



**The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:**

**Document Title:** Mapping Decision Points from School Based Incidents to Exclusionary Discipline, Arrest and Referral to the Juvenile Justice System

**Author(s):** Moriah Taylor, Jeremy Bilfield, Zoe Livengood, Melissa Sickmund

**Document Number:** 254625

**Date Received:** April 2020

**Award Number:** 2016-CK-BX-0018

**This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.**

**Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.**

**Final Summary Overview: Mapping Decision Points from School Based Incidents to Exclusionary Discipline, Arrest and Referral to the Juvenile Justice System (NIJ grant #2016-CK-BX-0018)**

Submitted by Moriah Taylor, Jeremy Bilfield, Zoe Livengood, and Melissa Sickmund  
National Center for Juvenile Justice, research division of the  
National Council of Juvenile and Family Court Judges

September 2019

**Purpose**

Policy and practice changes in the 1990s such as The Gun-Free Schools Act (1994) and widespread adoption of zero-tolerance policies introduced an era of over-reliance on exclusionary discipline and interventions from sources outside of the school (Marsh, 2014). These changes unintentionally introduced more low-risk youth to the justice system (Education Development Center, 2012). Since then, there have been efforts to address overreliance on exclusionary discipline and we have seen a national decline in the rates of exclusionary discipline (Bracey, Geib, Plant, O’Leary, Anderson, Herscovitch, O’Connell, & Vanderploeg, 2013; Teske, Huff, & Graves, 2013; U.S. Department of Education, Office for Civil Rights, 2016; U.S. Department of Education, Office for Civil Rights, 2018). However, there is still more work to be done. Responses to student behaviors can serve to help youth improve their decisionmaking capabilities and maintain student safety, but research indicates that responses to more subjective behaviors are administered disproportionately to youth of color, economically disadvantaged youth, and youth with developmental needs (Smolkowski, Girvan, McIntosh, Nese, & Horner, 2016) The overrepresentation of youth of color, especially black youth, as recipients of exclusionary discipline is well documented (e.g., Blake, Butler, Lewis & Darensbourg, 2011; Fabelo et al., 2011; U.S. Department of Education, 2014). More information is needed to understand how decisions to address school-based incidents are made.

The purpose of this two-phase study was to examine factors influencing decisions that lead from a school-based incident to exclusionary discipline, an arrest, and a referral to the juvenile court. We hypothesize that multiple external factors such as race identity, socioeconomic status and others will

negatively impact the decisionmaker, and generate harsher punishments for those who are in these vulnerable groups.

## **Project Design and Methods**

Phase 1 involved interviewing groups of key stakeholders including school administrators, district administrators, discipline coordinators, juvenile court judges and other staff, law enforcement officers, Positive Behavior Intervention Support coordinators, and child welfare agencies to understand their approaches to behavior management. Phase 2 involved secondary analysis of data from a local school district(s) and the juvenile court with jurisdiction in two counties.

### **Phase 1: Mapping Decisionmaking Practices**

Researchers conducted structured interviews with a convenience sample of 10 jurisdictions across the U.S. between October 2017 and March 2018 to gather information concerning their decisionmaking practices for school behavior incidents. Participants were asked about the information and policies that form the basis for discipline decisions, how decisions are documented, and the extent to which data on the decisions are collected and reviewed. To achieve a more complete understanding of each jurisdiction's decisionmaking procedures researchers supplemented the structured interview data with a systematic review of each district's behavior policy to extract information regarding behavior management policies. The school and community climate and practices vary widely, and while there are some common themes identified in this research, the findings are not able to be generalized and attributed to national approaches to school behavior management.

For each jurisdiction, researchers used the information from the structured interviews to create diagrams outlining the sequence of potential decisions involved in responding to behavior incidents. The research team authored a report describing the findings from the structured interviews (Taylor, Deal, Ehrman, and Siegel, 2019).

## **Phase 2: Analysis of Discipline Records**

The goal of Phase 2 was to obtain secondary data from school districts and courts with juvenile jurisdiction to further explore the factors that may lead to exclusionary discipline, potential arrest, and referral to court. The analyses focused on behavior incident records for middle and high schools in each district. NCJJ acquired data from school districts and courts in two different counties. For ease of delineation, each site is discussed separately.

**Data Collection: Site One.** Three school districts (15 schools with 10,823 students in the 2016-17 school year) in one county agreed to share their discipline data. The data file contained 9,922 incident records between the 2011-2018 school years. Researchers selected incidents occurring during the study period (2016-17 and 2017-18 school years) and specified the unit of count as a student with a discipline record on a day, selecting the most severe behavior. This provided 4,519 records to review. Similarly, the juvenile court provided a data file with all of the records referred by School Resource Officers within the same time frame. The court provided 45 records. Researchers were tasked with linking school behavior records with court records, but were only able to do so for 11 records. After learning about inconsistencies in data collection practices, we were limited to working with one school district's data and due to sample size, were unable to conduct any meaningful analysis.

