Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 25-1175

Education Governance and Race: An Analysis of School Board Discourse Using Large Language Models

Kylie L. Anglin
University of Connecticut

Alvin Christian
University of Michigan

Brian A. Jacob
University of Michigan

John D. Singleton
University of Rochester

Despite growing attention to school boards, it is unclear whether they primarily operate as bureaucratic forums, policy-making bodies, or arenas for contentious debate—particularly on issues of race. Recent controversies suggest increasing public engagement and conflict, but little evidence documents how often questions of race arise in board deliberations. This study analyzes over 40,000 meeting minutes from 2018–2022 to examine the prevalence, framing, and drivers of race-related discourse. Using large language models, natural language processing, and human coding, we find that race-related conversations are relatively uncommon but responsive to national events, particularly in politically competitive, suburban districts. Our analysis highlights the variable nature of local governance and the value of meeting minutes and computational tools for understanding this variability.

VERSION: April 2025

Suggested citation: Anglin, Kylie L., Alvin Christian, Brian A. Jacob, and John D. Singleton. (2025). Education Governance and Race: An Analysis of School Board Discourse Using Large Language Models. (EdWorkingPaper: 25 -1175). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/ew42-t954

Education Governance and Race: An Analysis of School Board Discourse Using Large Language Models

Kylie L. Anglin[1], Alvin Christian[2], Brian Jacob[2] & John Singleton[3]

[1]University of Connecticut

[2]University of Michigan

[3]University of Rochester

April 2025

Despite growing attention to school boards, it is unclear whether they primarily operate as bureaucratic forums, policy-making bodies, or arenas for contentious debate—particularly on issues of race. Recent controversies suggest increasing public engagement and conflict, but little evidence documents how often questions of race arise in board deliberations. This study analyzes over 40,000 meeting minutes from 2018–2022 to examine the prevalence, framing, and drivers of race-related discourse. Using large language models, natural language processing, and human coding, we find that race-related conversations are relatively uncommon but responsive to national events, particularly in politically competitive, suburban districts. Our analysis highlights the variable nature of local governance and the value of meeting minutes and computational tools for understanding this variability.

Education Governance and Race: An Analysis of School Board Discourse Using Large Language Models

In recent years, American school board meetings have seemingly transformed from sleepy, bureaucratic affairs into arenas of contentious debate. In 2021, for example, when more than 250 individuals signed up to speak at a school board meeting in Loudoun County, Virginia, the meeting ended prematurely, with one man arrested, another ticketed for trespassing, and a third receiving a minor injury (Wilder et al., 2021). Similarly, in Palm Beach, Florida, meeting disruptions in one school district became so routine that school board members voted to ban "shouting, heckling, jeering, hissing, booing, engaging in speech that defames individuals or stymies or blocks meeting progress or loud, excessive or prolonged applause that disrupts the meeting" (Kokal, 2023). Race and diversity have been at the center of these controversies, with critical race theory (CRT) emerging as a particularly polarizing issue. Indeed, in October 2021, the National School Board Association released a letter requesting federal law enforcement to protect school board members who had received threats due to "propaganda purporting the false inclusion of critical race theory within classroom instruction and curricula" (NSBA, 2021, p. 1). Concerns about CRT continue to receive national attention, with President Donald Trump, for example, recently pledging to cut federal funding for schools "pushing critical race theory" (Klein, 2024). If such events reflect patterns in U.S. school districts more generally, then there is good reason to believe that we have again reached a moment in national education politics where race has become a major issue and the subject of widespread conflict.

Despite the recent national attention directed towards school boards and race, however, minimal empirical evidence exists concerning how and the extent to which school boards engage in discussions of racial equity, diversity, and systemic racism. In part, this is because systematic data on school boards are notably sparse. School boards constitute the single largest group of elected officials in the United States, consisting of nearly 90,000 lay members who oversee the education of 50 million children and bear broad responsibilities for district governance. Despite this, there are currently no nationally representative datasets that focus on the "work" school boards perform—such as the topics school boards discuss, how they interact with parents, teachers, administrators and outside actors, or what decisions and actions they take. As a result, prior scholarship on school boards has variously relied on small-scale, qualitative studies (e.g., Sutherland, 2022), infrequent larger-scale surveys (e.g., National School Board Association, 2018), or readily observable downstream outcomes like student test scores (e.g., Billings et al., 2022).

In this paper, we address the paucity of school board data by systematically collecting and analyzing a novel dataset comprising over 40,000 meeting minutes from 500 school districts across the United States. We use these data to examine how U.S. school boards engage with the topic of race, assessing the changes in the frequency of race-related discussions, which speakers are involved, the stance these speakers take, and how these patterns change in response to nationally significant race-related events, such as the 2020 murder of George Floyd and debates over CRT. This analysis is possible because a significant proportion of school boards' work is completed during public meetings, where members deliberate on policies, address community concerns, and make decisions. Meeting minutes represent the official record of these meetings—capturing the substance of discussions, the topics covered, noteworthy voting results, and actions taken.

