in Exclusionary Discipline

Due to differences in student demographics and referral patterns across states, exclusionary discipline for major and minor referrals were examined by student race within each state. In Georgia, exclusionary responses were common across all racial groups for both major and minor referrals. For major referrals in Georgia, 70.4% of Black student referrals and 81.4% of White student referrals resulted in exclusion. For minor referrals in Georgia, 46.3% of Black student referrals resulted in exclusion compared to 47.8% of White student referrals, indicating relatively similar patterns across groups, but an elevated rate of exclusionary discipline in the state. In Massachusetts, exclusionary responses were less frequent overall, and racial differences were more pronounced. While 18.9% of Black student minor referrals resulted in exclusion,

Latíné students had a much lower rate of 8.9%, and White students had a rate of 17.6%. For major referrals, exclusionary discipline was more common among Latíné students, 54.7%, compared to Black, 30%, or White students, 32.6%, though the overall rates were still lower than in Georgia. In Minnesota, referrals for Black students were less likely to result in exclusionary discipline. For major referrals, 24.4% of Black student referrals resulted in exclusion compared to 39.3% for White students, despite Black students receiving major referrals at a disproportionate rate in the state. For minor referrals exclusionary discipline occurred for only 11.1% of Black student referrals, lower than the 17.4% for White students.

In Oregon, the exclusionary discipline rates for both major and minor referrals were higher for Black students than for other groups. For minor referrals, 52.3% of Black student referrals resulted in exclusion compared to 43% for Latíné and 41.1% for White students. For major referrals, 74.7% of Black student referrals resulted in exclusion compared to 64.7% for Latíné and 60.4% for White students, indicating a consistently higher exclusionary discipline rate for Black students in Oregon. These findings highlight that while some states exhibit more consistent exclusion rates across racial groups, others, particularly Massachusetts and Oregon, show notable racial variation in exclusionary discipline for both major and minor referrals. A full table of exclusion rates by student characteristics can be found on Table 10.

Table 10

Exclusion rates of major and minor referrals by student demographics by state

	Major Referrals		Minor Referrals	
	Non-Exclusion	Exclusion	Non-Exclusion	Exclusion
Georgia				
Black	3523 (9%)	27542 (70.4%)	5286 (34.1%)	7167 (46.3%)
Latine	425 (9%)	3677 (78.2%)	981 (33.8%)	1340 (46.2%)
Other	206 (4%)	3546 (69.2%)	393 (13.3%)	1557 (52.9%)
White	962 (9.4%)	8295 (81.4%)	1459 (31%)	2252 (47.8%)
Female	2044 (9.3%)	15851 (72.1%)	3221 (34.9%)	4017 (43.6%)
Male	3585 (8.7%)	30423 (73.9%)	5927 (31.5%)	8803 (46.8%)
No IEP	4505 (8.8%)	37839 (73.9%)	7364 (32.5%)	10397 (45.9%)
IEP	1124 (9.4%)	8435 (70.7%)	1784 (33.2%)	2423 (45.1%)

Table 10 continued

	Non-Exclusion	Exclusion	Non-Exclusion	Exclusion
Massachusetts				
Black	802 (43.9%)	549 (30%)	842 (65.9%)	242 (18.9%)
Latine	3002 (30.7%)	5348 (54.7%)	8416 (82.9%)	900 (8.9%)
Other	517 (37.3%)	581 (41.9%)	864 (67.1%)	246 (19.1%)
White	4959 (43.3%)	3729 (32.6%)	7139 (68.3%)	1840 (17.6%)
Female	3017 (36.5%)	3616 (43.8%)	5609 (76.1%)	929 (12.6%)
Male	6587 (38.2%)	6802 (39.4%)	12126 (72.8%)	2495 (15%)
No IEP	4902 (33.6%)	6325 (43.4%)	11366 (74.7%)	2016 (13.2%)
IEP	4725 (43.1%)	4095 (37.3%)	6370 (72.3%)	1410 (16%)
Minnesota				
Black	8356 (61.4%)	3329 (24.4%)	1796 (76.7%)	259 (11.1%)
Latine	1941 (57.2%)	955 (28.1%)	2313 (80.3%)	201 (7%)
Other	2943 (64.2%)	1024 (22.3%)	1374 (76.5%)	127 (7.1%)
White	1754 (45.6%)	1512 (39.3%)	3627 (64.4%)	980 (17.4%)
Female	6160 (58.4%)	2719 (25.8%)	3224 (71.5%)	515 (11.4%)
Male	9204 (58.3%)	4362 (27.6%)	6844 (68.5%)	1259 (12.6%)
No IEP	11383 (58.1%)	5071 (25.9%)	8475 (70.3%)	1356 (11.2%)
IEP	3981 (58.9%)	2010 (29.7%)	1593 (65.4%)	418 (17.2%)
Oregon				
Black	359 (19.7%)	1358 (74.7%)	381 (41.2%)	484 (52.3%)
Latine	4351 (28.8%)	9781 (64.7%)	6575 (40.9%)	6920 (43%)
Other	1555 (30%)	3238 (62.5%)	2163 (50.7%)	1696 (39.8%)
White	7306 (32.4%)	13607 (60.4%)	11050 (49.4%)	9196 (41.1%)
Female	4002 (28.5%)	9111 (65%)	4844 (41.2%)	5491 (46.7%)
Male	9648 (30.1%)	20158 (63%)	15472 (47.5%)	13303 (40.9%)
No IEP	9946 (28.7%)	22638 (65.3%)	15910 (45.6%)	15165 (43.4%)
IEP	3719 (32.5%)	6673 (58.4%)	4417 (47%)	3636 (38.7%)

Note: The percentages are calculated included referrals that were categorized as neither exclusionary nor non-exclusionary leading to percentages that do not sum to 100%. In addition, percentages are calculated by student race, gender, and IEP status separately.

Student Demographic Differences in Receiving a Major Referral

As the data suggested differences in rates of exclusion based on student characteristics with higher rates of exclusionary discipline for major referrals compared to minor referrals, risk ratios (RR) and raw differential representation (RDR) were calculated using White students, male students, and students without IEPs as reference groups to quantify disproportionality in the risk of receiving a major referral compared to a minor referral. The risk of receiving a major referral, a minor referral, and their ratio were calculated to examine over-representation of certain groups in receipt of a major referral as well as potential under-representation for minor referrals. The results show Black students were over-represented in the receipt of a major referral

and in their ratio of major to minor referrals compared to other students. The complete results by race, IEP status, and gender are available in Table 11.

Table 11

Combined overall RR and RDR by student demographic for receipt of major or minor referral

	Major Risk	Major RR	Major RDR ₁₀₀	Minor Risk	Minor RR	Major vs Minor Risk	Major: Minor RR
Race							
Black	0.85	1.19	13.6	0.38	0.65	2.22	1.83
Latine	0.75	1.03	2.5	0.61	1.03	1.21	1.00
Other	0.77	1.07	5.2	0.50	0.84	1.55	1.28
White	0.72	1.00	0	0.59	1.00	1.21	1.00
Gender							
Female	0.77	0.99	-1.20	0.49	0.91	1.58	1.08
Male	0.78	1.00	0	0.53	1.00	1.46	1.00
IEP Status							
No IEP	0.76	1.00	0	0.51	1.00	1.49	1.00
IEP	0.81	1.07	5	0.53	1.03	1.54	1.03

Note: Risk ratios are calculated at the student level.

Across all students with behavior referrals, Black students were 19% more likely to receive a major referral than White students (RR = 1.19) and had a raw differential representation of 13.6 per 100 students, meaning that approximately 14 more Black students per 100 were given a major referral than would be expected if referrals were assigned at the same rate as White students. Students in the “Other” racial category also experienced elevated risk (RR = 1.07; RDR₁₀₀ = 5.2), while Latiné students had a comparable referral rate to White students (RR = 1.03; RDR₁₀₀ = 2.5). In addition, Black students were roughly 83% more likely than White students to receive a major referral versus a minor referral compared to White

students and other race students were 28% more likely to receive a major referral. The results of the major to minor risk comparison show striking differences by race.

In contrast, gender differences in receipt of major referrals were minimal. Female students had a slightly lower likelihood of receiving a major referral compared to male students ($RR = 0.99$; $RDR_{100} = -1.2$), suggesting no meaningful difference. Similarly, students with IEPs had a modestly increased risk of receiving a major referral ($RR = 1.07$; $RDR_{100} = 5$), indicating that five more students per 100 with IEPs received major referrals than expected based on the rate for peers without IEPs. These findings suggest that racial disproportionality in assignment of major referrals remains a key equity concern, particularly for Black students, while disproportionality by gender and disability status is less pronounced when examined at the student level.

State Differences in Receipt of a Major Referral

The data analysis had already shown Georgia schools to use exclusionary discipline at a higher rate than other states and that Georgia contained the majority of Black students in the sample, but analysis was conducted to examine which states were driving the disproportionality in major referrals. State specific analysis of risk ratios and raw differential representation (RDR) revealed that racial disparities in the likelihood of receiving a major referral varied substantially by state. Across all four states, Black students were more likely to receive a major referral than their White peers, but the magnitude of disparity differed by context.

In Minnesota, the racial disparity was most pronounced, with Black students having a risk ratio (RR) of 1.46 and an RDR_{100} of 29.1. This indicated that nearly 29 more Black students per 100 received a major referral than would be expected put another way, Black students were 46% more likely to receive a major referral in Minnesota. Latiné and other race students in

Minnesota also exhibited elevated risk, though to a lesser extent, $RR = 1.13$ and $RR = 1.08$, respectively. The ratio of major to minor referral risk was 3.74, meaning Black students in Minnesota were over 300% more likely to receive a major referral than a minor referral. Students in the other race category in Minnesota were also at increased risk of major referral with a 275% increase.

In Massachusetts, Black students again had the highest risk ($RR = 1.20$; $RDR_{100} = 14.0$), followed closely by Latiné ($RR = 1.19$; $RDR_{100} = 13.3$) and other race students ($RR = 1.16$; $RDR_{100} = 11.0$). When examining the comparative major to minor referral risk ratio against White students, Black, Latiné, and other race students were all at elevated risk of major referral compared to minors. These findings suggest that disproportionality in receipt of a major referral extended beyond Black students.

In Oregon, Black students also had an elevated risk of major referral ($RR = 1.19$; $RDR_{100} = 13.4$), with students missing race data and those in the other race category showing moderate disparities ($RDR_{100} = 11.0$ for both groups). Latiné students had a smaller but still present disparity in receipt of major referral ($RR = 1.07$; $RDR_{100} = 4.4$). Similar to the other states, all of the racial groups had a higher risk ratio of major referral to minor referral compared to White students.

In Georgia, the risk ratio for Black students was the lowest among the four states ($RR = 1.08$), though the RDR_{100} was still 6.4, indicating modest disproportionality in receiving a major referral. In contrast, Latiné and other race students in Georgia had a lower likelihood of receiving major referrals than White students ($RR = 0.93$ and 0.96 , respectively), with corresponding negative RDR_{100} values. Georgia also had the smallest difference in major to minor risk ratio between racial groups with Black and White students at equal risk. This suggests that although

Georgia used major referrals and exclusion at a higher rate, they were implementing major referrals equally. This suggests the other states in the sample were the main drivers for over-representation of receipt of a major referral for Black students.

These results highlight the importance of examining disproportionality within state contexts, as patterns of disproportionality can vary in magnitude and in direction across demographic groups. The complete list of RDR₁₀₀ and RR by race per state can be found on Table 12 and visualization of risk ratios by state can be seen in Figure 7.

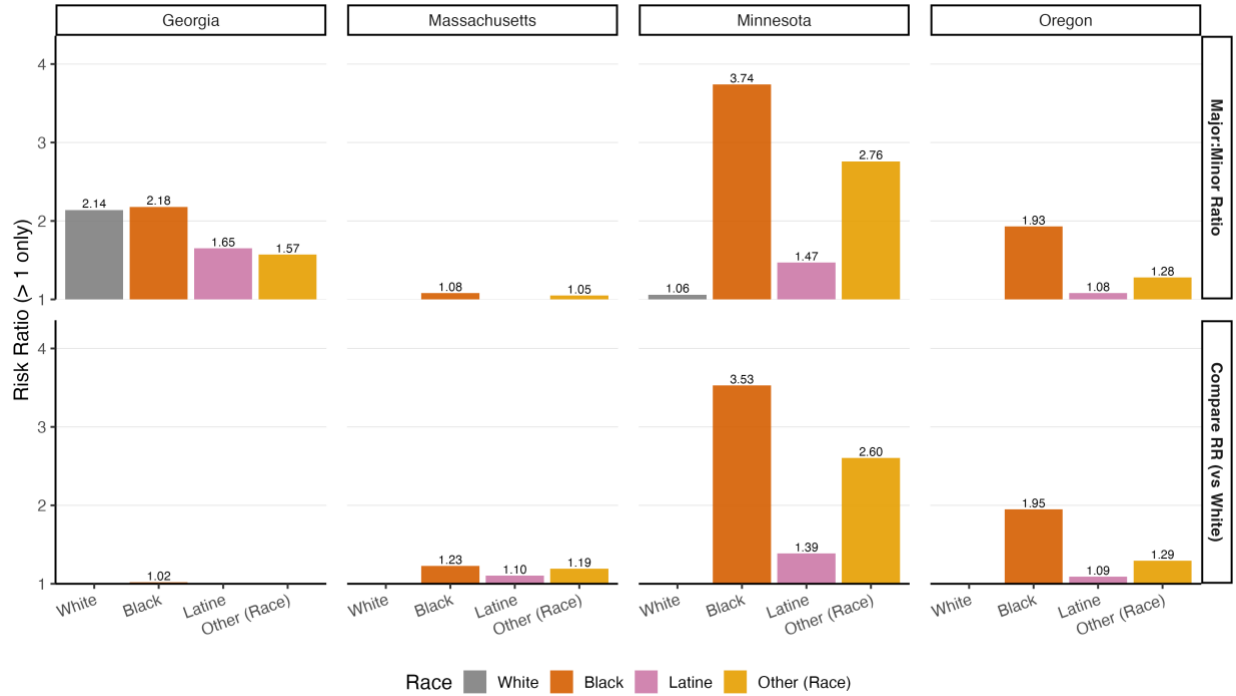
Table 12

Risk Ratio and Raw Differential Representation (RDR) by Race and State for Risk of Major or Minor Referral

	Major Referral Students	Total Students	Major Risk (%)	Major RR	Major RDR	RDR ₁₀₀	Minor RR	Major: Minor Ratio	Major: Minor RR
Georgia									
Black	12,276	14,445	84.98	1.08	924.5	6.4	1.06	2.18	1.02
White	4,148	5,280	78.56	1.00	0.0	0.0	1.00	2.14	1.00
Other	1,992	2,655	75.03	0.96	-92.9	-3.5	1.30	1.57	0.73
Latine	2,004	2,736	73.25	0.93	-145.0	-5.3	1.21	1.65	0.77
Massachusetts									
Black	149	179	83.24	1.20	25.1	14.0	0.98	1.08	1.23
Latine	792	959	82.59	1.19	127.5	13.3	1.08	0.97	1.10
Other	147	183	80.33	1.16	20.1	11.0	0.97	1.05	1.19
White	1,529	2,207	69.28	1.00	0.0	0.0	1.00	0.88	1.00
Minnesota									
Black	1,121	1,222	91.73%	1.46	355.6	29.1	0.42	3.74	3.53
Other	649	786	82.57%	1.32	156.4	19.9	0.51	2.76	2.60
Latine	591	837	70.61%	1.13	67.0	8.0	0.81	1.47	1.39
White	803	1,282	62.64%	1.00	0.0	0.0	1.00	1.06	1.00
Oregon									
Black	395	477	82.81%	1.19	63.9	13.4	0.61	1.93	1.95
Other	1,118	1,440	77.64%	1.12	118.1	8.2	0.86	1.28	1.29
Latine	3,400	4,583	74.19%	1.07	215.4	4.7	0.98	1.08	1.09
White	4,825	6,948	69.44%	1.00	0.0	0.0	1.00	0.99	1.00

Note: Based on a sample of students who received at least one behavior referral.