**Data Collection: Site Two.** Researchers received data from one school district (26,231 students in the 2016-17 school year) and the court with juvenile jurisdiction. The school district data was shared with researchers at the state Administrative Office of the Courts who identified all of the court records for the youth in the school file and provided a data file that combined school discipline/juvenile court delinquency cases for the study. The unit of count was a student with at least one behavior incident on a day. Each student incident on a day was characterized by the most severe behavior documented for that day. We also removed any incidents involving only attendance issues. These refining steps left providing 10,726 school records to review (Appx. Table 1). The court provided 5,896 court records for the students in the education file. This includes any students' history with the court prior to the incident. To identify court cases that were the most likely impetus for a court case, researchers identified court cases that were

petitioned within +/- 14 days of a school incident and manually compared offenses to determine if they were matches. (See Appx. Tables 1 and 2 and Figure 1 for descriptive statistics.) Only 84 court cases were identified as the result of a school offense.

## **Data Analysis and Findings**

We applied paired student t-testing and Bonferroni corrections to determine if there were differences in the demographic characteristics of the school incident sample, the matched court case sample, and the school student population. We found a significant difference in the number of low socioeconomic status (SES) individuals and individuals with prior court interaction between the court and school incident samples, both with  $p < 0.005$  (Appx. Table 3). Similarly, we found a significant difference in the racial make-up of the student population in comparison to both sample groups, with  $p$ 's  $< 0.005$  (Appx. Table 4). These results indicate that there are more non-white students in both the incident sample and in the court sample, and further that there is a significantly greater than average number of individuals with low SES and prior court experience in the court sample than in the school sample. In addition to these tests, more appropriate Kruskal-Wallis tests were applied to look at differences in ordinal values related to offenses and punishments received between white and non-white students in our samples (Appx. Table 5). However, no axis of observation was significant at  $p < 0.05$  either before or after the Bonferroni correction.

To move beyond detecting directly observable differences in the treatment of students, we created several Bayesian ordinal regression models to predict the variable  $y^*$  (the mental assessment of a school administrator), which then predicts the level of punishment given to a student as a response to an incident where  $y^*$  is predicted by creating parameters defined by several indicator variables, including racial minority, non-male (female or other gender expression), low SES, homelessness, special education requirements, and prior court interaction. These models also used the total number of incidents a youth had on that day as well as the severity of the incident. Each model was designed with a clear philosophy motivating it (Appx. Table 6).

After the fitting of 12 Bayesian models to the testing set of the data, all parameters were found to have converged around a value with an R-hat of 1.00. Leave One Out (LOO) testing shows that the Pareto K values were at least ok or  $k < 0.7$ , though the majority (nearly 99%) were good ( $k < 0.5$ ), and those few that were bad ( $k > 0.7$ ) were re-fit excluding those k values as outliers. Furthermore, measures of the predicted number of parameters against the actual number of parameters were near equal or less than 0. All of which is to say that the models were appropriately fit to the data itself (Appx. Table 7).

In addition, each model was tested for accuracy against a testing set to prevent overfitting and as an additional method of model analysis. Based on these criteria, we believed that the model which best fit the data was model #3 or the model which contained interaction between the parameters and measured grouping based on incident severity (or the intercepts/breakpoints varied between these groups).

As the model of best fit has been established, there is a wide variety of information that we are able to glean from the model of best fit. First, we cannot clearly delineate when specific responses such as out-of-school-suspension (OSS) or no action will be administered, indicating that the decision whether to punish or not for most incidents seems to be nearly arbitrary and varies on the particular discretion of the decisionmaker (Appx. Figure 2).

Second, the school district perceives drug possession offenses as more severe than the juvenile justice system does, as we see the severity order change. Drug use and possession is treated similarly to drug trafficking by school decisionmakers. It is clear that the models had difficulty separating the instances of punishment in a way that is visible statistically (Appx. Figure 3).

Third, student demographics may play a role in determining their punishment. For instance, for first time offenders, white students receive a protective effect from their total number of incidents on a day, where non-whites don't (Appx. Figure 4). In addition, we see first time white students receive a protective effect from having prior interaction with the courts, whereas non-white students don't. However, non-white students with low SES seem to receive a less severe punishment than their white counterparts with low SES. Students (both white and non-white) without prior school behavior records in

our data, are less likely to receive more severe responses if they are homeless or have had prior court interaction.

We also used a model to assess differences between responses for students without prior school discipline records in the data (first-timer sample) and found that the influence of race on the determination of punishment is subtle (Appx. Figure 5):

1. Non-white students are more likely to receive more severe responses simply for being non-white:
2. Non-white students see a harmful effect in homelessness compared to their white peers:
3. White students with more total incidents on a day or prior court interaction are less likely to receive a severe response than their non-white peers with the same characteristics.