Methodologically, we applied a combination of large language models (LLMs), natural language processing, and human coding to school board meeting minutes recorded between 2018 and 2022. We broadly operationalized "race-related conversations" as any statement where race and ethnicity were referenced, including associated issues such as discrimination, equity, inclusion, multiculturalism, and/or representation. We identified the stance of each statement as either *affirmative,* in favor of equity/diversity initiatives, or *oppositional,* a rejection/critique of these initiatives. We also identified the source of each statement—attributing them to either an *official*, such as school board members or invited speakers, or to the *public*, arising within the public comment section of a meeting. After aggregating from the statement to the meeting level, we subsequently used these measures to assess the frequency and framing of race-related discussions across school boards nationwide, examining variation by school district demographics, geographic region, and local political context.

This analysis revealed significant heterogeneity. First, while most school board meetings featured no race-related discussions (87% of meetings during the five-year period), a small proportion of districts generated the majority of such content. For instance, 26% of districts never discussed race-related content during the study period, while one-third of districts produced nearly 90% of all such content. Second, when race was discussed, affirmative statements dominated. We found that 12% of school board meetings contained affirmative race-related statements, compared to just 1% that contained oppositional race-related statements. Third, affirmative statements were predominantly spoken by officials, while oppositional critiques were more equally distributed between these officials and members of the public. Fourth, the frequency of race-related conversations correlated with district characteristics; discussions of race were most common in suburban, diverse, and politically competitive districts.

Finally, we also found that school boards are significantly responsive to broader societal and political debates. Using our five-year panel dataset, we assessed changes in school board discourse in the aftermath of George Floyd's murder in 2020, and the national conversations surrounding CRT in 2021. These data revealed that affirmative race-related statements surged in 2020, but this predominantly occurred in districts which leaned Democratic and/or were racially and ethnically diverse, and where discussions regarding race were already relatively more frequent. In contrast, oppositional statements observed in 2021 appeared in both left-leaning and right-leaning districts and were primarily concentrated in politically competitive districts.

Ultimately, this study makes two key contributions to the literature. First, we bring sizeable empirical data to bear on an issue of significant national interest: the extent to which race-related discourse in school board meetings is changing. To date, this topic has been predominantly examined anecdotally, via insightful but limited accounts. In so doing, we provide a valuable descriptive foundation for understanding how school boards address issues of race. Second, we highlight the methodological value of meeting minutes as an appropriate data source and demonstrate the contributions that rigorous LLM applications can make to education governance research.

Background

The Democratic Nature of School Board Governance

Although rules vary by state, school boards typically consist of seven unpaid democratically-elected officials (Tracy & Durfy, 2007). While critics have taken issue with school boards' lack of racial diversity (Diem et al., 2015; Hess & Meeks, 2010), and have highlighted the low participation rates in board elections (Kogan, 2022; Kogan et al., 2021), school boards' capacity to democratically represent local needs is a key argument in their favor (Hess & Meeks, 2010). Indeed, not only are board members democratically elected but their meetings "provide opportunities for local stakeholders to have direct input over policy and/or observe transparent policy deliberation" (J. Collins, 2021, p. 343) (Collins, 2021). With few exceptions, sunshine laws require school board business to be conducted in open meetings (J. E. Collins, 2021; Tracy & Durfy, 2007). These meetings also provide dedicated time for citizen input, although the extent to which members of the public use this opportunity varies (Campbell, 2005).

School Board Activities

Researchers have catalogued the common duties carried out by school boards, including the recruitment and termination of superintendents and staff, curricula approval, financial and contractual oversight, staff negotiation, and the management of contractual services (Ehrensal & First, 2008). Beyond such discrete duties, however, there is limited agreement on school boards' broader role in shaping the education system. Indeed, education reforms have typically circumvented school boards, preferring schools themselves to be the unit of change—for example, via charter school formation, school choice initiatives, and school accountability reforms (Anderson, 2003; Cohodes & Parham, 2021; Jacob, 2017). However, the National School Board Association (2023) identifies policy-making and improving student achievement as school boards' key roles. Others believe school boards' key role is to act as an interface between schools and the public, promoting connection and listening to public concerns (Ehrensal & First, 2008; Hochschild, 2005). In practice, research on how school boards use their power is limited. Nevertheless, surveys and case studies suggest that boards spend only three percent of their time on policy development and oversight (Olson & Bradley, 1992). Indeed, a long-standing critique of school boards is that they are "mired in minutia" (Nowakowski & First, 1989).