Figure 7
Ratio of Major to Minor Referral Risk by Race and State



Note: Major: Minor Ratio is calculated using major and minor referral risk by group meaning the ratio is the within-group likelihood of receiving a major referral whereas the comparison risk is the between-group risk using White students as the reference group. The comparison shows how much more likely a group is to receive a major referral versus a minor referral compared to White students.

Intersectionality by Race, Gender, and IEP Status

Finally, an analysis of risk ratio and RDR for the interaction between student demographics, for example. Black male students with IEPs, was calculated. The results suggested increased RR and RDR for Black students overall, but particularly for those with IEPs. Black males with IEPs, (RR = 1.25, RDR₁₀₀ = 17.6), and Black females with IEPs, (RR = 1.23, RDR₁₀₀ = 16.1) were at the highest risk of receiving a major referral. The full interaction list can be found on Table 13.

Table 13

Risk Ratio and Raw Differential Representation by Race, Gender, and IEP Status Interaction for Major Referrals

	Students with Major Referrals	Total Students	Group Risk	Risk Ratio	RDR (Raw)	RDR ₁₀₀
Black Female IEP	734	841	0.87	1.23	135.40	16.1
Black Female None	5,088	6,028	0.84	1.19	795.70	13.2
Black Male IEP	1,835	2,066	0.89	1.25	363.62	17.6
Black Male None	6,339	7,470	0.85	1.19	1,023.39	13.7
Latine Female IEP	332	456	0.73	1.02	7.30	1.6
Latine Female None	2,268	3,127	0.73	1.02	43.78	1.4
Latine Male IEP	902	1,137	0.79	1.11	93.23	8.2
Latine Male None	3,355	4,496	0.75	1.05	157.36	3.5
Other Female IEP	221	273	0.81	1.14	26.75	9.8
Other Female None	1,144	1,534	0.75	1.05	52.16	3.4
Other Male IEP	560	684	0.82	1.15	73.19	10.7
Other Male None	2,004	2,608	0.77	1.08	148.66	5.7
White Female IEP	577	787	0.73	1.03	16.53	2.1
White Female None	3,165	4,534	0.70	0.98	-63.48	-1.4
White Male IEP	1,800	2,301	0.78	1.10	163.37	7.1
White Male None	5,880	8,262	0.71	1.00	0.00	0.0

Note: Risk is calculated relative to White, Male, Students without IEPs.

Overall, the results indicated Black students were consistently at elevated risk of receiving a major referral compared to White students, a pattern that held across intersections with gender and IEP status. Given that major referrals were more likely to result in exclusionary discipline, this disparity has important implications for interpreting disproportionality in the minor referral sample. Specifically, if Black students are disproportionately assigned major referrals, they may be underrepresented in the minor referral dataset—even though they may continue to experience disproportionate exclusionary disciplinary consequences for minor referrals. To address this, the next phase of the analysis focused on exclusionary responses to minor referrals. Risk ratios (RR) and raw differential representation per 100 students (RDR₁₀₀) were calculated to quantify disparities in exclusionary discipline outcomes among students who received at least one minor referral. The descriptive analyses were followed by multilevel

logistic modeling to test whether exclusionary discipline was predicted by student and school characteristics, accounting for nested data structures with random intercepts for both student and school and a fixed effect for state.

RQ4: Differences in Exclusionary Discipline for Minor Referrals

The primary research question aimed to determine the presence of disproportionality in exclusionary discipline for minor referrals by student race, gender, and IEP status. The data showed that roughly 40% of minor referrals resulted in exclusionary discipline. The risk of ever receiving exclusionary discipline varied by student race, gender, and IEP status. Risk ratios were calculated at the student level both aggregated across states and by state. An analysis of exclusionary discipline responses to minor referrals aggregated across states revealed Black students were approximately 22% more likely than White students to receive at least one exclusionary consequence. This equates to a risk ratio (RR) of 1.22 and a raw differential representation per 100 students (RDR_{100}) of 11.5 students. Students in the “other” racial category—which included Native, Asian, Multi-Racial, and Pacific Islander students—were about 10% more likely to receive exclusionary discipline ($RR = 1.10$; $RDR_{100} = 5.3$) across all states. In contrast, Latiné students were slightly less likely to receive exclusionary discipline than White students ($RR = 0.96$; $RDR_{100} = -2.0$).

Gender disparities were also evident in the receipt of minor referrals. Female students were about 10% less likely than male students to ever receive exclusionary discipline in response to a minor referral ($RR = 0.90$; $RDR_{100} = -5.5$). While students with IEPs were marginally more likely to receive exclusionary discipline than students without IEPs ($RR = 1.05$, $RDR_{100} = 3.0$), this difference corresponds to 3 additional students with IEPs excluded. The full table of RR and RDR by student demographic can be found in Table 14.

Table 14
Disparity in Exclusionary Discipline for Minor Referrals by Race, Gender, and IEP Status

	Total Students	Excluded Students	Group Risk (%)	Risk Ratio	RDR	RDR ₁₀₀
Race						
Black	6,263	4,076	65.08	1.22	720.24	11.50
Latine	5,586	2,877	51.50	0.96	-111.72	-2.00
Other	2,516	1,481	58.86	1.10	133.35	5.30
White	9,320	4,990	53.54	1.00	0.00	0.00
N/A	1,693	642	37.92	0.71	-264.11	-15.60
Gender						
Female	9,013	4,686	51.99	0.90	-495.71	-5.50
Male	16,296	9,369	57.49	1.00	0.00	0.00
Other	12	4	33.33	0.58	-2.90	-24.17
IEP Status						
No IEP	20,896	11,458	54.83	1.00	0.00	0.00
IEP	4,625	2,675	57.84	1.05	138.75	3.00

Note: The results are aggregated across states and measured using Risk Ratio and Raw Differential Representation. The table shows minor referrals and exclusion for minor referrals using student characteristics. The reference group is White, Male, and Non-IEP students.

Exclusionary Discipline for Minor Referrals by State

An analysis of exclusionary discipline by state revealed notable differences in risk of exclusion by race, gender, and IEP status. Black students were more likely to receive at least one exclusionary consequence for a minor referral in Georgia, Minnesota, and Oregon, with risk ratios of 1.09 – 1.11 across these three states and RDR₁₀₀ values ranging from 4 to 7. In contrast, in Massachusetts, Black students had a slightly lower risk (RR = 0.89, RDR₁₀₀ = -4). Latine students were consistently less likely to receive exclusionary discipline for minor referrals compared to White students across all states, with risk ratios ranging from RR = 0.59 to RR = 0.99. Students classified in the other racial category had variable risk of exclusion; for instance, students in the other group in Georgia showed elevated risk (RR = 1.11), while those in Minnesota had substantially lower risk (RR = 0.60, RDR₁₀₀ = -15), though these differences may be due to differences in composition of the other race group by state. Across all four states,

female students were consistently at a lower risk of receiving exclusionary discipline for a minor referral. Risk ratios for female students ranged from $RR = 0.74$ in Massachusetts to $RR = 0.93$ in Georgia and Oregon, indicating a consistently protective effect against exclusion.

Students with IEPs, in contrast, had elevated risk of exclusionary discipline in every state, with the highest disparity observed in Minnesota ($RR = 1.34$, $RDR_{100} = 11$), suggesting that students with disabilities were more likely to receive at least one exclusionary response to minor referrals. Although students with IEPs were more likely to have at least one exclusionary consequence, their individual referrals were less likely to result in exclusionary discipline when compared to students without IEPs. On average, across states, referrals for students with IEPs were approximately 11% less likely to lead to exclusion. This pattern suggests that students with IEPs may be more frequently referred for behavior and that those referrals may be treated less punitively on a per-referral basis.

State-level differences in exclusionary discipline per referral by race also emerged. In Georgia and Minnesota, although Black and Latiné students were more likely to have at least one exclusionary referral, their individual referrals were not more likely to result in exclusion than those for White students. This pattern may be driven by higher referral rates overall for students, states, or school-level practices that generate more referrals for specific student groups. In contrast, in Massachusetts, Black students were not more likely to receive at least one exclusionary referral, but their individual referrals were more likely to result in exclusion—suggesting different local referral and discipline patterns. Oregon exhibited the most pronounced disparity, where referrals for Black students were 27% more likely to lead to exclusionary discipline than those for White students, pointing to substantial disproportionality in the disciplinary response to minor referrals in the state. The combined demographic risk ratio and

RDR₁₀₀ figures for exclusionary discipline for minor referrals can be found on Table 15. In addition, Figure 7 shows the Risk Ratios of receiving a major referral, exclusionary discipline for major referrals as well as for minor referrals by student demographic and state.

Table 15

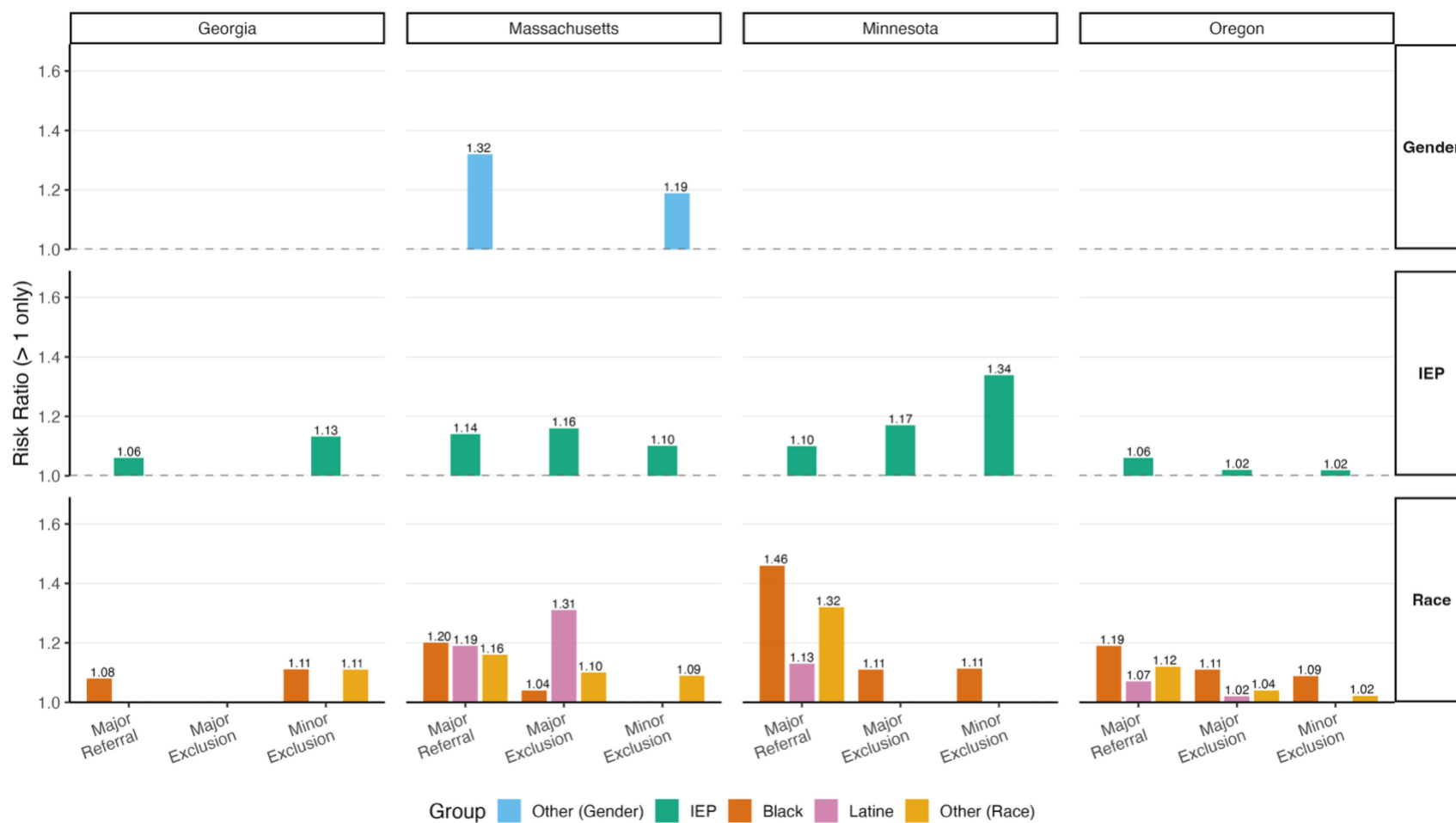
Disproportionality in Exclusionary Discipline for Minor Referrals by State

	Excluded Students	Minor Referral Students	Group Risk (%)	Risk Ratio	RDR (Raw Count)	RDR ₁₀₀
Georgia						
Black	3,770	5,620	67.1	1.11	393.4	7
Other	850	1,268	67.0	1.11	88.8	7
Latine	697	1,212	57.5	0.95	-36.4	-3
Female	2307	3892	59.3	0.93	-194.6	-5
IEP	1176	1704	69.0	1.13	136.3	8
Massachusetts						
Black	50	138	36.2	0.89	-5.5	-4
Latine	292	815	35.8	0.88	-40.8	-5
Other	62	140	44.3	1.09	5.6	4
Female	341	1089	31.3	0.74	-119.8	-11
IEP	365	901	40.5	1.10	36.0	4
Minnesota						
Black	127	300	42.3	1.11	12.0	4
Other	91	402	22.6	0.60	-60.3	-15
Latine	53	235	22.6	0.59	-35.3	-15
Female	208	724	28.7	0.81	-50.7	-7
IEP	142	340	41.8	1.34	37.4	11
Oregon						
Black	129	205	62.9	1.09	10.3	5
Latine	1,797	3,157	56.9	0.99	-31.6	-1
Other	516	873	59.1	1.02	8.7	1
Female	1,830	3,308	55.3	0.93	-132.3	-4
IEP	992	1,680	59.0	1.02	16.8	1

Note: RR and RDR calculated using student data by state with White, male students without IEPs serving as the reference group.

Figure 8

Risk Ratios for Major Referral and Major and Minor Referral Exclusion by State and Student Group



Intersectional Differences in Exclusionary Discipline

The intersection of race, gender, and IEP status continued to demonstrate disparities in the likelihood of exclusionary discipline following a minor referral. Specifically, Black male and female students without IEPs were 23% more likely to have a referral result in exclusionary discipline compared to White male students without IEPs, the reference group. In contrast, female Latiné students with IEPs had the lowest risk of exclusionary discipline, approximately 21% less likely to receive an exclusionary response to a referral. The full table of RR and RDR₁₀₀ for intersectional student demographic groups can be found in Table 16. These findings underscore the compounded risk associated with multiple marginalized identities, particularly for Black students. However, it is important to interpret these results within the broader context of state-level disciplinary practices. Georgia and Oregon were more likely to use exclusionary discipline overall, and Georgia accounted for the majority of Black student referrals in the dataset. As a result, state-specific disciplinary norms and the demographic composition of the sample may influence the observed disparities in risk ratios.