Not only do we see a directly negative effect within the first-timer sample, but also a variety of interactions with multiple parameters with the net result being that it appears that white students receive more leniency than students of color. Even the benefit conferred by an individual being merely “sent to the office” for their first incident is lost to students of color. We see this effect carry through in the descriptive aspects of our data, and may at least partially imply why there is a disparity in punishment that is not proportional to the racial breakdown of the greater school district.

### **Implications for Policy and Practice**

Teachers and administrators should actively seek better ways to manage student behavior, especially non-compliant behavior. The policy of the school district and the state laws explicitly state that exclusionary discipline should be used as little as possible unless other corrective action has not worked, but still permit exclusionary methods for behaviors that do not threaten the safety of others. Many districts have worked with programs like Positive Behavior Intervention and Supports to implement practices to effectively manage behavior and the study site is currently doing just that. It is imperative to train teachers and give them the necessary tools to appropriately handle misbehavior. All school personnel should understand the potential adverse effects of exclusionary discipline. Cogshall et al. (2012) suggests that teachers should receive training in regards to brain development, culture, trauma, and health and how these factors affect students’ behavior. Other researchers suggest using restorative justice

practices (Fishman & Hack, 2012; Smolkowski et al., 2016; Teasley, 2014; Thompson, 2016), positive behavior supports (Clonan et al., 2007; Feuerborn & Tyre, 2016; Thompson, 2016), and student-teacher relationships (Blake et al., 2011; Coggshall et al., 2012; Jones et al., 2013) in place of exclusionary discipline practices. School-based interventions should be focussed on support for teachers and helping them respond to youth in appropriate ways, dissuading reporting youth for minor insubordination.

Ideally, national school discipline data would be standardized, with all schools receiving training on what, when, and how to enter data properly. Efforts should be made to improve data collection and use in order to ensure that discipline is administered fairly and that non-compliance is responded to appropriately. The Department of Education collects discipline for the Civil Rights Data Collection, but school level documentation is inconsistent. There needs to be very clearly documented instructions for schools and those entering data to improve the quality of the data. Although resources exist to accomplish this, it is unclear if those entering data are trained in using these resources. Improving school discipline data will also require schools to establish clear data collection policies, provide training to enter data, and review their data routinely. Once data are collected, administrators can use the data to have a better idea of how they discipline and how this affects students. The goal should be to move away from exclusionary discipline as much as possible in favor of more useful ways of managing student behavior.

### **Limitations and Future Research**

Although the findings of this study are informative and have practical implications there are limitations that can be addressed in future research that will allow for a more comprehensive understanding of variation in decisionmaking practices

**Data from more sites is required.** This study only included one school district's data because the other sites we intended to include did not use definitions of in-school suspension or collect data on behaviors that received non-exclusionary responses consistently. In order for this type of study to occur, there needs to be a way to ensure that school and district wide data are collected in a manner that lends to comparison and study across districts (Skiba, 2013). Including more variables, such as time of day and teacher/administrator attributes would also be important information for future research.



**Data on all behavior incidents and responses should be documented.** We could only analyze the behaviors and responses that were documented limiting our insight into other minor incidents and informal discipline. We also have limited contextual information about the practices in the sample site and the complexities of that school system. This could be mediated by conducting qualitative analysis with sites contributing secondary data to further understand practices. A longitudinal study focused on a large number of students in different districts with different backgrounds might offer some insight about how school discipline is approached in schools around the country and how school discipline affects students long-term.

**Indicators of court involvement because of a school incident should be tracked.** Because our data had too few court records that derived from school incidents, we were not able to draw many conclusions based on this information. Though it is a positive thing that there were not many cases to analyze, sampling more school districts and the courts with jurisdiction would provide insight into school discipline and subsequent court involvement.

## References

- Blake, J. J., Butler, B. R., Lewis, C. W., & Darnesbourg, A. (2011). Unmasking the inequitable discipline experiences of urban black girls: Implications for urban educational stakeholders. *The Urban Review*, 43, 90-106. doi: 10.1007/s11256-009-0148-8
- Bracey, J. R., Foley Geib, C., Plant, R., O'Leary, J. R., Anderson, A., Herscovitch, L., O'Connell, M., & Vanderploeg, J. J. (2013). Connecticut's comprehensive approach to reducing in-school arrests: Changes in statewide policy, systems coordination and school practices. *Family Court Review*, 51(3), 427-434. doi: 10.1111/fcre.12039
- Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement*. Retrieved from The Council of State Government's website: <https://csgjusticecenter.org/youth/breaking-schools-rules-report/>
- Kruskal, W., & Wallis, W. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260), 583-621. doi:10.2307/2280779
- Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. N. T., & Horner, R. H. (2016). Vulnerable decision points for disproportionate office discipline referrals: Comparisons of discipline for African American and White elementary school students. *Behavioral Disorders*, 41(4), 178-195. doi: 10.17988/bedi-41-04-178-195.1
- Taylor, M.; Deal, T.; Ehrman, S.; and Siegel, G. (2019). *Mapping Decision Points from School-based Incidents to Exclusionary Discipline, Arrest and Referral to the Juvenile Justice System: An Analysis of Mapping Interviews*. Pittsburgh, PA: National Center for Juvenile Justice, National Council of Juvenile and Family Court Judges.
- <http://www.ncjj.org/publication/Mapping-Decision-Points-from-School-based-Incidents-to-Exclusionary-Discipline-Arrest-and-Referral-to-the-Juvenile-Justice-System-An-Analysis-of-Mapping-Interviews.aspx>