Given the power granted to them, however, school boards can still undertake actions to address pressing challenges, including racial inequities. In 2021, for example, in the wake of national protests for racial equity, the National School Board Association (2021) recommended that boards should increase their focus on several areas, including: within-district funding equity across schools, recruiting and retaining teachers of color, selecting culturally responsive curricula, providing professional development opportunities to address teachers' implicit biases, and incorporating restorative justice practices. To date, however, there remains no empirical research concerning which U.S. school boards discuss race in the course of their duties, nor whether the character and frequency of such discussions have changed in recent years.

Recent School Board Controversies

Several scholars have examined recent school board controversies, concerning issues like COVID-19, gender, and CRT (Graves, 2024; Holman et al., 2024; Kitchens & Goldberg, 2024; Shah et al., 2024). Quantitative work has assessed how the rise in controversial issues affected voters' decisions at the ballot box (Jacob, 2024; Shah et al., 2024) or has correlated school board member partisanship and racial demographics with district decision-making (Kitchens & Goldberg, 2024). To the best of our knowledge, Holman et al. (2024) is the only existing study that provides extensive evidence regarding the prevalence of school board meeting conflicts. Using transcripts of videoed school board meetings, these researchers discovered that most school boards experience some kind of conflict, that this conflict coincides with national events, and that the prevalence of such conflict is correlated with district demographics. By contrast, our approach in this paper employs a more nationally representative sample than Holman et al. (2024) and narrows the focus from broadly defined conflict to the specific issue of race. Concurrently, we demonstrate the benefit of using LLMs to harness the insights contained in thousands of previously unexamined meeting minutes.

Study Context

We focus our analysis on a period of intense social change in the United States, from 2018 to 2022. This period includes the widespread protests following the killing of George Floyd and national conversations surrounding CRT. We include 2018 as a baseline year, providing a valuable point of comparison for the particularly tumultuous years that follow. Approximately two and a half years following our starting time point, on May 25, 2020, George Floyd, a 46-year-old Black man was killed by Minneapolis city police. Protests began in Minneapolis the next day (Reuters, 2020), and quickly intensified after the police officers were formally charged. On July 20, 2020, thousands of U.S. workers staged an 8-minute walk-off in a "Strike for Black Lives.". Further protests, including an occupation of the intersection where George Floyd was killed, continued in the following years (Cobb et al., 2024).

Just months after George Floyd's death, in September 2020, Chris Rufo, a conservative activist, argued on national television that publicly funded institutions were inappropriately teaching people to apply CRT to their lives. Soon after, the Trump administration banned forms of diversity training which included "divisive concepts," such as "the United States is fundamentally racist or sexist," or that "an individual, by virtue of his or her race or sex, is inherently racist, sexist, or oppressive, whether consciously or unconsciously" (Exec. Order No. 13950, 2020). By 2021, several states had banned the teaching of CRT in public schools (Schwartz, 2021).

As local institutions responsible for setting district policies, these overlapping events forced school boards to navigate a shifting terrain of public opinion, state mandates, and community pressures. Thus, by focusing on this period, our study captures how national events filtered down to local decision-making, providing a unique window into the responsiveness of school boards to broader societal forces.

Summary of Approach

To systematically examine race-related discussions in school board meetings, we first identified a stratified random sample of 543 U.S. public school districts. We then collected over 40,000 meeting minutes from these districts and used optical character recognition (OCR) to extract text from the documents. Next, we transformed this natural language text data into a structured dataset, and applied LLMs to identify race-related discussions, the stance of the statements (affirmative, oppositional, or both), and the type of speaker (official or public comment). Finally, these measures were used to assess the relationship between school board discourse, district characteristics, and national events.

Sample and Data

We stratified public school districts by two criteria: district size (measured by the number of students) and 2016 Republican presidential vote share. We separated our sample by quintiles for each criterion, generating 25 strata. This ensured that our sample included large Democratic-leaning districts as well as small Republican-leaning districts (and everything in between). Within each stratum, we randomly selected 5% of observations, resulting in the final sample of 543 public school districts.

For each school district, employees of a third-party data collection service visited the district's website and downloaded all available school board meeting minutes from 2018 to 2022—a process that return 40, 316 documents and succeeded in accessing the relevant documentation for 92% of districts. As shown in Table 1, the average district enrollment in our sample was larger than the national average (reflecting our stratified sampling) but was relatively representative in terms of urbanicity and student demographics. The analytic sample consisted of 500 districts (the 92% of sampled districts where we were able to identify publicly posted meeting minutes): 23% in the Northeast, 37% in the Midwest, 24% in the South, and 17% in the West. Demographically, the average district in our sample has 25% Black or Hispanic students and 48% of students who receive Free or Reduced-Price Lunch (FRPL). Approximately 40% of the sample districts are rural, 22% in towns, 29% suburban, and 8% urban.