Table 16

Risk Ratios & RDR for Exclusion for Minor Referrals by Race, Gender, and IEP Status

	Excluded Students	Total Students	Group Risk (%)	Risk Ratio	RDR	RDR ₁₀₀
Black Female IEP	194	289	67.1	1.23	36.1	12.5
Black Female None	1,382	2,231	61.9	1.13	162.9	7.3
Black Male IEP	505	799	63.2	1.16	67.9	8.5
Black Male None	2,006	2,974	67.5	1.23	380.7	12.8
Latine Female IEP	135	311	43.4	0.79	-34.8	-11.2
Latine Female None	854	1,818	47.0	0.86	-140.0	-7.7
Latine Male IEP	414	745	55.6	1.02	6.7	0.9
Latine Male None	1,494	2,770	53.9	0.99	-19.4	-0.7
Other Female IEP	71	123	57.7	1.06	3.8	3.1
Other Female None	402	740	54.3	0.99	-2.2	-0.3
Other Male IEP	219	350	62.6	1.14	27.7	7.9
Other Male None	796	1,323	60.2	1.10	72.8	5.5
White Female IEP	243	464	52.4	0.96	-10.7	-2.3

Table 16 continued							
White Female None	1,210	2,540	47.6	0.87	-177.8	-7.0	
White Male IEP	847	1,422	59.6	1.09	69.7	4.9	
White Male None	2,717	4,971	54.7	1.00	0.00	0.0	
<i>Note:</i> Risk Ratios and RDR are calculated using white male students without IEPs as the reference group. The data contains minor referral observations with complete race and ethnicity data as well as removes the other group for gender due to small sample size.							

State Level Differences in Odds of Exclusionary Discipline

In order to test the differences in rates of exclusionary discipline, a binary multi-level regression model was fit to the minor referral data. The final analytic sample for the multi-level model examining exclusion from minor referrals included 100 schools, 18,977 students, and 74,059 minor referral observations. An iterative model-building process was employed to identify the best-fitting model, using AIC, BIC, and likelihood ratio tests for nested models. Initial models that included all four states revealed that Georgia exhibited substantially higher exclusionary discipline rates compared to the other states. Referrals for students in Georgia were over 30 times more likely to result in exclusionary discipline than similar referrals in other states (OR = 31.86, $p < .001$) though this dropped to (OR = 2.77, $p = 0.05$) when removing schools that only used exclusionary discipline. An exploratory model comparing state-level differences in exclusion also showed that students in Massachusetts were significantly less likely to receive exclusionary discipline compared to those in Oregon (OR = 0.20, $p = .001$).

To further investigate this disparity in exclusionary discipline, school-level patterns were examined. Out of an initial 131 schools, 26 schools were found to be either 100% exclusion schools or 0% exclusion schools and an additional 5 had fewer than 10 students and 20 referrals. These schools were omitted from the sample to improve model fit. In Georgia, 22 out of 60 schools (37%) used only exclusionary discipline in response to minor referrals. In Oregon only 1 out of 45 were full exclusion schools. Both states also had zero exclusion schools, 2 in Georgia

and 1 in Oregon. These stark differences in school discipline practices prompted a modeling strategy that grouped Georgia separately from the other three states. The resulting model, comparing Georgia to all other states combined, provided the best overall fit and is reported as the primary model in this analysis.

Student Demographic Differences in Odds of Exclusionary Discipline

The main model consisted of a reference group of White Male students without IEPs in Georgia, with race, gender, IEP status, and receipt of a prior referral as student level predictors and percent Black student enrollment, location type (i.e. urban, suburban, or rural) as school level predictors. Random intercepts were specified at the student and school levels, with fixed effects included for state, student characteristics, and school-level predictors. Race and gender were not significant predictors of an exclusionary discipline outcome. Referrals for Black students in Georgia did not differ significantly from referrals for White students in Georgia in the odds of receiving exclusionary discipline (OR = 01.00, $p = 0.98$), nor did Latiné students (OR = 1.05, $p = 0.18$) or students categorized as other race (OR = 0.92, $p = 0.10$).

Referrals for female students also did not differ significantly from males for the odds of exclusionary discipline (OR = 0.99, $p = 0.75$). However, referrals for students with IEPs had significantly lower odds of exclusion compared to their non-IEP peers (OR = 0.84, $p < .001$), suggesting reduced likelihood of exclusionary discipline in referrals for students receiving special education services. Referrals for students with a prior referral were significantly more likely to result in exclusionary discipline (OR = 1.44, $p < .001$), indicating a cumulative discipline risk where a student is more likely to be excluded after receiving multiple referrals. School-level centered percent Black enrollment was found to be slightly indicative of lower exclusion risk (OR = 0.98, $p = 0.02$) showing that referrals in schools with higher Black

enrollment were less likely to result in exclusion. Urbanicity, where the school was located, was not a significant predictor of exclusionary discipline, though the confidence intervals and standard errors showed large variability in practices within the groups of urban, rural, and suburban locale. Compared to Georgia, referrals for students in the combined group of Minnesota, Oregon, and Massachusetts had significantly lower odds of exclusionary discipline (OR = 0.12, $p < .001$), demonstrating state-level variation in discipline practices. The model summary can be found in Table 17 along with the distribution of exclusion rates by school in the state groups in Figure 4.

Table 17
Predicting Exclusion by Race, IEP Status, and Gender with State Fixed Effects

	Exclusionary Discipline for Minor Referrals				
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	β	<i>SE</i>
White Male, No IEP, No Prior Referral, Urban, Georgia	2.77	0.99 – 7.80	0.05	1.02	0.53
Black	1.00	0.90 – 1.12	0.98	0.00	0.06
Latine	1.05	0.98 – 1.12	0.18	0.05	0.04
Other Race	0.92	0.84 – 1.02	0.10	-0.08	0.05
Female	0.99	0.94 – 1.05	0.75	-0.01	0.03
IEP (Yes vs No)	0.84	0.78 – 0.89	<0.001	-0.18	0.03
Prior Referral (Yes vs No)	1.44	1.36 – 1.52	<0.001	0.36	0.03
Black Enrollment % (Centered)	0.98	0.96 – 1.00	0.02	-0.02	0.01
Rural vs Urban	0.78	0.33 – 1.86	0.58	-0.24	0.44
Suburb vs Urban	0.80	0.30 – 2.13	0.65	-0.22	0.50
Other States vs Georgia	0.12	0.004 – 0.095	<0.001	-2.15	0.54
Random Effects					
σ^2_{ϵ} Referral	3.29				
σ^2_{τ} Student	0.47			ICC Student	0.07
σ^2_{μ} School	2.97			ICC School	0.44
ICC Total	0.51				
N Schools	91				
N Students	17218				
Observations	67989				
Marginal R ² / Conditional R ²	0.048 / 0.534				
<i>Note:</i> The Confidence Interval (CI) is set to 95% and shows the Odds Ratios (OR) from the multilevel logistic regression model testing for exclusionary discipline by student and school level predictors.					

The intraclass correlation coefficient (ICC) for the model was 0.51, indicating 51% of the variance in exclusionary discipline outcome was attributable to differences across schools and students while the school clustering ICC was 0.44, representing a much greater share of the explained variance than the student clusters. The marginal $R^2 = 0.048$ and conditional $R^2 = 0.534$ suggest that while fixed effects explained a small portion of the variance, 5%, a substantial amount of variability, 53%, was accounted for by between-school and between-student variability effects in the model. The model suggests that student demographic characteristics and school predictors only account for a small amount of the variance in exclusionary discipline, and differences between individuals as well as between schools account for the majority of the variance in exclusion. In addition, the majority of the variation in random effects stems from school-level differences suggesting that behavior management varies significantly by school. The model results suggest school level interventions, such as PBIS may moderate the effect of student level predictors on odds of exclusion.

RQ5: Moderation of Exclusionary Discipline from PBIS Implementation

To examine whether school-level PBIS implementation moderated the odds of exclusionary discipline for minor referrals, 2 three-level multi-level logistic regression models were estimated using a centered, continuous Tier 1 and Tier 2 fidelity of implementation scores (TFI). The mean Tier 1 score was about 24 out of 30 while the mean Tier 2 score was 16 out of 26. Both models accounted for student-level and school-level clustering using random intercepts and included a fixed effect for state—Georgia vs. Other States—to address policy and reporting differences. The overall distribution of school-level fidelity scores for Tier 1 and Tier 2 can be observed in Figure 2 and the score distribution by state can be seen in Figure 3.

The Tier 1 model demonstrated substantial between-school variance in the odds of exclusionary discipline. The final analytic sample in the model for Tier 1 included 9,312 students and 34,656 minor referrals from 49 schools, with Georgia representing 60% of schools. The intraclass correlation coefficient (ICC) was 0.56, indicating 56% of the variance in exclusionary discipline outcome was attributable to differences between school and student clustering. The model's marginal $R^2 = 0.088$, suggested fixed effects alone explained only 9% of the variance, while the conditional $R^2 = 0.595$, indicated the full model (fixed and random effects) explained 60% of the variance in exclusionary outcomes though most of the explained variance was due to differences between students and between schools. In addition, the school level clustering explained the majority of the variance as the ICC_{School} was 0.48 compared to 0.56 total ICC.

The main effects of the PBIS model showed higher Tier 1 fidelity was associated with increased odds of exclusion ($OR = 1.09, p < .001$). However, when accounting for the interaction by state, this relationship reversed. The interaction model revealed a statistically significant moderation effect for schools in Georgia ($OR = 0.91, p < .01$). For schools in the group containing Massachusetts, Minnesota, and Oregon, each one-point increase in Tier 1 fidelity score above the mean was associated with a 12% increase in the odds of exclusion ($OR = 1.12, p < 0.001$). The results suggest higher PBIS implementation fidelity at Tier 1 was associated with lower odds of exclusionary discipline in Georgia, but not in the other states. Additionally, model comparison criteria favored the interaction model ($AIC = 29,515.8$; $BIC = 29,642.6$) over the main effects model ($AIC = 29,546.8$; $BIC = 29,665.1$), suggesting improved fit when accounting for differences by state. The state dependent relationship for odds of exclusionary discipline moderated by PBIS fidelity of implementation can be observed in Table 18.

Table 18

PBIS Tier 1 Fidelity: Predicting Exclusion with and without State Interaction

	Main Effects Model (Tier 1)					Interaction Model (Tier 1 $\times$ State)				
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	β	<i>SE</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	β	<i>SE</i>
White, Male, No IEP, Georgia	1.12	0.25 – 4.99	0.88	0.12	0.76	1.76	0.33 – 9.26	0.51	0.56	0.85
Tier1 (Centered)	1.09	1.07 – 1.12	<0.001	0.09	0.01	0.91	0.85 – 0.97	<0.01	-0.10	0.03
Black	1.05	0.92 – 1.20	0.47	0.05	0.07	1.04	0.91 – 1.19	0.54	0.04	0.07
Latíné	1.11	1.00 – 1.23	0.05	0.10	0.05	1.10	1.00 – 1.22	0.06	0.10	0.05
Other	0.97	0.83 – 1.12	0.65	-0.03	0.08	0.97	0.84 – 1.13	0.73	-0.03	0.08
Female	1.02	0.94 – 1.11	0.60	0.02	0.04	1.02	0.94 – 1.11	0.59	0.02	0.04
IEP	0.81	0.74 – 0.89	<0.001	-0.21	0.05	0.81	0.74 – 0.89	<0.001	-0.21	0.05
Prior Referral	1.40	1.30 – 1.51	<0.001	0.34	0.04	1.40	1.30 – 1.51	<0.001	0.34	0.04
Rural vs City	1.45	0.34 – 6.14	0.61	0.37	0.74	1.36	0.30 – 6.20	0.69	0.31	0.77
Suburb vs City	1.77	0.41 – 7.68	0.45	0.57	0.75	1.43	0.31 – 6.67	0.65	0.36	0.78
Black % Enrolled (Centered)	0.98	0.96 – 1.00	0.08	-0.02	0.01	0.97	0.95 – 1.00	0.02	-0.03	0.01
Other States	0.25	0.06 – 0.96	0.04	-1.40	0.69	0.16	0.04 – 0.71	0.02	-1.82	0.76
Other States x Tier 1(Centered)						1.24	1.15 – 1.33	<0.001	0.21	0.04
Tier 1 (Other States)						1.12	1.10 – 1.15	<0.001	0.12	0.01
Random Effects										
σ^2_{ϵ} Referral	3.29					3.29				
σ^2_{τ} Student	0.53		ICC _{Student}	0.077		0.54		ICC _{Student}	0.074	
σ^2_{μ} School	3.08		ICC _{school}	0.446		3.58		ICC _{School}	0.482	
ICC _{Total}	0.52					0.56				
N _{Schools}	49					49				
N _{Students}	9312					9312				
Observations	34656					34656				
Marginal R ² / Conditional R ²	0.061 / 0.553					0.088 / 0.595				
<i>Note:</i> The table displays a main effects model with Georgia as the reference group followed by an interaction model to examine the interaction between Tier 1 implementation and state group. The mean Tier 1 score was 23.74 out of 30.										

Tier 2 PBIS fidelity of implementation

The Tier 2 PBIS model also revealed considerable between-school variability in exclusion outcomes, with an intraclass correlation coefficient (ICC) of 0.56, suggesting 56% of the variance in exclusionary discipline was attributable to school-level differences. The final analytic sample for PBIS Tier 2 fidelity included 5,573 students and 21,711 minor referrals across 31 schools, the majority of the schools were located in Georgia, 68%. The model's marginal $R^2 = 0.062$, and the conditional $R^2 = 0.590$, reflecting the minimal explanatory power of the model coefficients and the increase in explained variance through school and student clustering. In particular, school level differences in exclusionary discipline practices explained the majority of the variance in the model.

In the main effects model, there was no significant effect of Tier 2 PBIS implementation fidelity ($OR = 1.02$, $p = 0.27$). For Tier 2, the interaction model showed that this relationship did not vary significantly by state. The interaction term (Other States $\times$ Tier 2 Fidelity) was not significant ($OR = 0.96$, $p = 0.30$), suggesting that the effect of Tier 2 fidelity on exclusion was relatively consistent across Georgia and the comparison states. The lack of high-fidelity Tier 2 schools may have impacted the findings as the mean score for Tier 2 fidelity was 16 out of 26 compared to an average of 24 out of 30 for Tier 1 fidelity. Like with the previous models, the majority of the variance was explained by student and school clustering. Model comparison indices again favored the interaction model ($AIC = 16,092.0$; $BIC = 16,211.8$) over the main effects model ($AIC = 16,094.1$; $BIC = 16,197.9$). The results of Tier 2 moderation models are summarized in Table 19.