- Teske, S. C., Huff, B., & Graves, C. (2013). Collaborative role of courts in promoting outcomes for students: The relationship between arrests, graduation rates, and school safety. *Family Court Review*, 51(3), 418-426. doi: 10.1111/fcre.12038
- U.S. Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS). *Public School Data File, 2007-08*. Retrieved from [https://nces.ed.gov/surveys/sass/tables/sass0708\\_035\\_s1s.asp](https://nces.ed.gov/surveys/sass/tables/sass0708_035_s1s.asp)
- U.S. Department of Education Office for Civil Rights. (2014, March). *Data snapshot: school discipline, Civil Rights Data Collection* (Issue Brief No. 1). Retrieved from <http://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf>.
- U.S. Department of Education Office for Civil Rights. (2016). A first look, 2013-14 Civil Rights Data Collection. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf>
- U.S. Department of Education Office for Civil Rights. (2018). Master List of 2015 - 2016 CRDC Definitions. Retrieved from <https://ocrdata.ed.gov/DataDefinitions>
- U.S. Department of Education Office for Civil Rights. (2018). School climate and safety, 2015–16 Civil Rights Data Collection. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/docs/school-climate-and-safety.pdf>

# Appendix

**Table 1**

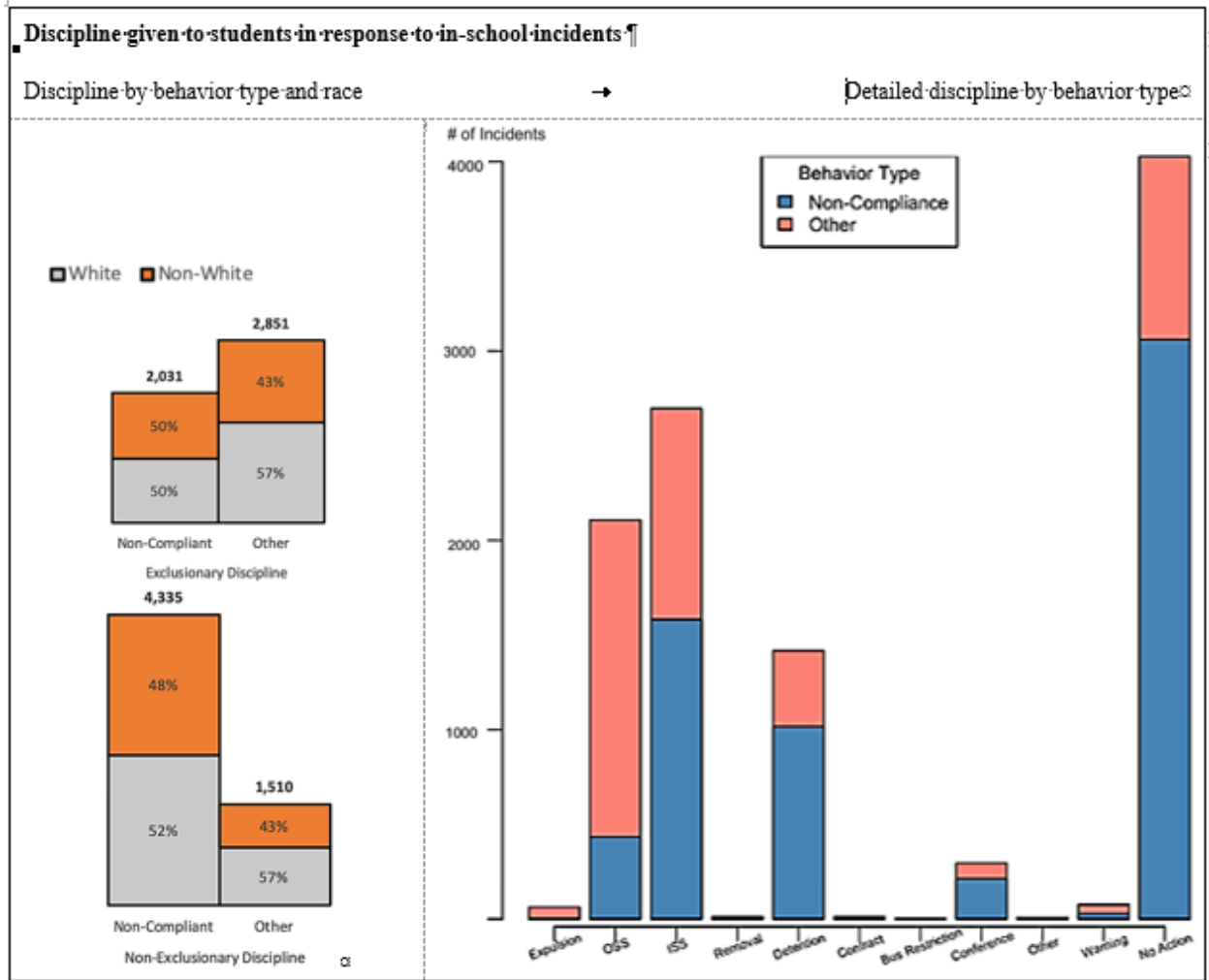
Number of youth in the general student population of the school district and five different sample groups												
	Total	Male	Non-male	White	Non-white	Homeless	Low-SES	Special-Ed	Court-priors	Mode incidents on a day	Mode incident level	Mode discipline level
General population	25,080	12,545	12,535	17,695	7,385	—	—	—	—	—	—	—
Full sample	10,726	8,095	2,628	5,692	5,034	749	6,914	3,382	2,102	1	5	11
1 <sup>st</sup> incident	5,220	3,762	1,457	2,766	2,454	333	3,203	1,355	894	1	5	11
Court incidents	84	65	19	40	44	16	73	17	34	1	3	2
White	5,692	4,401	1,288	5,692	—	326	3,114	1,952	1,205	1	6	11
Nonwhite	5,034	3,694	1,340	—	5,034	423	3,800	1,430	897	1	6	11