Table 19 PBIS Tier 2 Fidelity: Predicting Exclusion with and without State Interaction										
	Main Effects Model (Tier 2)					Interaction Model (Tier 2 × State)				
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>β</i>	<i>SE</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>	<i>β</i>	<i>SE</i>
White, Male, No IEP, Urban, Georgia	3.42	0.46 – 25.22	0.23	1.23	1.02	3.63	0.48 – 27.49	0.21	1.29	1.03
Tier 2 (Centered)	1.02	0.99 – 1.06	0.27	0.02	0.02	1.03	0.99 – 1.08	0.15	0.03	0.02
Black	1.08	0.90 – 1.29	0.40	0.08	0.09	1.08	0.90 – 1.29	0.40	0.08	0.09
Latiné	1.10	0.97 – 1.26	0.14	0.10	0.07	1.11	0.97 – 1.26	0.12	0.10	0.07
Other	0.89	0.73 – 1.09	0.27	-0.11	0.10	0.89	0.73 – 1.09	0.28	-0.11	0.10
Female	1.02	0.92 – 1.13	0.68	0.02	0.05	1.02	0.92 – 1.13	0.70	0.02	0.05
IEP	0.81	0.71 – 0.91	0.001	-0.21	0.06	0.81	0.71 – 0.91	0.001	-0.22	0.06
Prior Referral	1.41	1.27 – 1.56	<0.001	0.34	0.05	1.41	1.28 – 1.57	<0.001	0.35	0.05
Rural vs City	3.58	0.52 – 24.71	0.20	1.28	0.99	1.09	0.14 – 8.52	0.93	0.09	1.05
Suburb vs City	1.25	0.17 – 9.20	0.83	0.22	1.02	0.61	0.09 – 4.28	0.62	-0.50	1.00
Percent Black (Centered)	0.99	0.96 – 1.02	0.40	-0.01	0.01	0.97	0.94 – 1.00	0.05	-0.03	0.01
Other States	0.11	0.02 – 0.69	0.02	-2.20	0.93	0.11	0.02 – 0.69	0.02	-2.23	0.94
Other States x Tier 2						0.96	0.90 – 1.03	0.30	-0.04	0.04
Tier 2 (Other States)						1.00	0.94 – 1.05	0.87	-0.00	0.03
Random Effects										
σ_{ϵ}^2 Referral	3.29					3.29				
σ_{τ}^2 Student	0.46			ICC _{Student}	0.062	0.46			ICC _{Student}	0.061
σ_{μ}^2 School	3.69			ICC _{School}	0.495	3.77			ICC _{School}	0.501
ICC _{Total}	0.56					0.56				
N _{Schools}	31					31				
N _{Students}	5573					5573				
Observations	21711					21711				
Marginal R ² / Conditional R ²	0.064 / 0.586					0.062 / 0.590				
<i>Note:</i> The table displays the main effects model with Georgia as the reference group and an interaction model between state and Tier 2 PBIS implementation. The mean Tier 2 score was 16 out of 26.										

Chapter IV: Discussion

Analysis of the referral dataset consisting of major and minor referrals spanning the 2022-2023 and 2023-2024 school years from 4 states, Minnesota, Oregon, Georgia, and Massachusetts, shows state, school, and student level differences impact the odds of receiving exclusionary discipline for minor referrals. Across models, school and state differences explained a large portion of the variance, highlighting the potential for school systems and state policies to impact the use of exclusionary discipline practices. The overall take away from the analysis was the high percentage of minor referrals resulting in exclusionary discipline and the high variability in exclusionary discipline practices by school and state. In an analysis of disproportionality of ever receiving a major referral or exclusionary discipline showed Black students were at increased risk of ever being excluded due to minor referrals compared to White students. The main predictive model, examining referral level odds of exclusion suggested school level interventions for school staff may be important in reducing their use of exclusionary discipline practices, which is important as 40% of minor referrals resulted in exclusionary discipline.

In addition, the study highlighted the importance of using a variety of metrics to examine disproportionality. Through RR and RDR₁₀₀ the study showed the overrepresentation of Black students and those of other race in the receipt of major referrals as well as for ever receiving exclusionary discipline in major and minor referrals. Students with IEPs were also at increased risk of ever being excluded. When examining the odds of exclusion at the referral level and accounting for student, school, and state variables the increased odds of exclusion disappears. The risk ratio data also shows important trends in exclusionary discipline by state. Black students had an elevated risk of major referral and exclusionary discipline for minor referrals in

aggregated counts across states. Although the risk of major referral was higher for Black students in all states, Black students were at especially elevated risk of major referral in Minnesota and Oregon. Similarly, aggregated across states, Black students had higher risk of ever receiving exclusionary discipline, but again there were difference by state. In Massachusetts Black students had a lower risk than White students of exclusionary discipline but in Minnesota, Georgia, and Oregon Black students had an elevated risk of ever receiving exclusionary discipline. There may also be differences in risk based on an incident level risk ratio though the risk ratios reported were calculated at the student level.

One challenge in the analysis was the lack of a population-wide sample. Because the data came from behavior referrals from middle schools implementing PBIS interventions, only students with referrals were included. The restricted dataset meant that only a probability of exclusion conditional on receipt of a minor referral could be examined. Although the gender demographics of most schools in the sample were relatively consistent with almost equal enrollment numbers, the distribution of male and female students in the referral dataset demonstrated the variation in composition from enrollment data. The referral data contained more than double the number of male student referrals compared to female student referrals, with only 35% of students in the sample identified as female. This trend continued when filtered by minor referrals as well, suggesting male students were more likely to receive referrals in general as well as minor referrals.

For students with IEPs, the national average enrollment in public schools of students with disabilities is between 12-14%, but in the minor referral dataset students with disabilities were about 17% of the students and contributed about 25% of the total referrals, demonstrating a potential overrepresentation in the sample and suggesting students with IEPs may receive more

referrals than students without IEPs. Race and ethnicity enrollment varied significantly by state and locale. The sample demographics showed differences between national enrollment trends and the sample which contained 37% White students, 25% Black, 22% Latiné, and 10% Other race including Asian, Native, and multi-racial in the sample. By comparison in 2022 nationally 44% of enrolled students were White, 29% Latiné, 15% Black, 5% Asian, 5% Multi-racial, and 1% Native (National Center for Education Statistics). These differences may suggest an overall overrepresentation of Black students in the referral dataset compared to enrollment, though the data was only collected from four states.

Importantly, the race and ethnicity composition varied widely by state with Georgia containing the majority of Black students in the sample, 90%, while Oregon consisted of a large portion of the Latiné students, 56%. Due to the variation in composition of students in the sample, race and ethnicity differences in odds of exclusion may be partially explained by state-level differences. Furthermore, school level differences may be influenced by the number of schools represented in each state, while the minor referral model contained data from 100 schools, the majority of the data was from Georgia and Oregon with Minnesota and Massachusetts contributing a smaller portion of the referral data.

The Impact of Teacher Implicit Bias on Subjective Referrals

In addition to the variation in student demographic composition and school distributions by state, referral data is complex and may be subjective in nature. Although the focus of this study was exclusionary discipline for minor referrals and not subjective behavior referrals, the nature of behavior referrals can lead to biased outcomes. An analysis of the behavior labels for the minor referrals demonstrated that the majority of the minor referrals were written for subjective behaviors. By definition, subjective referrals are impacted by our own thoughts,

beliefs, and biases and because referrals are written by teachers who inherently have their own biases, it can impact analysis of referrals through a higher number of referrals for students from specific backgrounds as well as more severe referral consequences based on student characteristics.

Labeling of behaviors into different categories may also change by teacher perception of certain student groups (Gilliam et al., 2016; Lorenzetti & Johnson, 2023). In one study examining the language used by teachers for referrals, the authors found teacher description of behavior varied by student race with teachers using more negative emotions when Black student was involved (Markowitz et al., 2023). Additionally, referrals can contain multiple behaviors and actions taken, the study first examined the distribution of all challenging behaviors and administrative actions listed before creating a variable to indicate the presence of an exclusionary discipline response, which was the main outcome variable in the study.

This study was consistent with previous referral research that demonstrated the prevalence of subjective behavior referrals as the driving factor in disproportionality (Girvan et al., 2017; Ritter & Anderson, 2018). Defiance and disruption, followed by disrespect—all subjective behaviors—were the most common challenging behaviors for both major and minor referrals, indicating a tendency for teachers to refer students due to subjective behaviors. The subjective behaviors are strongly correlated with disproportionality by race in school discipline (Girvan et al., 2017; Legette & Anyon, 2023; Little & Welsh, 2022; Welsh & Rodriguez, 2024). Black students in particular have been shown to be at increased risk of receiving a referral and suspension (Morris & Perry, 2017). Although analysis of subjective behaviors was not a specific goal of the study, the results show behaviors that are typically categorized as subjective were the

most common in the middle school setting, hinting at the possibility of differences in discipline practices by teacher.

The potential for teacher implicit bias is exacerbated by the fact that most minor referrals are written for classroom behaviors as over 79% of referrals listed as occurring in the classroom. In the referral sample, the average teacher wrote 21 minor referrals with a range of 1 to 2000, though the median was 7 referrals, showing the majority of teachers wrote a small number of referrals and the average was impacted by a few high referral outliers. The extreme outliers likely represent a single school individual that input all the school's referrals. The potential for teacher implicit bias to impact referral rates has led to interventions aimed at supporting teachers to recognize their own biases through reflection. McIntosh and colleagues (McIntosh et al., 2021b) proposed a teacher intervention to reduce the subjective referrals, finding that directly targeting teacher behavior had promise for reducing disproportionality. Another intervention encouraged an empathetic mindset and found promising results leading to a reduction in suspension rates (Okonofua et al., 2016). Additionally, Gallo and colleagues found that teachers and behavior staff were receptive to implicit bias training and that professional development had the potential to reduce exclusionary discipline disproportionality (Gallo et al., 2025). The results of teacher focused supports suggest interventions to aid teacher behavior management and mindsets may help to reduce subjective referrals and suspensions.

Differences Between Major and Minor Referrals

Once a referral is written, school behavior staff and administrators determine disciplinary actions, which can range from parent contact or a conference with a staff member to OSS and expulsion. The severity of the disciplinary action should vary based on the severity of the challenging behavior, but the disciplinary actions taken are also subject to bias from school staff.

A review of the data found substantial differences in the consequences used in response to major and minor referrals. For instance, as expected major referrals led to more ISS (21%) and OSS (22%) compared to ISS (7%) and OSS (3%) for minor referrals. The most common response to minor behavior referrals was conference with staff (34%) followed by detention (24%) and parent contact (22%). Unsurprisingly, exclusionary discipline responses were more common for major behavior referrals. Parent contact and detention were both substantially different between major and minor referrals with 13% of major referrals leading to detention and 9% resulting in parent contact. One potential issue is referrals are often recorded by hand and entered into the online system manually leading to potential errors in coding. For example, parent contact is likely a standard response to major referrals but may not have been recorded in the online system if other, more severe, consequences were also included.

Other consequences, such as bus suspension, additional support, Saturday school, alternative placement, restitution, community service, and expulsion were rarely used for either type of referral as they each composed less than 1% of the disciplinary actions taken. For minor referrals, the primary exclusionary discipline response was detention, followed by ISS. However, detention as defined in the code book provided by ECS could refer to after school detention, detention during academic time, or lunch detention, which would range from exclusionary in terms of lost instructional time or non-exclusionary if referring to after school detention. Overall, the results demonstrated the increased risk of exclusion for students receiving major referrals, but showed minor referrals result in exclusion as well. In fact, 40% of the minor referrals resulted in some form of exclusionary discipline.

Examining the differences between major and minor referrals for exclusionary discipline showed exclusionary discipline was used more often for major referrals, 56%, but was also

common for minor referrals, 39%. Despite reduced use of exclusionary discipline for minor referrals, they frequently resulted in exclusion from the learning environment. This exclusion from the learning environment leads to lost instructional time, which as prior research has shown can widen the achievement gaps, particularly if specific student groups are excluded at a higher rate based on race, gender, or IEP status (Bell & Puckett, 2023; Hwang, 2018; Losen et al., 2017; Morris & Perry, 2016).

Before focusing on exclusion in minor referrals, risk associated with receipt of a major referral and exclusion for major referral were examined to identify over-representation of student groups in major referrals. When examining exclusion by student race/ethnicity, gender, and IEP status, patterns emerged that suggested increased exclusionary discipline rate by race but not by gender or IEP status. In fact, while students with IEPs showed an increased risk of receiving a major referral and for exclusion at the student level, a referral for a student with an IEP had lower odds of exclusionary discipline. This suggests students with IEPs may receive more referrals which eventually lead to at least one instance of exclusion, but each referral is at lower odds of exclusionary discipline. One possible explanation is that challenging behavior may lead to receiving special education services, but once a student has an IEP, they are less likely to receive exclusionary discipline (Hurwitz et al., 2021). The differences in disciplinary responses to students with disabilities may also be impacted by the disability label because disability manifests differently potentially leading to differences in student perception and behavior. Research has shown that students with an EBD or ADHD/Other Health Disabilities label are more likely to be suspended whereas students with autism spectrum disorder (ASD) are less likely to be suspended (Forman & Markson, 2022; Krezmien et al., 2017; Miller & Meyers, 2015). Research has also shown as a aggregated group, students with IEPs may be at reduced

risk of exclusion as confirmed by the results of the study. Exploring differences in exclusionary discipline by disability category would be beneficial for future studies.

There were race differences for risk of receiving a major referral where Black students, particularly in Minnesota and Oregon had higher risk of receiving major referrals compared to White students. In both states, Black students were almost twice as likely to receive a major referral compared to a minor referral. When examining the risk of exclusionary discipline, the data may be partially explained by state context though. The risk of exclusion was higher in Georgia compared to other states and Georgia had the highest portion of the Black student population in the sample meaning the risk of exclusionary discipline in Georgia had an oversized effect when aggregated across states. Yet when calculating risk ratios disaggregated by state again Minnesota and Oregon had the highest risk of exclusion due to major referral for Black students. This can be explained by Georgia's higher exclusion rate being applied to White and Black students equally. The differences in discipline practice by state suggests that state level policy or culture may contribute to discipline patterns observed in the sample. When examining the odds a referral results in exclusion using a multilevel model to account for school and state differences the differences by race were negligible. Overall, the student and school variables explained only a small portion of the odds of exclusion. In fact, the majority of the explanation of the variance in odds of exclusion were at the school level, showing that school practices were the biggest predictor of exclusion. Future studies should examine referral level predictors as well as additional student and school variables to attempt a more precise prediction model.

Intersectionality

Although race, gender, and IEP status individually were not particularly strong indicators of exclusionary discipline, the intersection of race, gender, and IEP status likely impact

exclusionary discipline risk. Examining the intersection of student demographics suggests Black students with IEPs were at an increased risk of ever receiving a major referral. Additionally, Latiné males with IEPs and students in the other race category with IEPs were at increased risk of receiving a major referral compared to students from those groups without IEPs (Samimi et al., 2023). Studies have shown differences in suspension rates based on the interaction between IEP status and student race which suggest Black students, particularly those with an EBD label are excluded at a higher rate (Krezmien et al., 2017; Sullivan et al., 2014). The interaction between exclusionary discipline for Black students and the EBD label is further exacerbated by the increased likelihood of Black students to be labeled with EBD compared to students of other races (Bal et al., 2019). The risk ratios from the study demonstrate Black students with IEPs may be more likely to receive exclusionary discipline for minor referrals.

In contrast, across race and ethnicity demographics, female students tended to have a slightly lower risk of ever receiving a major referral compared to males. White female students were the only group less likely than White males to ever receive a major referral. The findings were largely consistent when examining exclusionary discipline in minor referrals for intersectional student demographics. Black females with IEPs and Black males without IEPs were the most at risk for exclusionary discipline and White females without IEPs remained the least at risk. These findings are consistent with other studies that have found an increased risk of suspension for Black students and students with IEPs (Samimi et al., 2023; Smith et al., 2023).

Minor Referral Exclusionary Discipline

Although major referrals and suspensions have been well-documented to have disparities based on race and disability, research on minor referrals is less common. In this study, the trend

of disproportionality found in RR and RDR₁₀₀ for major referrals continued in minor referral which showed Black students at an increased risk of ever receiving exclusion compared to White students. The risk ratios for exclusionary discipline for minor referrals by student gender showed a minimal difference between male and female students and between students with IEPs and without IEPs. However, when examining the risk calculated at the state-level the differences appeared, particularly for students with IEPs. Consistent with previous studies, the data showed female students were slightly less likely to be excluded (Wilkerson & Afacan, 2022) but students with IEPs were slightly more likely to be excluded for minor referrals (Cunningham et al., 2024), based on risk ratios. The increased risk of ever receiving exclusion for students with IEPs was present across states but was particularly high in Minnesota. However the odds of a referral for a student with an IEP leading to exclusion were lower than expected. Because of these findings, the interaction between IEP and race/ethnicity should be explored further as risk ratios suggested increased odds of exclusion for Black students with IEPs, though this may be a result of state demographic differences and the distribution of Black students with IEPs in high exclusion states.

The primary differences in risk of exclusionary discipline were for race and ethnicity but with variation by state. For example, Massachusetts was the only state where Black students were less likely than White students to ever receive exclusionary discipline for minor referrals. The other race category—which contained Native, Asian, Multi-racial, and Pacific Islander students—had a slightly elevated risk of exclusion in all states except Minnesota, where they were less likely than White students to receive exclusionary discipline. The variation in risk in the other race category may be due to the composition of the category. In some states the other race category may primarily consist of Asian students while in a different state Native and multi-

racial may be the primary groups in the category. These differences in demographic composition suggest school racial composition may be correlated with exclusionary discipline practices. Smith and colleagues examined the differences in suspension rate by school racial composition and found higher rates of suspension in schools with majority Black students, but when controlling for school-level factors they found Black students had higher risk in majority White school (Smith et al., 2023). The higher suspension rates experienced in majority Black schools may also be due to interactions with higher poverty rates as well as under-resourced schools, both of which are associated with segregated schools (Owens et al., 2022). Due to the role of school demographics, an examination of exclusionary discipline practices should consider differences in school enrollment population.

In this study to account for differences between students, schools, and states, a multilevel model was fit to the data to examine predictors of exclusionary discipline. The findings showed the best indicator of exclusionary discipline was the presence of a prior referral. This is consistent with a meta-analysis of suspension research showing a receipt of a suspension is a strong predictor of future suspensions (Noltemeyer et al., 2015). However, the model also highlighted that between school variation explained the majority of the difference in odds of exclusionary discipline. Furthermore, the model demonstrated the differences between state disciplinary practices. In Georgia and Oregon, students were more likely to receive exclusionary discipline in response to minor behavior referrals, though in different forms. Georgia was more likely to use ISS and OSS and Oregon schools were more likely to use detention. This further highlights the exclusion continuum and the variability in the amount of instructional time students miss when they are removed from the learning environment. In particular, this study

shows the rapid process by which a student can be removed from the learning environment for an extended period of time as a result of a minor behavior.

State policies and practices around behavior management are likely important factors in determining the use of exclusionary discipline. In Massachusetts, a 2014 law went into effect known as Chapter 222, which limited the use of suspension. In addition, the law required principals to send a written explanation of suspension for students in a K-3 program to the superintendent prior to suspension, in essence limiting suspension for early grades. This law reduced the use of exclusionary discipline in Massachusetts and according to one study, found positive effects on ELA achievement as well as a decrease in 9th and 10th grade dropout (Cleveland, 2023). This suggests that limiting suspensions in early grades may have a lasting impact on the use of exclusionary discipline in the state. In another study by Craig and Martin (2023) examining the effects of a suspension ban for non-violent behavior in New York City found the policy had minimal impact on achievement directly, but improved school culture by improving teacher-student relationships and perceptions of safety in school. Taken together, these examples highlight the potential for policy to impact the use of exclusionary discipline.

PBIS Implementation

Positive behavior interventions and supports (PBIS) have been touted as an intervention to reduce the use of suspension and improve school climate (Gage et al., 2018; Lee & Gage, 2020; Sugai & Horner, 2020). Although PBIS is typically categorized into three tiers, this study focused on Tier 1 and Tier 2. Tier 1 PBIS provides a universal school-wide framework for behavior management focused on setting community expectations and providing teachers with proactive management strategies. Tier 2 is more specialized and aimed at supporting students with demonstrated need (Lee & Gage, 2020). Improving Tier 1 systems should reduce the use of

exclusionary discipline due to detailed universal behavior management systems and proactive strategies. Similarly, Tier 2 interventions should reduce exclusionary discipline through targeted support for students with behavioral needs. Tier 3 is focused on supporting students with the highest needs who are unlikely to be receiving minor referrals but instead are likely engaging in major behaviors. An analysis of the impact of PBIS on suspension found full fidelity of all three tiers may have a greater effect of reducing suspension than just tier 1 universal interventions alone (Fleming et al., 2024). This study sought to explore all three tiers, but due to low reporting of Tier 3, limited the investigation to Tier 1 and Tier 2. The study incorporated school fidelity ratings, which are calculated using a tier-specific PBIS scoring rubric with higher scores corresponding to higher quality PBIS implementation. Tier 1 fidelity scores had an inverse relationship with odds of exclusionary discipline in Georgia but not aggregated across all states. The model demonstrated the potential of PBIS as a protective factor for reducing the use of exclusion. Tier 2 scores were not found to be significant in this study, but future research may examine more schools to gain a more precise estimate.

The findings show a need to further standardize the PBIS fidelity process as well as examine differences in state practices in PBIS ratings. Although the Tiered Fidelity Inventory Checklist (TFI) asks the same questions across states it is possible states and schools interpret the questions differently. A variety of interpretations may lead to variation in school scoring. For this study, the small number of schools in the sample, particularly for Minnesota, Massachusetts, and Oregon, resulted in a model with high variance based on few schools. The results of this study suggested PBIS implementation may be a proactive system to reduce the use of exclusionary discipline practices, particular in Georgia, which demonstrated a higher rate of exclusion. Additionally, previous research has suggested PBIS interventions can lead to reductions in

inequities in behavior discipline through teacher training and proactive responses to student behavior (Austin et al., 2024; McIntosh et al., 2021a; R.N.T. Nese et al., 2022; Okonofua et al., 2016).

Implications for Research

The findings of the study suggest several areas for further investigation. First, the differences in the frequency of administrative actions highlight varying behavior discipline practices by schools and states. Second, the presence of exclusionary discipline in minor referrals suggests students are being removed from instructional time for behaviors that are considered minor. The primary forms of exclusion for minor referrals consisted of detention and in-school suspension with Oregon in particular using detention over 5 times more than the next most common state and Georgia being the most likely state to use ISS. The differences highlight the need for research on how administrative policies and school-level practices shape exclusionary responses to minor referrals.

Further study should also examine how “detention” is defined across schools. In this study detention was categorized as exclusionary discipline but detention may present differently depending on school context. In one school it may constitute lunch detention in one—which would not exclude a student from the learning environment—but in another school detention may present similarly to ISS, where a student is removed from instruction. Although non-exclusionary responses were more common for minor referrals in middle school, the use of exclusionary discipline outcomes has the potential to impede student’s learning as a result of minor behavior infractions (Losen et al., 2017). Furthermore, the consequences are on a spectrum of exclusionary discipline where an administrative action, like time in the office, may result in less time out of the classroom than ISS, yet in this study those outcomes were treated

equally. Future research may examine the factors that predict the severity of the discipline response.

The implementation and usage of restorative practices is another area for future research. The main predictor of exclusionary discipline was receipt of a prior referral, which is in line with the idea of escalating responses to repeated behavior, yet exclusionary discipline does not reduce the likelihood of future challenging behavior. The low frequency of restorative practices and restitution suggests schools are not implementing restorative practices to reduce the odds of future challenging behaviors. Future research should explore the barriers to implementation and the current restorative practices in place at schools.

Although the study highlighted the potential for PBIS implementation to have a protective effect on exclusionary discipline for Tier 1 PBIS, the results varied by state. PBIS implementation fidelity is measured through a fidelity score from the Tiered Fidelity Inventory (TFI) (Center on PBIS, 2025) where schools evaluate their practices based on items such as “school policies and procedures describe and emphasize proactive, instructive, and/ or restorative approaches to student behavior that are implemented consistently.” The TFI scores are based on a rubric from where higher fidelity indicate more robust interventions. The findings of the regression model with school PBIS Tier 1 fidelity scores showed Georgia was the only state where PBIS showed a significant protective effect, and the majority of PBIS fidelity score observations came from the state. Future research should examine the role of high-fidelity implementation of PBIS with a larger sample to confirm the protective effect on reducing exclusionary discipline.

Finally, the mixed findings on student race, ethnicity, gender and IEP status point to the need for additional studies of disproportionality in minor referrals. Although the unadjusted risk

ratios demonstrated higher rates of exclusionary discipline for Black students, students with IEPs, and intersectional student groups (i.e. Black male students with IEPs), adjusted predictive model at the referral level did not show the same patterns. This discrepancy suggests that variation in school practices may be more consequential than student demographics in predicting exclusionary discipline. Further research should examine additional school variables to determine in what school contexts does disproportionality emerge in minor referrals. Future models may also incorporate additional referral and student level predictors to improve the predictive ability of the model and increase the amount of explained variance.

Implications for Practice

For practitioners, the findings demonstrate the importance of aligning disciplinary responses to the nature of the challenging behavior. By definition, minor referrals should be managed by classroom teachers through behavior management strategies and relationship building (NH MTSS-B, 2021; PBIS Apps, 2024). Removing students from instruction for minor behaviors—through detention or in-school suspension—risks impeding learning without addressing the underlying needs of the student. Schools should examine their behavior discipline practices to ensure students are not losing instructional time. The results of the study also highlight the potential of restorative practices and the underusage of the strategy. Restorative practices allow students to grow and learn from their mistakes, reduce the likelihood of future behaviors, and boost a student’s sense of belonging in school, which is an important part of preparing students for success. Exclusion on the other hand, does not teach students the skills to succeed in the academic environment and does not reduce the likelihood of future challenging behavior.

Schools should increase the use of data-based decision making with behavior analysis practices to examine student discipline data and reflect on school practices. For students with challenging behaviors, schools need a team to analyze the behavior to identify and address the root cause of the challenging behavior, whether that is a lagging student academic skill or social-emotional need. By developing behavior analysis processes teachers and staff can work to remedy the cause of the challenging behavior.

In terms of teacher training and support, schools need to provide behavior management training opportunities to enable teachers to effectively manage low level behaviors in their classrooms. In addition, schools should promote staff training on implicit bias as the majority of minor referrals were for subjective behaviors which are prone to the effects of implicit bias. Schools also need to ensure that behavior management trainings are aligned with expectations across staff, to allow for more consistency of behavior discipline across the school staff.

School administrators should examine their school PBIS practices to ensure high fidelity implementation across all tiers. The regression findings showed the potential for PBIS Tier 1 to reduce exclusionary discipline, likely through systematic processes that address challenging behavior and prepare school staff for effective behavior management. Although the finding was limited to schools in Georgia with Tier 1 scores, the study suggests schools should seek to adhere to the fidelity framework provided in the Tiered Fidelity Inventory. In order to improve implementation fidelity schools should dedicate professional development time to learning the PBIS framework and strategies. In addition, teachers and staff would benefit from increased time to collaborate across teams to ensure school-wide adoption of PBIS practices. PBIS is most effective when all teachers understand the best practices and are committed to implementation.

Limitations

The data showed patterns of exclusionary discipline from minor behavior referrals for four US states with one of the primary findings suggesting significant differences in disciplinary practices between states. By increasing the number of states in the sample, the model would be better able to demonstrate trends in the US. The current sample was limited in size, with only four states, inhibiting the ability to cluster by state in statistical analysis. Instead, the model incorporated fixed effects by examining a state level factor. A larger sample of states would allow for deeper analysis of trends in regional differences as well as the potential to improve understanding of state policy impact on exclusionary discipline.

The inconsistency in the use of minor referrals across schools was an additional challenge. Minor referrals were chosen as the focus of the study to understand the practices in place for addressing minor student behavior; however, not all schools used the minor referral label. The sample contained 155 unique schools, but of those, only 100 were included in analysis once filtering for the presence of minor referrals and missing data, demonstrating that almost a third of schools in the sample did not record any minor behavior referrals. The lack of standardization in the approach to classifying behavior referrals poses a challenge to understanding the trends in school responses to minor behaviors. Although PBIS has definitions for behaviors, schools may interpret what constitutes a minor referral differently. Without a standardized system, one school may classify a behavior as minor whereas others designate the referral as major, leading to differences in administrative response and potentially skewing the likelihood of exclusionary discipline for minor referrals. Another challenge with the data was the potential for inaccurate counting given that SWIS data is inputted by hand by the school, which may explain discrepancies observed in the data such as a teacher who wrote 2000 referrals in a single school year.

In addition to constraints regarding the labeling of behavior referrals, the sample was also constrained in the information provided. In order to improve the precision of the odds ratio estimates additional variables should be included in the model. For example, at the referral level variables such as time and location of the referral should be examined. In the present study, the student level data was limited to the student's grade, gender, IEP status, race, and ethnicity. SES information was not present at the student level, yet research has shown a strong association between socioeconomic status and suspension (Leung-Gagné et al., 2022; Sabol et al., 2022). Similarly, at the school level, Title 1 status—a proxy for percent low-income—was not available nor were district identification numbers, which inhibited examining district level differences in exclusionary discipline rather than solely by school and state. Examining trends by district would allow the study to observe if behavior management was more similar across states or by district.

The model would also benefit from incorporating additional school level variables such as enrollment of different student groups. Although the sample did contain NCES data on school enrollment by gender and race, the data did not contain information about enrollment of students with disabilities. Furthermore, the enrollment data by race led to school estimates of enrollment that often differed from the possible enrollment based on the referral data and student id information as well as discrepancies between school reported enrollment and NCES. Due to these discrepancies the school enrollment data, particularly for the student racial groups with a relatively small number of students such as Native and Pacific Islander students were not used in analysis. In future studies, NCES data should be prioritized to create additional school level variables for the model. In this study, to develop a working multilevel model that converged with the predictors, the race/ethnicity information was combined into a single variable with students

listed as Asian, Native, Multi-racial, or Pacific-Islander falling into the other race category. A more in-depth analysis into differences by racial group would require a greater sample size.

Similarly, interactions between race, gender, and IEP status were primarily examined through Risk Ratios (RR) and Raw Differential Representation (RDR) due to the inability of the multi-level model to converge while incorporating the interaction terms.

Although the predictive model found no racial differences, the composition of students and their unadjusted risk ratios varied significantly between states with Georgia containing the majority of Black students in the sample while Oregon and Massachusetts had relatively few Black students and a larger number of Latiné students, particularly in Oregon. The differences in student demographics may have been obscured by state level differences in managing behavior. Furthermore, Black students were at increased risk of receiving a major referral compared to minor referrals suggesting Black students may have been overrepresented in the major referral category and underrepresented in minor referrals. This underrepresentation may have impacted the ability to observe disproportionality for minor referrals. Additionally, the dataset only contained students who received at least one minor referral, meaning the sample demographics may not have reflected the enrollment data for the sample schools.

A potential confounding factor is school connectedness. Connection to school staff, peers, and community have been shown to some of the most influential factors in building school connection, which may reduce the rate of challenging behavior (Allen et al., 2018; Anyon et al., 2018). A meta-analysis of school connectedness research also found students who feel a sense of belonging in school and in the classroom have improved academic and SEL outcomes (Korpershoek et al., 2020). This study did not examine any measures of sense of belonging, but

connection may have played a role in the rates and severity of challenging behaviors by school leading to differences in exclusionary discipline.