Notes: Column listings include Total Records (count of all records), and Counts for Gender Identity, Racial Identifier, Homelessness, Low-SES Status, Special Ed, and Court Priors for these youth. This is to say exactly how many records meet those categories in the data. Furthermore, we have included the average number of total offenses for each sample division, as well as the most common incident level and most common discipline level.

**Table 2**

Ratios for the groups of interest in the sample							
	Total	Male:↵ Non-male	White:↵ Non-white	Homeless:↵ Homeless	HighSES:↵ LowSES	None:↵ Special-Ed	None:↵ Court-Priors
General population	25,080	1.0:1.0	2.40:1	—	—	—	—
Full sample	10,726	3.08:1	1.13:1	13.32:1	0.55:1	2.17:1	4.10:1
1 <sup>st</sup> incident	5,220	2.58:1	1.13:1	14.68:1	0.63:1	2.85:1	4.84:1
Court incidents	84	3.42:1	0.91:1	4.25:1	0.15:1	3.94:1	1.47:1
White	5,692	3.42:1	—	16.46:1	0.83:1	1.92:1	3.72:1
Nonwhite	5,034	2.76:1	—	10.90:1	0.32:1	2.52:1	4.61:1

**Figure 1**



**Table 3**

<b>Full-Sample characteristics compared with Court-Sample</b>						
	<b>Race-Minority</b>	<b>Gender-Minority</b>	<b>Homeless</b>	<b>Low-SES</b>	<b>Special-Ed</b>	<b>Court-Priors</b>
Mean difference	-0.164	-0.073	-0.207	0.299	-0.248	-0.316
Interval	0.054	0.111	-0.035	-0.150	0.201	-0.101
T-value	-0.990	0.414	-2.795	-6.015	2.548	-3.866
P-value	0.325	0.680	0.006*	0.000*	0.013*	0.000*
Bonferroni	1	1	0.773	0.000*	0.152	0.003*
*p<0.05						

**Table 4**

<b>Ratio of white to non-white students Full-Sample and Court-Sample compared with general student population</b>		
	<b>Full-Sample</b>	<b>Court-Sample</b>
Mean difference	0.169	0.120
Interval	0.186	0.339
T-value	31.154	4.178
P-value	0.000*	0.000*
Bonferroni	0.000*	0.009*
*p<0.05		

**Table 5**

<b>Offense seriousness and discipline severity for white compared with non-white students in Full and 1st-Time Samples</b>				
	<b>Full-Sample</b>		<b>1st-Time-Sample</b>	
	<b>Offense</b>	<b>Discipline</b>	<b>Offense</b>	<b>Discipline</b>
Chi-squared	3.117	2.600	2.036	0.979
P-value	0.077	0.107	0.154	0.322
Bonferroni	0.930	1	1	1
*p<0.05				

**Table 6**

<p><b>Parameters &amp; Variables of Analysis</b> - Unobserved assessment in decisionmaker’s mind (<math>y^*</math>), breakpoints/intercepts, racial minority status, low SES status, non-male status, special education requirements, homelessness, prior court interaction, incident severity, total number of incidents, first incidents within the sample time frame, specific racial/ethnic group.</p>
<p><b>Model 1</b> - shows the relationship between the relevant indicator parameters and simplex/monotonic parameters with the predicted variable in a way that makes the least amount of assumptions on the unobserved <math>y^*</math>.</p>
<p><b>Model 2</b> - attempts to measure potential interaction terms between the racial minority variable and the first incident recorded during the sample time frame variable in the case of the full training sample.</p>
<p><b>Model 3</b> - removes incident level as a monotonic predictor, treating it as a grouping variable where each incident level has unique intercepts, meaning <math>y^*</math> different intercepts/breakpoints based on the offense level.</p>
<p><b>Model 4</b> - Model 3, but instead of grouping based on offense, it groups on specific race or ethnicity data and removes racial minority indicators.</p>
<p><b>Models 5 &amp; 6</b> - shows a situation in which the only provided information was incident-related information (5) and non-incident-related information (6). This allows us to examine a situation in which the decisionmaker’s <math>y^*</math> was only influenced by either types of information.</p>

**Table 7**

Model Fits to Full Sample and First-time Sample					
Models	ELPD-diff-LOO	p-LOO--actual	Percentage Correct	Percentage More Severe	Percentage Less Severe
<b>Full Sample</b>					
3. w/ interaction, group offense	0.00	-25.577	41.392%	16.035%	42.573%
2. w/interaction, no grouping	-14.620	-37.430	41.361	15.258	43.381
4. w/interaction, group race	-35.400	-12.961	41.765	15.475	42.759
1. w/o interaction, no grouping	-46.489	-9.992	41.206	15.786	43.008
5. w/only offense info	-67.538	-10.053	41.361	15.507	43.132
6. w/o offense info	-586.748	-1.107	36.855	1.150	61.995
<b>First-Time Sample</b>					
3. w/ interaction, group offense	0.000	-16.386	48.054%	6.254%	45.692%
4. w/interaction, group race	-3.288	-14.192	47.224	6.445	46.331
1. w/o interaction, no grouping	-8.569	-11.123	46.650	6.382	46.969
5. w/only offense info	-8.662	-11.133	46.075	5.424	48.500
2. w/interaction, no grouping	-10.830	-23.821	47.352	6.190	46.458
6. w/o offense info	-249.211	-1.199	45.246	0.000	54.754
<p><b>Note:</b> This figure lists the ELPD difference LOO; a measure of model—predicted number versus actual number of parameters; a measure of model validity—percentage of correct model predictions; and percentage of incorrect predictions that were more severe or less severe than the actual punishment given.</p>					