The final portion of the study examined the potential of PBIS interventions as a protective factor to reduce exclusionary discipline for minor referrals. The study explored the effect of Tier 1 and Tier 2 PBIS intervention fidelity, with findings hinting at benefits with increased PBIS fidelity, which varied by state. The Tier 2 findings in particular should be interpreted with caution as very few schools were represented in the sample and roughly 63% of those schools were from Georgia. Tier 1 data was present for about half of the schools, 52 out of 100, with Georgia again representing the largest proportion of reporting schools. Tier 3 PBIS, which focuses on supporting students with the highest needs, was unable to be included in the multi-level model due to a low number of schools reporting the TFI fidelity scores for Tier 3. The incomplete PBIS fidelity scores inhibited the ability to examine the potential of full intervention implementation to reduce the use of exclusionary discipline for minor referrals.

Future Research

The findings of this study highlight the differences in the use of exclusionary discipline for minor referrals between schools and states. The study found an increased unadjusted risk for Black students to receive major referrals compared to minor referrals as well as increased risk of exclusionary discipline for minor referrals. Future research should examine protective factors and mechanisms leading to disproportionality in minor referrals. For example, a measure examining student sense of belonging should be included in future studies to examine the potential for connection to school as a protective factor for risk of minor referrals and exclusionary discipline. The role of student socioeconomic level and school Title I status may also help explain differences in school discipline practices as low-income schools have been

associated with higher rates of exclusionary discipline (Camacho & Krezmien, 2019). Future analysis should also examine the interaction between SES and race and ethnicity as research suggests an interaction between the two (Sabol et al., 2022).

In addition to expanding the study to examine a wider number of states, future research should analyze the patterns of discipline for elementary and high school students. Elementary students may receive fewer referrals overall but may have a higher number of minor referrals as students spend more time with the classroom teacher. On the other hand, high school students may have more challenging behaviors generally leading to higher odds of exclusion. The presence of disparities in exclusionary discipline in either elementary or high school have the potential to negatively impact student outcomes through reduced sense of belonging in school leading to disengagement as well as exclusionary discipline resulting in missed academic instructional time (Craig & Martin, 2023).

Another potential area for future research is the addition of referral, student, and school covariates in the model to improve the model's explained variance. The fixed effects of the main model only explained about 9% of the variance in odds of exclusion. This increased to around 55% when accounting for student and school clustering, but the model would ideally explain more variance in odds of exclusion, which would require additional variables to be added to the model.

High PBIS implementation at Tier 1 was found to be a significant predictor of reduced exclusionary discipline in Georgia but examining the level at which PBIS implementation is sufficient to lower exclusion as well as which components of PBIS implementation are the greatest predictors of reduced exclusion are important in generalizing PBIS interventions to reduce exclusionary discipline. Tier 1 PBIS contains measures of teaming, school behavior

processes, and proactive classroom strategies, but this study examined the composite score without delving into which specific components of Tier 1 PBIS were the most important for reducing exclusionary discipline. Tier 1 and Tier 2 were the focus of this study due reporting from schools, but Tier 3 PBIS strategies should be examined to understand the differences in impact on exclusionary discipline practices. Tier 2 interventions, which target students who have demonstrated a need for increased support, may have a stronger protective factor against the odds of future exclusion due to the student already being at risk of receiving a referral. In this study the Tier 2 findings were largely inconclusive due missing data and a reduced sample size of only 44 schools. Tier 3 strategies were not examined at all due to a lack of reporting. Future studies should increase the number of schools in the sample to reach a threshold of schools reporting PBIS implementation to include their scores in a statistical model.

Recommendations

The goal of behavior management is to reduce the frequency and severity of challenging behaviors to create a space where students can learn. Unfortunately, behavior discipline can often lead to disproportionate exclusion from the learning environment for specific student groups. By improving our understanding of patterns in minor behavior management, we can create more equitable schools. Furthermore, behavior discipline processes are an important component of the PBIS intervention. PBIS implementation fidelity survey items focus on ensuring school-wide behavior expectations and flowcharts for handling student behavior referrals, which develop the processes needed for effective proactive response to challenging behavior. The study has provided evidence for the effectiveness of high-fidelity Tier 1 PBIS interventions to reduce exclusionary discipline in minor referrals. Schools would benefit from increased focus on implementing Tier 1 PBIS with support from agencies such as the PBIS

Center to ensure high-fidelity implementation. The variation in disciplinary practices between schools with high Tier 1 fidelity scores demonstrates the need for further implementation checks across schools to ensure PBIS interventions are implemented accurately. One potential mechanism to improve consistency of disciplinary practices across schools is to create uniform definitions for administrative actions and checks for accurate implementation across schools within a district. Similarly, developing a consensus, particularly within a building, on the definitions of challenging behaviors to ensure a subjective behavior such as defiance or disrespect are the same regardless of classroom.

In addition, the primary responses to minor referrals were parent contact and conference with staff, with exclusionary practices such as detention and ISS more commonly used than restorative practices. Restorative practices, which allow students to repair relationships and mend harm in the classroom community, were rarely used in the sample, accounting for less than 1% of administrative actions for minor referrals. Developing school systems to facilitate the use of restorative practices may provide a valuable tool for schools to reduce the use of exclusionary discipline while allowing students to develop skills needed to succeed in the classroom through the development of social understanding and a sense of belonging. Restorative practices may also improve the student's relationship with their teachers and peers as well as a sense of connection to their school, which may reduce challenging behaviors.

Summary

Student risk of ever receiving a major referral and exclusionary discipline was elevated for Black students as well as those with IEPs, but when accounting for student and school level predictors the odds of a referral resulting in exclusion was not significantly different by race and the odds of a referral for students with an IEP resulting in exclusionary discipline were lower

than students without IEPs. However, exclusionary discipline rates varied significantly between schools and by state with students in Georgia having significantly higher odds of referrals leading to exclusion whereas those in Massachusetts had lower odds of receiving exclusionary discipline. Across states and student demographics, 40% of minor referrals resulted in exclusionary discipline, highlighting the potential for lost instructional time based on behavior discipline patterns. In addition, the study demonstrates the rapid pipeline of minor behaviors to exclusion which can quickly impact a student's academic opportunity by removing them from the classroom for low level behaviors.

The multi-level model demonstrated that school differences in exclusionary discipline practices explained the greatest amount of variance in odds of a referral leading to exclusionary discipline, which highlights the possibility of intervening at the school level to change behavior discipline practices. The study also examined the impact of PBIS intervention fidelity on the rate of exclusionary discipline as all schools in the sample were PBIS schools. Although the state differences remained, the multi-level model indicated schools with higher PBIS Tier 1 fidelity of implementation in Georgia were significantly less likely to result in the use of exclusionary discipline. These findings demonstrate the potential for PBIS implementation to reduce exclusionary discipline practices and disparities for minor referrals.

APPENDIX

Full citations of R packages used in Analysis

- Arel-Bundock, V., Greifer, N., Heiss, A. (2024). “How to Interpret Statistical Models Using `marginaleffects` for R and Python.” *Journal of Statistical Software*, 111(9), 1-32. <https://doi.org/10.18637/jss.v111.i09>.
- Barrett, T., Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., Hocking, T., Schwendinger, B., Krylov, I. (2025). *data.table: Extension of `data.frame`*. R package version 1.17.0, <https://CRAN.R-project.org/package=data.table>.
- Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). “Fitting Linear Mixed-Effects Models Using `lme4`.” *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>.
- Bates, D., Maechler, M., Jagan, M. (2024). *Matrix: Sparse and Dense Matrix Classes and Methods*. R package version 1.7-0, <https://CRAN.R-project.org/package=Matrix>.
- Fox, J., Weisberg, S. (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. <https://www.john-fox.ca/Companion/>.
- Fox, J., Weisberg, S., Price, B. (2022). *carData: Companion to Applied Regression Data Sets*. R package version 3.0-5, <https://CRAN.R-project.org/package=carData>.
- Gohel, D., Moog, S., Heckmann, M. (2025). *officer: Manipulation of Microsoft Word and PowerPoint Documents*. R package version 0.6.8, <https://CRAN.R-project.org/package=officer>.
- Gohel, D., Skintzos, P. (2024). *flextable: Functions for Tabular Reporting*. R package version 0.9.7, <https://CRAN.R-project.org/package=flextable>.
- Grolemund, G., Wickham, H. (2011). “Dates and Times Made Easy with `lubridate`.” *Journal of Statistical Software*, 40(3), 1-25. <https://www.jstatsoft.org/v40/i03/>.
- Iannone, R., Cheng, J., Schloerke, B., Hughes, E., Lauer, A., Seo, J., Brevoort, K., Roy, O. (2024). *gt: Easily Create Presentation-Ready Display Tables*. R package version 0.11.1, <https://CRAN.R-project.org/package=gt>.
- Lenth, R. (2025). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.11.1, <https://CRAN.R-project.org/package=emmeans>.
- Long, JA. (2022). *jtools: Analysis and Presentation of Social Scientific Data*. R package version 2.2.0, <https://cran.r-project.org/package=jtools>.

- Lüdecke, D. (2018). “ggeffects: Tidy Data Frames of Marginal Effects from Regression Models.” *Journal of Open Source Software*, 3(26), 772. <https://doi.org/10.21105/joss.00772>.
- Lüdecke, D. (2024). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.16, <https://CRAN.R-project.org/package=sjPlot>.
- Lüdecke, D., Patil, I., Ben-Shachar, M., Wiernik, B., Waggoner, P., Makowski, D. (2021). “see: An R Package for Visualizing Statistical Models.” *Journal of Open Source Software*, 6(64), 3393. <https://doi.org/10.21105/joss.03393>.
- Makowski, D., Lüdecke, D., Patil, I., Thériault, R., Ben-Shachar, M., Wiernik, B. (2023). “Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption.” *CRAN*. <https://easystats.github.io/report/>.
- Müller, K., Wickham, H. (2023). *tibble: Simple Data Frames*. R package version 3.2.1, <https://CRAN.R-project.org/package=tibble>.
- R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Tierney, N., Cook, D. (2023). “Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations.” *Journal of Statistical Software*, 105(7), 1-31. <https://doi.org/10.18637/jss.v105.i07>.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
- Wickham, H. (2023). *forcats: Tools for Working with Categorical Variables (Factors)*. R package version 1.0.0, <https://CRAN.R-project.org/package=forcats>.
- Wickham, H. (2023). *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.5.1, <https://CRAN.R-project.org/package=stringr>.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, LD., François, R., Golemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H. (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, H., Bryan, J. (2023). *readxl: Read Excel Files*. R package version 1.4.3, <https://CRAN.R-project.org/package=readxl>.
- Wickham, H., François, R., Henry, L., Müller, K., Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>.

- Wickham, H., Henry, L. (2023). *purrr: Functional Programming Tools*. R package version 1.0.2, <https://CRAN.R-project.org/package=purrr>.
- Wickham, H., Hester, J., Bryan, J. (2024). *readr: Read Rectangular Text Data*. R package version 2.1.5, <https://CRAN.R-project.org/package=readr>.
- Wickham, H., Miller, E., Smith, D. (2023). *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. R package version 2.5.4, <https://CRAN.R-project.org/package=haven>.
- Wickham, H., Vaughan, D., Girlich, M. (2024). *tidyr: Tidy Messy Data*. R package version 1.3.1, <https://CRAN.R-project.org/package=tidyr>.
- Yanagida, T. (2025). *misty: Miscellaneous Functions 'T. Yanagida'*. R package version 0.7.1, <https://CRAN.R-project.org/package=misty>.

REFERENCES

- Allen, K., Kern, M. L., Vella-Brodrick, D., Hattie, J., & Waters, L. (2018). What Schools Need to Know About Fostering School Belonging: A Meta-analysis. *Educational Psychology Review*, 30(1), 1–34. <https://doi.org/10.1007/s10648-016-9389-8>
- Allman, K. L., & Slate, J. R. (2011). School Discipline in Public Education: A Brief Review of Current Practices. *The Connexions Project*. <https://files.eric.ed.gov/fulltext/EJ973838.pdf>
- Amemiya, J., Mortenson, E., & Wang, M.-T. (2020). Minor infractions are not minor: School infractions for minor misconduct may increase adolescents' defiant behavior and contribute to racial disparities in school discipline. *American Psychologist*, 75(1), 23–36. <https://doi.org/10.1037/amp0000475>
- 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>
- Anyon, Y., Atteberry-Ash, B., Yang, J., Pauline, M., Wiley, K., Cash, D., Downing, B., Greer, E., & Pisciotta, L. (2018). It's All about the Relationships": Educators' Rationales and Strategies for Building Connections with Students to Prevent Exclusionary School Discipline Outcomes. *Children & Schools*, 40(4), 221–230.
- Austin, S. C., McIntosh, K., Smolkowski, K., Santiago-Rosario, M. R., Arbuckle, S. L., & Barney, N. E. (2024). Examining differential effects of an equity-focused schoolwide positive behavioral interventions and supports approach on teachers' equity in school discipline. *Journal of School Psychology*, 104, 101284. <https://doi.org/10.1016/j.jsp.2024.101284>
- Bal, A., Betters-Bubon, J., & Fish, R. E. (2019). A Multilevel Analysis of Statewide Disproportionality in Exclusionary Discipline and the Identification of Emotional Disturbance. *Education and Urban Society*, 51(2), 247–268. <https://doi.org/10.1177/0013124517716260>
- Barrett, N., McEachin, A., Mills, J. N., & Valant, J. (2021). Disparities and Discrimination in Student Discipline by Race and Family Income. *Journal of Human Resources*, 56(3), 711–748. <https://doi.org/10.3368/jhr.56.3.0118-9267R2>
- Barton, R., & Nishioka, V. (2014). Discipline Disparities: Implications for School Practice and Policy. *Principal's Research Review*, 9(4).
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>

- Bell, C., & Puckett, T. (2023). I Want to Learn but They Won't Let Me: Exploring the Impact of School Discipline on Academic Achievement. *Urban Education*, 58(10), 2658–2688. <https://doi.org/10.1177/0042085920968629>
- Bera, L., Souchon, M., Ladsous, A., Colin, V., & Lopez-Castroman, J. (2022). Emotional and Behavioral Impact of the COVID-19 Epidemic in Adolescents. *Current Psychiatry Reports*, 24(1), 37–46. <https://doi.org/10.1007/s11920-022-01313-8>
- Bergh, B., & Cowell, J. (2013). Discipline Referral Outcomes: Meeting the Needs of Students. *NCPEA Education Leadership Review*, 14(3).
- Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and white adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, 109(4), 532–545. <https://doi.org/10.1037/edu0000155>
- Bottiani, J. H., Kush, J. M., McDaniel, H. L., Pas, E. T., & Bradshaw, C. P. (2023). Are We Moving the Needle on Racial Disproportionality? Measurement Challenges in Evaluating School Discipline Reform. *American Educational Research Journal*, 60(2), 293–329. <https://doi.org/10.3102/00028312221140026>
- 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(2), 508–520. <https://doi.org/10.1037/a0018450>
- Brobbe, G. (2018). Punishing the Vulnerable: Exploring Suspension Rates for Students with Learning Disabilities. *Intervention in School and Clinic*, 53(4), 216–219. <https://doi.org/10.1177/1053451217712953>
- Camacho, K. A., & Krezmien, M. P. (2019). Individual- and School-Level Factors Contributing to Disproportionate Suspension Rates: A Multilevel Analysis of One State. *Journal of Emotional and Behavioral Disorders*, 27(4), 209–220. <https://doi.org/10.1177/1063426618769065>
- Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and Indirect Aggression During Childhood and Adolescence: A Meta-Analytic Review of Gender Differences, Intercorrelations, and Relations to Maladjustment. *Child Development*, 79(5), 1185–1229. <https://doi.org/10.1111/j.1467-8624.2008.01184.x>
- Carter, P., L., Skiba, R., Arredondo, M., & Pollock, M. (2016). You Can't Fix What You Don't Look At: Acknowledging Race in Addressing Racial Discipline Disparities. *Urban Education*, 29. <https://doi.org/10.1177/0042085916660350>
- Center on PBIS. (2021). *What is PBIS?* <https://www.pbis.org/pbis/what-is-pbis>