Figure 2

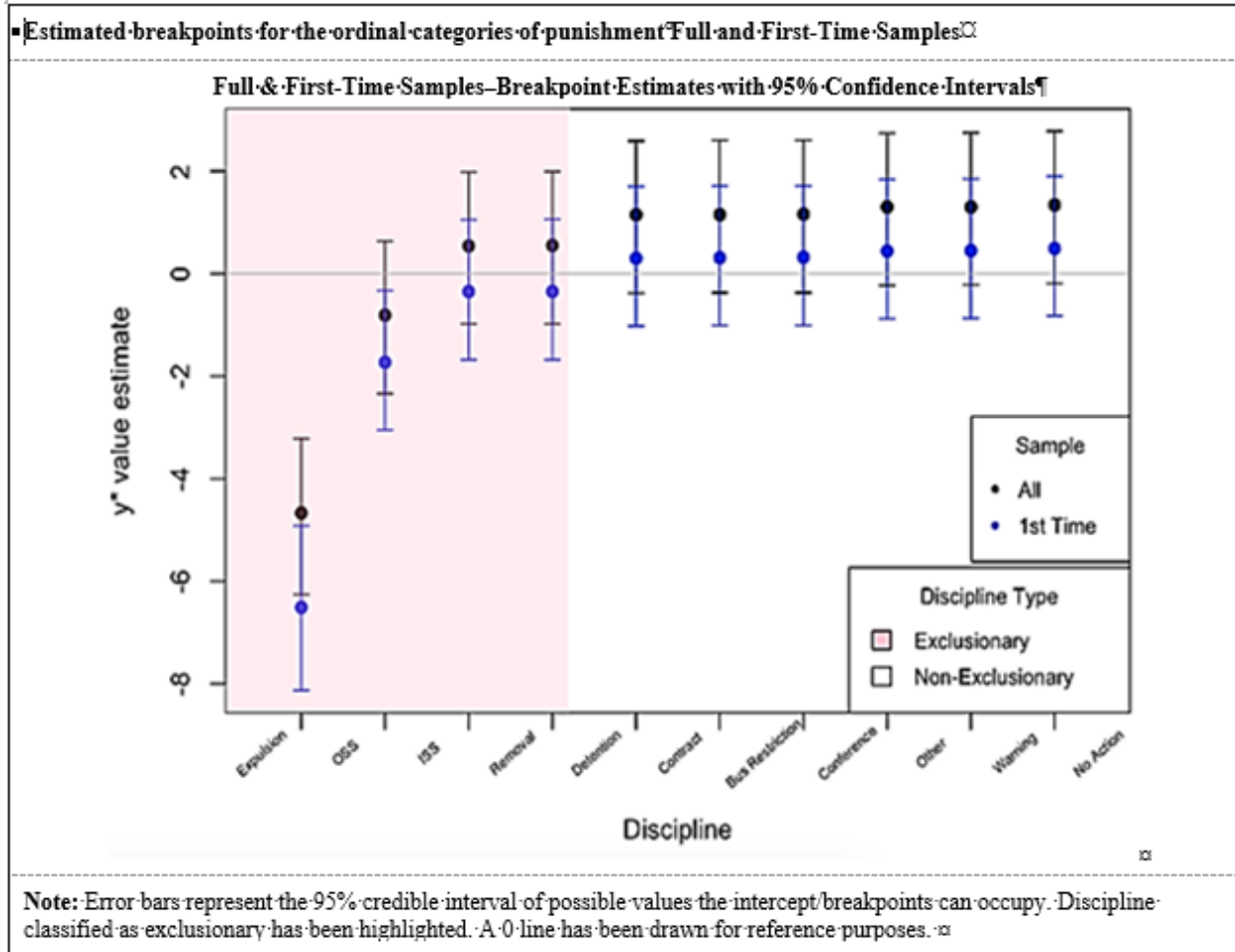




Figure 3

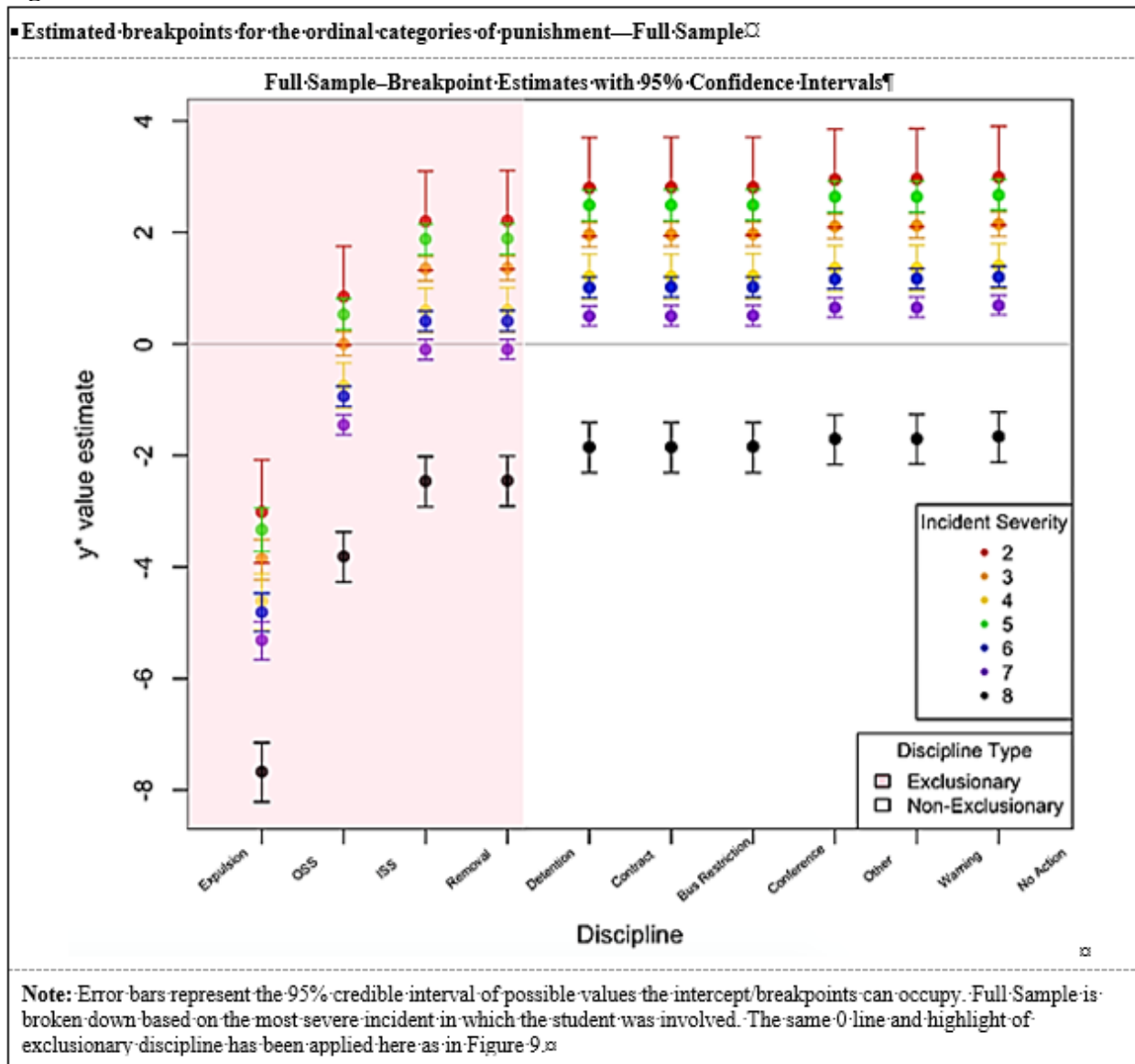


Figure 4

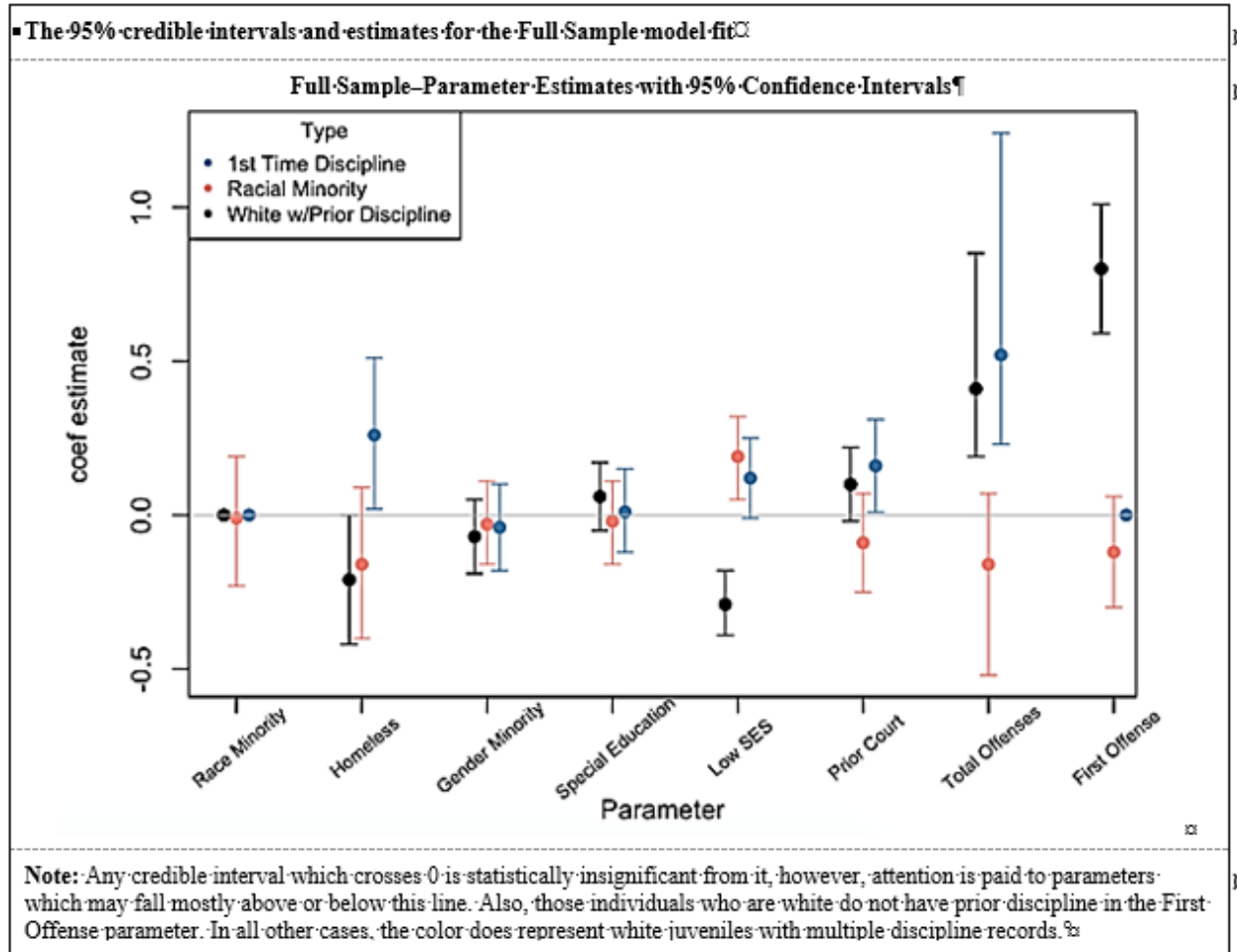


Figure 5