- Center on PBIS. (2025). Tiered Fidelity Inventory (TFI) 3 Manual. *Center on PBIS*, 3. <https://files.pbisapps.org/pub/pdf/TFI-3-Manual.pdf>
- Cholewa, B., Hull, M. F., Babcock, C. R., & Smith, A. D. (2018). Predictors and academic outcomes associated with in-school suspension. *School Psychology Quarterly*, 33(2), 191–199. <https://doi.org/10.1037/spq0000213>
- 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>
- Civil Rights Data Collection. (2023). *2020-21 Civil Rights Data Collection Student Discipline and School Climate in U.S. Public Schools*. US Department of Education Office of Civil Rights. <https://civilrightsdata.ed.gov/>
- Cleveland, C. (2023). *Rethinking Discipline: The effects of state discipline reform laws on students*. Harvard University.
- Craig, A., & Martin, D. C. (2023). *Discipline Reform, School Culture, and Student Achievement EdWorkingPaper*, 23(722). Annenberg Institute at Brown University. <https://doi.org/10.26300/x14m-2f97>
- Cribb Fabersunne, C., Lee, S. Y., McBride, D., Zahir, A., Gallegos-Castillo, A., LeWinn, K. Z., & Morris, M. D. (2023). Exclusionary School Discipline and School Achievement for Middle and High School Students, by Race and Ethnicity. *JAMA Network Open*, 6(10). <https://doi.org/10.1001/jamanetworkopen.2023.38989>
- Cunningham, M. C., McDermott, L., & Cruz, R. A. (2024). Do I Belong Yet? The Relationship Between Special Education, In-School Suspension, Belonging, and Engagement. *Remedial and Special Education*. <https://doi.org/10.1177/07419325241277884>
- Curran, F. C. (2020). A Matter of Measurement: How Different Ways of Measuring Racial Gaps in School Discipline Can Yield Drastically Different Conclusions About Racial Disparities in Discipline. *Educational Researcher*, 49(5), 382–387. <https://doi.org/10.3102/0013189X20923348>
- Davis, K. M., Nation, M., Christopher, C., & Fisher, B. W. (2024). The Role of Social Emotional Competencies in Student Discipline and Discipline Disparities. *Journal of School Violence*, 23(2), 250–263. <https://doi.org/10.1080/15388220.2023.2297040>
- Del Toro, J., & Wang, M.-T. (2022). The roles of suspensions for minor infractions and school climate in predicting academic performance among adolescents. *American Psychologist*, 77(2), 173–185. <https://doi.org/10.1037/amp0000854>
- Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel Modeling Using R* (Second Edition). CRC Press/Taylor & Francis Group.

- Fisher, B. W., Dawson-Edwards, C., Higgins, E. M., & Swartz, K. (2020). Who belongs in school? Examining the link between Black and White racial disparities in sense of school belonging and suspension. *Journal of Community Psychology*, 48(5), 1481–1499. <https://doi.org/10.1002/jcop.22342>
- Fleming, J. I., Grasley-Boy, N. M., Gage, N. A., Lombardo, M., & Anderson, L. (2024). Effects of Tiered SWPBIS Fidelity on Exclusionary Discipline Outcomes for Students with Disabilities: A Conceptual Replication. *Journal of Positive Behavior Interventions*, 26(1), 3–13. <https://doi.org/10.1177/10983007231193173>
- Forman, K., & Markson, C. (2022). Students with Disabilities and Suspension Rates: A Cautionary Tale for School Districts. *Journal of Leadership and Inclusion*, 20–24.
- Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage. <https://www.john-fox.ca/Companion/>
- Freiberg, J., & Lapointe, J. (2006). Research-Based Programs for Preventing and Solving Discipline Problems. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of Classroom Management* (pp. 735–787). Routledge.
- Gage, N. A., Adamson, R., MacSuga-Gage, A. S., & Lewis, T. J. (2017). The Relation Between the Academic Achievement of Students with Emotional and Behavioral Disorders and Teacher Characteristics. *Behavioral Disorders*, 43(1), 213–222.
- Gage, N. A., Grasley-Boy, N., Peshak George, H., Childs, K., & Kincaid, D. (2019). A Quasi-Experimental Design Analysis of the Effects of School-Wide Positive Behavior Interventions and Supports on Discipline in Florida. *Journal of Positive Behavior Interventions*, 21(1), 50–61. <https://doi.org/10.1177/1098300718768208>
- Gage, N. A., Katsiyannis, A., Carrero, K. M., Miller, R., & Pico, D. (2021). Exploring Disproportionate Discipline for Latinx Students With and Without Disabilities: A National Analysis. *Behavioral Disorders*, 47(1), 3–13. <https://doi.org/10.1177/0198742920961356>
- Gage, N. A., Whitford, D. K., & Katsiyannis, A. (2018). A Review of Schoolwide Positive Behavior Interventions and Supports as a Framework for Reducing Disciplinary Exclusions. *The Journal of Special Education*, 52(3), 142–151. <https://doi.org/10.1177/0022466918767847>
- Gagnon, J. C., Gurel, S., & Barber, B. R. (2017). State-Level Analysis of School Punitive Discipline Practices in Florida. *Behavioral Disorders*, 42(2), 65–80. <https://doi.org/10.1177/0198742916688652>
- Gale, A., & Nepomnyaschy, L. (2024). School Connectedness and Mental Health Among Black Adolescents. *Journal of Youth and Adolescence*, 53(5), 1066–1077. <https://doi.org/10.1007/s10964-023-01898-0>

- Gallagher, E. K., Dever, B. V., Hochbein, C., & DuPaul, G. J. (2019). Teacher Caring as a Protective Factor: The Effects of Behavioral/Emotional Risk and Teacher Caring on Office Disciplinary Referrals in Middle School. *School Mental Health, 11*(4), 754–765. <https://doi.org/10.1007/s12310-019-09318-0>
- Gallo, J. C. R., Austin, S. C., & McIntosh, K. (2025). Effects of Implicit Bias Training on Educator Beliefs and Attitudes Related to Racial Bias. *Psychology in the Schools, 62*(10), 3988–3999. <https://doi.org/10.1002/pits.23586>
- Gayl, C. L. (2017). *How State Planning for the Every Student Succeeds Act (ESSA) Can Promote Student Academic, Social, and Emotional Learning: An Examination of Five Key Strategies* [CASEL Brief]. Collaborative for Academic, Social, and Emotional Learning.
- Gerlinger, J., Viano, S., Gardella, J. H., Fisher, B. W., Chris Curran, F., & Higgins, E. M. (2021). Exclusionary School Discipline and Delinquent Outcomes: A Meta-Analysis. *Journal of Youth and Adolescence, 50*(8), 1493–1509. <https://doi.org/10.1007/s10964-021-01459-3>
- Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions? *Yale University Child Study Center*.
- Gion, C., McIntosh, K., & Smolkowski, K. (2018). Examination of American Indian/Alaska Native School Discipline Disproportionality Using the Vulnerable Decision Points Approach. *Behavioral Disorders, 44*(1), 40–52. <https://doi.org/10.1177/0198742918773438>
- Girvan, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2017). The relative contribution of subjective office referrals to racial disproportionality in school discipline. *School Psychology Quarterly, 32*(3), 392–404. <https://doi.org/10.1037/spq0000178>
- Girvan, E. J., McIntosh, K., & Smolkowski, K. (2019). Tail, Tusk, and Trunk: What Different Metrics Reveal About Racial Disproportionality in School Discipline. *Educational Psychologist, 54*(1), 40–59. <https://doi.org/10.1080/00461520.2018.1537125>
- Gopalan, M., Asson, S., Cattell, M., & Frankenberg, E. (2024). Historical Patterns and Trends in Racial/Ethnic Disproportionality in School Discipline in the United States. *Socius: Sociological Research for a Dynamic World, 10*. <https://doi.org/10.1177/23780231241285105>
- Gopalan, M., & Nelson, A. A. (2019). Understanding the Racial Discipline Gap in Schools. *AERA Open, 5*(2). <https://doi.org/10.1177/2332858419844613>
- Gregory, A., Huang, F. L., Anyon, Y., Greer, E., & Downing, B. (2018). An Examination of Restorative Interventions and Racial Equity in Out-of-School Suspensions. *School Psychology Review, 47*(2), 167–182. <https://doi.org/10.17105/SPR-2017-0073.V47-2>

- Gregory, A., & Roberts, G. (2017). Teacher Beliefs and the Overrepresentation of Black Students in Classroom Discipline. *Theory Into Practice*, 56(3), 187–194. <https://doi.org/10.1080/00405841.2017.1336035>
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The Achievement Gap and the Discipline Gap: Two Sides of the Same Coin? *Educational Researcher*, 39(1), 59–68. <https://doi.org/10.3102/0013189X09357621>
- Griffith, A. N. (2023). “They Do Us Wrong”: Bringing Together Black Adolescent Girls’ Voices on School Staff’s Differential Treatment. *Journal of Black Psychology*, 49(5), 598–645. <https://doi.org/10.1177/00957984231191848>
- 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>
- Herry, E., & Mulvey, K. L. (2025). Students in the margins: Lesbian, gay, and bi+ students’ perceptions of school climate and discrimination. *International Journal of LGBTQ+ Youth Studies*, 22(2), 273–295. <https://doi.org/10.1080/19361653.2024.2310649>
- Holbein, J. B., & Ladd, H. F. (2017). Accountability pressure: Regression discontinuity estimates of how No Child Left Behind influenced student behavior. *Economics of Education Review*, 58, 55–67. <https://doi.org/10.1016/j.econedurev.2017.03.005>
- Huang, F. L. (2018). Do Black students misbehave more? Investigating the differential involvement hypothesis and out-of-school suspensions. *The Journal of Educational Research*, 111(3), 284–294. <https://doi.org/10.1080/00220671.2016.1253538>
- Huang, F. L., Gregory, A., & Ward-Seidel, A. R. (2023). The Impact of Restorative Practices on the Use of Out-of-School Suspensions: Results from a Cluster Randomized Controlled Trial. *Prevention Science*, 24(5), 962–973. <https://doi.org/10.1007/s11121-023-01507-3>
- Hurwitz, S., Cohen, E. D., & Perry, B. L. (2021). Special Education Is Associated with Reduced Odds of School Discipline Among Students with Disabilities. *Educational Researcher*, 50(2), 86–96. <https://doi.org/10.3102/0013189X20982589>
- Hwang, N. (2018). Suspensions and Achievement: Varying Links by Type, Frequency, and Subgroup. *Educational Researcher*, 47(6), 363–374. <https://doi.org/10.3102/0013189X18779579>
- James, A. G., Noltemeyer, A., Ritchie, R., Palmer, K., & University, M. (2019). Longitudinal disciplinary and achievement outcomes associated with school-wide PBIS implementation level. *Psychology in the Schools*, 56(9), 1512–1521. <https://doi.org/10.1002/pits.22282>
- Kang-Brown, J., Trone, J., Fratello, J., & Daftary-Kapur, T. (2023). A Generation Later: What We’ve Learned about Zero Tolerance in Schools. *Vera Institute of Justice*.

- Kaufman, J. S., Jaser, S. S., Vaughan, E. L., Reynolds, J. S., Di Donato, J., Bernard, S. N., & Hernandez-Brereton, M. (2010). Patterns in Office Referral Data by Grade, Race/Ethnicity, and Gender. *Journal of Positive Behavior Interventions*, 12(1), 44–54. <https://doi.org/10.1177/1098300708329710>
- Kelley, B., Jamieson, C., & Perez, Z. Jr. (2021). *50-State comparison: School discipline policies*. Education Commission of the States: ECS. <https://www.ecs.org/50-state-comparison-school-discipline-policies/>
- Kim, J., McIntosh, K., Mercer, S. H., & Nese, R. N. T. (2018). Longitudinal Associations Between SWPBIS Fidelity of Implementation and Behavior and Academic Outcomes. *Behavioral Disorders*, 43(3), 357–369. <https://doi.org/10.1177/0198742917747589>
- Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M., & De Boer, H. (2020). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Research Papers in Education*, 35(6), 641–680. <https://doi.org/10.1080/02671522.2019.1615116>
- Kremer, K. P., Flower, A., Huang, J., & Vaughn, M. G. (2016). Behavior problems and children's academic achievement: A test of growth-curve models with gender and racial differences. *Children and Youth Services Review*, 67, 95–104. <https://doi.org/10.1016/j.childyouth.2016.06.003>
- Krezmien, M. P., Travers, J. C., & Camacho, K. (2017). Suspension rates of students with autism or intellectual disabilities in Maryland from 2004 to 2015. *Journal of Intellectual Disability Research*, 61(11), 1011–1020. <https://doi.org/10.1111/jir.12406>
- Kunesh, C. E., & Noltemeyer, A. (2019). Understanding Disciplinary Disproportionality: Stereotypes Shape Pre-Service Teachers' Beliefs About Black Boys' Behavior. *Urban Education*, 54(4), 471–498. <https://doi.org/10.1177/0042085915623337>
- Lacoe, J., & Steinberg, M. P. (2019). Do Suspensions Affect Student Outcomes? *Educational Evaluation and Policy Analysis*, 41(1), 34–62. <https://doi.org/10.3102/0162373718794897>
- Lee, A., & Gage, N. A. (2020). Updating and expanding systematic reviews and meta-analyses on the effects of school-wide positive behavior interventions and supports. *Psychology in the Schools*, 57(5), 783–804. <https://doi.org/10.1002/pits.22336>
- Legette, K. B., & Anyon, Y. (2023). Just go to the office! An intersectional exploration of the role of race and gender in discipline referral reasons. *Race, Ethnicity and Education*. <https://doi.org/10.1080/13613324.2023.2192946>
- Lehmann, P. S. (2023). Classroom Disruptiveness and Racial, Ethnic, and Gender Disparities in Office Discipline Referrals. *Journal of School Violence*, 22(4), 581–594. <https://doi.org/10.1080/15388220.2023.2255131>

- Leung-Gagné, M., McCombs, J., Scott, C., & Losen, D. (2022). *Pushed out: Trends and disparities in out-of-school suspension*. Learning Policy Institute. <https://doi.org/10.54300/235.277>
- Lincove, A., Jane, Mata, C., & Cortes, K. E. (2024). The Effects of a Statewide Ban on School Suspensions. *Annenberg Institute at Brown University*. <https://doi.org/10.26300/RZKW-Y763>
- Little, S. J., & Welsh, R. O. (2022). Rac(e)ing to punishment? Applying theory to racial disparities in disciplinary outcomes. *Race Ethnicity and Education*, 25(4), 564–584. <https://doi.org/10.1080/13613324.2019.1599344>
- Liu, X., Qian, W., Liang, P., Xie, T., & Wang, J. (2022). The relationship between peer victimization and social anxiety among adolescents: A moderated mediation model with cognitive flexibility and affect intensity. *Psychology, Health & Medicine*, 27(3), 663–675.
- Lorenzetti, N. L., & Johnson, H. (2023). The Relationship Between Teachers' Implicit Racial Attitudes and Their Labeling of Classroom Behavior. *Critical Questions in Education*, 14(1).
- Losen, D.J., Hodson, C., Ee, J., & Martinez, T. (2014). Disturbing Inequities: Exploring the Relationship Between Racial Disparities in Special Education Identification and Discipline. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, 5(2). <https://doi.org/10.58464/2155-5834.1224>
- Losen, D. J., & Martinez, P. (2020). *Lost Opportunities: How Disparate School Discipline Continues to Drive Differences in the Opportunity to Learn*. Learning Policy Institute.
- Losen, D. J., Sun, W.-L., & Keith, M. A. (2017). Suspended Education in Massachusetts: Using Days of Lost Instruction Due to Suspension to Evaluate Our Schools. In *Civil Rights Project—Proyecto Derechos Civiles*. Civil Rights Project / Proyecto Derechos Civiles. <https://eric.ed.gov/?id=ED573327>
- Lüdecke, D. (2024). *sjPlot: Data Visualization for Statistics in Social Science*. <https://CRAN.R-project.org/package=sjPlot>
- Markowitz, D. M., Kittelman, A., Girvan, E. J., Santiago-Rosario, M. R., & McIntosh, K. (2023). Taking Note of Our Biases: How Language Patterns Reveal Bias Underlying the Use of Office Discipline Referrals in Exclusionary Discipline. *Educational Researcher*, 52(9), 525–534. <https://doi.org/10.3102/0013189X231189444>
- Mata, C. (2024). What are we learning from school suspension bans? *The Brookings Institute*. <https://www.brookings.edu/articles/what-are-we-learning-from-school-suspension-bans>

- May, S., Ard, W., Todd, A. W., Horner, R. H., Glasgow, A., Sugai, G., & Sprague, J. R. (2005). *School-wide information system*. Educational and Community Supports, University of Oregon.
- McConnell, S. (2022). Mind the Gap: Addressing the Disproportionate Rate of Discipline of Black Students in the Public Education System. *OUR Journal: ODU Undergraduate Research Journal*, 9. <https://doi.org/10.25778/ESAS-9Q42>
- McIntosh, K., Girvan, E. J., Fairbanks Falcon, S., McDaniel, S. C., Smolkowski, K., Bastable, E., Santiago-Rosario, M. R., Izzard, S., Austin, S. C., Nese, R. N. T., & Baldy, T. S. (2021a). Equity-focused PBIS approach reduces racial inequities in school discipline: A randomized controlled trial. *School Psychology*, 36(6), 433–444. <https://doi.org/10.1037/spq0000466>
- McIntosh, K., Girvan, E. J., McDaniel, S. C., Santiago-Rosario, M. R., St. Joseph, S., Fairbanks Falcon, S., Izzard, S., & Bastable, E. (2021b). Effects of an equity-focused PBIS approach to school improvement on exclusionary discipline and school climate. *Preventing School Failure: Alternative Education for Children and Youth*, 65(4), 354–361. <https://doi.org/10.1080/1045988X.2021.1937027>
- Meadan, H., Ayvazo, S., & Ostrosky, M. M. (2016). The ABCs of Challenging Behavior: Understanding Basic Concepts. *Young Exceptional Children*, 19(1), 3–15. <https://doi.org/10.1177/1096250614523969>
- Miller, C. E., & Meyers, S. A. (2015). Disparities in School Discipline Practices for Students with Emotional and Learning Disabilities and Autism. *Journal of Education and Human Development*, 4(1). <https://doi.org/10.15640/jehd.v4n1a23>
- Mittleman, J. (2018). Sexual Orientation and School Discipline: New Evidence From a Population-Based Sample. *Educational Researcher*, 47(3), 181–190. <https://doi.org/10.3102/0013189X17753123>
- Monahan, K. C., Vanderhei, S., Bechtold, J., & Cauffman, E. (2014). From the School Yard to the Squad Car: School Discipline, Truancy, and Arrest. *Journal of Youth and Adolescence*, 43(7), 1110–1122. <https://doi.org/10.1007/s10964-014-0103-1>
- Morgan, M. A., & Wright, J. P. (2018). Beyond Black and White: Suspension Disparities for Hispanic, Asian, and White Youth. *Criminal Justice Review*, 43(4), 377–398. <https://doi.org/10.1177/0734016817721293>
- Morris, E. W., & Perry, B. L. (2016). The Punishment Gap: School Suspension and Racial Disparities in Achievement. *Social Problems*, 63(1), 68–86.
- Morris, E. W., & Perry, B. L. (2017). Girls Behaving Badly? Race, Gender, and Subjective Evaluation in the Discipline of African American Girls. *Sociology of Education*, 90(2), 127–148. <https://doi.org/10.1177/0038040717694876>

- National Center for Education Statistics. (n.d.). *State Nonfiscal Survey of Public Elementary and Secondary Education* (Digest of Education Statistics 2023, p. Table 203.50). U.S. Department of Education. <https://nces.ed.gov/programs/coe/indicator/cge/racial-ethnic-enrollment>
- National Center for Education Statistics. (2024). *Percentage of students suspended and expelled from public elementary and secondary schools, by sex, race/ethnicity, and state or jurisdiction: School year 2020-2021*. U.S. Department of Education Office of Civil Rights. https://nces.ed.gov/programs/digest/d23/tables/dt23_233.40.asp
- Nese, R. N. T., Bastable, E., Gion, C., Massar, M., Nese, J. F. T., & McCroskey, C. (2020). Preliminary analysis of an instructional alternative to exclusionary discipline. *The Journal of At-Risk Issues*.
- Nese, R. N. T., Nese, J. F. T., McCroskey, C., Meng, P., Triplett, D., & Bastable, E. (2021). Moving away from disproportionate exclusionary discipline: Developing and utilizing a continuum of preventative and instructional supports. *Preventing School Failure: Alternative Education for Children and Youth*, 65(4), 301–311. <https://doi.org/10.1080/1045988X.2021.1937019>
- Nese, R. N. T., Santiago-Rosario, M. R., Malose, S., Hamilton, J., Nese, J. F. T., & Horner, R. (2022). Improving a universal intervention for reducing exclusionary discipline practices using student and teacher guidance. *Psychology in the Schools*, 59(10), 2042–2061.
- Nguyen, B. M. D., Noguera, P., Adkins, N., & Teranishi, R. T. (2019). Ethnic Discipline Gap: Unseen Dimensions of Racial Disproportionality in School Discipline. *American Educational Research Journal*, 56(5), 1973–2003. <https://doi.org/10.3102/0002831219833919>
- NH MTSS-B. (2021, September 22). *Defining Minor vs Major Student Behaviors*. NH MTSS-B Technical Assistance Center. <https://nhmtssb.org/wp-content/uploads/2023/02/Sample-Minor-vs.-Major-Behavior-Matrix-2021-09-22-br.pdf>
- Noltemeyer, 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>
- Okonofua, J. A., & Eberhardt, J. L. (2015). Two Strikes: Race and the Disciplining of Young Students. *Psychological Science*, 26(5), 617–624. <https://doi.org/10.1177/0956797615570365>
- Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences*, 113(19), 5221–5226. <https://doi.org/10.1073/pnas.1523698113>

- Owens, A., Reardon, S., Kalogrides, D., Jang, H., & Tom, T. (2022). Trend in Racial/Ethnic and Economic School Segregation, 1991-2020. *The Segregation Index, May*.
- Oyez. (n.d.). *Goss v. Lopez* [Online post]. Www.Oyez.Org. Retrieved February 3, 2025, from www.oyez.org/cases/1974/73-898
- PBIS Apps. (2024, January 17). *SWIS Referral Definitions*. PBIS Apps. <http://www.pbisapps.org/resource/referral-category-definitions>
- Pearman, F. A., Curran, F. C., Fisher, B., & Gardella, J. (2019). Are Achievement Gaps Related to Discipline Gaps? Evidence From National Data. *AERA Open*, 5(4). <https://doi.org/10.1177/2332858419875440>
- Perry, B. L., & Morris, E. W. (2014). Suspending Progress: Collateral Consequences of Exclusionary Punishment in Public Schools. *American Sociological Review*, 79(6), 1067–1087. <https://doi.org/10.1177/0003122414556308>
- Pimentel-Mannan, I. A., Nese, J. F. T., Newson, A., Kjellstrand, J., & Nese, R. N. T. (2024). Addressing discipline equity: The Inclusive Skill-Building Learning Approach (ISLA) an alternative to exclusionary discipline. *Preventing School Failure: Alternative Education for Children and Youth*, 68(4), 318–329. <https://doi.org/10.1080/1045988X.2023.2252761>
- R Core Team. (2024). *R: A Language and Environment for Statistical Computing* (4.4.1). R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reny, T., & Manzano, S. (2016). The Negative Effects of Mass Media Stereotypes of Latinos and Immigrants. *Media and Minorities*, 195–212. <https://doi.org/10.13109/9783666300882.195>
- 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>
- Sabol, T. J., Kessler, C. L., Rogers, L. O., Petitcherc, A., Silver, J., Briggs-Gowan, M., & Wakschlag, L. S. (2022). A window into racial and socioeconomic status disparities in preschool disciplinary action using developmental methodology. *Annals of the New York Academy of Sciences*, 1508(1), 123–136. <https://doi.org/10.1111/nyas.14687>
- Samimi, C., Buraik, L., Oteman, Q., & Marsalis, S. (2024). What is Known About LGBTQ Youth and Exclusionary Discipline? A Scoping Review. *Journal of Homosexuality*, 1–26. <https://doi.org/10.1080/00918369.2024.2420077>
- Samimi, C., Jefferson, N., Flanagan, S., & Anyon, Y. (2023). Intersections of Disproportion: A Critical Quantitative Examination of Dis/ability and Gender in Black Students' School Discipline Outcomes. *The Urban Review*, 55(4), 456–475. <https://doi.org/10.1007/s11256-023-00657-6>

- Samuels, W. E., Tournaki, N., Blackman, S., & Zilinski, C. (2016). Executive functioning predicts academic achievement in middle school: A four-year longitudinal study. *Journal of Educational Research*, 109(5), 478–490.
- Santiago-Rosario, M. R., Whitcomb, S. A., Pearlman, J., & McIntosh, K. (2021). Associations between teacher expectations and racial disproportionality in discipline referrals. *Journal of School Psychology*, 85, 80–93. <https://doi.org/10.1016/j.jsp.2021.02.004>
- Schiff, M. (2018). Can restorative justice disrupt the ‘school-to-prison pipeline?’ *Contemporary Justice Review*, 21(2), 121–139. <https://doi.org/10.1080/10282580.2018.1455509>
- Schollenberger, T. L. (2015). Racial disparities in school suspension and subsequent outcomes: Evidence from the nation longitudinal survey of youth. In *Closing the school discipline gap: Research for policymakers* (pp. 31–43). Teachers College Press.
- 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. <https://doi.org/10.3102/0002831214541670>
- Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race Is Not Neutral: A National Investigation of African American and Latino Disproportionality in School Discipline. *School Psychology Review*, 40(1), 85–107. <https://doi.org/10.1080/02796015.2011.12087730>
- Skiba, R. J., 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*, 34(4).
- Smith, L. H., Bottiani, J. H., Kush, J. M., & Bradshaw, C. P. (2023). The discipline gap in context: The role of school racial and ethnic diversity and within school positionality on out-of-school suspensions. *Journal of School Psychology*, 98, 61–77. <https://doi.org/10.1016/j.jsp.2023.02.006>
- Sone, S. (2023). *The Alarming Legacy of the 1994 Gun-Free Schools Act*. The Trace. <https://www.thetrace.org/newsletter/gun-free-schools-act-legacy-punishment/>
- St-Amand, J., Smith, J., & Rasmy, A. (2022). Empirical Validation of a Model for Predicting Students’ Sense of Belonging and School Engagement as a Function of Classroom Management Practices. *International Journal of Educational Psychology*, 11(2). <https://doi.org/10.17583/ijep.7508>
- Sugai, G., & Horner, R. (2002). The Evolution of Discipline Practices: School-Wide Positive Behavior Supports. *Child & Family Behavior Therapy*, 24(1–2), 23–50. https://doi.org/10.1300/J019v24n01_03

- Sugai, G., & Horner, R. H. (2020). Sustaining and Scaling Positive Behavioral Interventions and Supports: Implementation Drivers, Outcomes, and Considerations. *Exceptional Children*, 86(2), 120–136. <https://doi.org/10.1177/0014402919855331>
- Sullivan, A. L., Van Norman, E. R., & Klingbeil, D. A. (2014). Exclusionary Discipline of Students With Disabilities: Student and School Characteristics Predicting Suspension. *Remedial and Special Education*, 35(4), 199–210. <https://doi.org/10.1177/0741932513519825>
- Tierney, N., & Cook, D. (2023). Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations. *Journal of Statistical Software*, 105(7), 1–31. <https://doi.org/10.18637/jss.v105.i07>
- Vera, E. M., Polanin, J. R., Polanin, M., & Carr, A. L. (2018). Keeping Latina/o students in school: Factors related to academic intentions of students at risk for school dropout. *Journal of Latina/o Psychology*, 6(1), 34–48. <https://doi.org/10.1037/lat0000082>
- Wang, M.-T., Scanlon, C. L., & Del Toro, J. (2023). Does anyone benefit from exclusionary discipline? An exploration on the direct and vicarious links between suspensions for minor infraction and adolescents' academic achievement. *American Psychologist*, 78(1), 20–35. <https://doi.org/10.1037/amp0001030>
- Welsh, R. O., & Little, S. (2018). The School Discipline Dilemma: A Comprehensive Review of Disparities and Alternative Approaches. *Review of Educational Research*, 88(5), 752–794. <https://doi.org/10.3102/0034654318791582>
- Welsh, R. O., & Rodriguez, L. A. (2024). The Plight of Persistently Disciplined Students: Examining Frequent Flyers and the Conversion of Office Discipline Referrals Into Suspensions. *Educational Evaluation and Policy Analysis*, 46(1), 160–170. <https://doi.org/10.3102/01623737231155155>
- Whisman, A., & Hammer, P. (2014). The association between school discipline and academic performance: A case for positive discipline approaches. *West Virginia Board of Education*.
- Wilkerson, K. L., & Afacan, K. (2022). Repeated School Suspensions: Who Receives Them, What Reasons Are Given, and How Students Fare. *Education and Urban Society*, 54(3), 249–267. <https://doi.org/10.1177/00131245211009854>
- Wu, E. (2014). *The color of success: Asian Americans and the origins of the model minority*. Princeton university press.
- Zakszeski, B., Rutherford, L., Heidelberg, K., & Thomas, L. (2021). In pursuit of equity: Discipline disproportionality and SWPBIS implementation in urban schools. *School Psychology*, 36(2), 122–130. <https://doi.org/10.1037/spq0000428>

Zimmermann, C. R. (2018). The Penalty of Being a Young Black Girl: Kindergarten Teachers' Perceptions of Children's Problem Behaviors and Student–Teacher Conflict by the Intersection of Race and Gender. *The Journal of Negro Education*, 87(2), 154. <https://doi.org/10.7709/jnegroeducation.87.2.0154>

Zimmermann, C. R. (2024). Looking for Trouble: How Teachers' Racialized Practices Perpetuate Discipline Inequities in Early Childhood. *Sociology of Education*, 97(3), 219–232. <https://doi.org/10.1177/00380407241228581>