We collected all available meeting minutes, including regular board meetings and special board meetings, for every school board in our sample between 2018 and 2022. This generated over 40,000 documents, approximately 18 documents per district per year, though there are many districts where we were not able to identify meeting minutes for earlier years in our panel (see Table 2). Typically, these documents were in PDF format, and sometimes they had been photocopied, making it difficult to extract the text. To address this, we applied OCR implemented within LayoutParser (Shen et al., 2021) to extract and store each document's unique text and formatting information. We linked the data extracted from meeting minutes to three additional data sources. The Common Core of Data contained information on geography, enrollment, and student demographics for each district in our sample. The Stanford Education Data Archive (Reardon et al., 2016) contained information on district-level student academic outcomes, disaggregated by racial and economic sub-groups. Harvard's Voting and Election Science Team provided information on the partisanship of the surrounding geographic area.

Measures

Definition of Concepts

For each meeting, our aim was to determine whether it contained any of the following occurrences: (1) a race-related statement; (2) an affirmative race-related statement; (3) an oppositional race-related statement; (4) a race-related statement from an official; (5) a race-related statement from the public comment section of the meeting; (6) an affirmative and/or oppositional race-related statement from an official; and (7) an affirmative and/or oppositional race-related statement from the public comment section. (Each of these occurrences is defined below.) To differentiate by speaker and stance, we first segmented all the documents into 150-word segments of text—a size we judged to be large enough to determine speaker and stance while small enough to use LLMs efficiently. Later, we aggregated the segment-level classifications to the meeting level by creating binary variables for whether there are one or more of the previously outlined occurrences (e.g., any race-related statement, any affirmative statement from the public).

We broadly define "race-related conversations" as any statement where race and ethnicity—or associated issues such as discrimination, equity, inclusion, multiculturalism, and/or representation—are referenced. Statements about other aspects of identity (such as gender, sex, religion, or disability) are not incorporated into this definition, nor are statements about equity pertaining to non-demographic factors, such as equity in funding between athletics and fine arts, unless race is also specifically mentioned. However, general statements about equity would be coded as race-related unless the speaker is clearly referring to equity on a dimension other than race or ethnicity.

We further characterized each race-related statement according to whether it was uttered during the public comment section of a meeting or by an official (e.g., school board member, superintendent, or invited guest). We also established whether the statement included an affirmative stance concerning race in education or an oppositional critique of such a stance, or both. We use these terms in a very particular sense, where an affirmative stance pertains to any race-related statement other than those which critique equity and multicultural initiatives. Under this definition, the character of affirmative statements may range from symbolic (such as statements that simply express appreciation for diversity) to action-oriented (such as statements that indicate substantive efforts to increase the hiring of teachers of color). They may range, too, from politically neutral (such as statements acknowledging Black History Month) to politically liberal (such as statements expressing support for the Black Lives Matter movement). Oppositional statements are defined as those which critique and/or disavow racial equity and/or multicultural initiatives. Under this definition, critiques of CRT would be coded as oppositional, as would a statement from a board member stating that CRT is not taught in the district. However, if the board member continued to discuss how the district addresses multiculturalism, the statement would be coded as both affirmative and oppositional. (See Table 3 for examples of the above categories.)

The Use of Large Language Models (LLMs) in the Categorization of Documents

If we were conducting this study twenty years ago, we would have had little choice but to hand-code our corpus according to our concepts of interest. Because of the time-intensive nature of such a task, our sample would almost certainly have been limited in size and scope. Even five years ago, we would have likely aimed to train a series of supervised learning classifiers from scratch—labeling a subset of our corpus to serve as training examples and using these annotations to teach the model to identify the words and phrases associated with each of our concepts in order to predict the labels of new documents. Today, however, we have a new option; rather than training a model from scratch, we can rely on pre-trained generative LLMs to characterize text data.

Generative LLMs are natural language models that have been pre-trained on enormous samples of text, capable of producing ‘reasonable continuations’ of writing following users’ natural language prompts (Reynolds & McDonell, 2021). They achieve this by maximizing the probability of each subsequent term, conditional on the previous terms (Radford et al., 2018). This generative capacity can classify texts by indicating a classification task in a user’s prompt (Radford et al., 2018, 2019). For example, consider a simple prompt, such as: “Classify the following text as race-related (yes) or not (no). Text: [EXCERPT].” When the prompt is followed by an excerpt unrelated to race, the next most likely word—conditional on the terms provided in both the prompt and excerpt—is “no.”

While training a classifier model from scratch typically requires thousands of labeled examples to reasonably align with human codes, a generative LLM can be applied to this task without any labeled examples at all (a process termed “zero-shot learning”). Alternatively, a researcher may provide the LLM with 2 to 5 select examples (“few-shot learning”). When a researcher wishes to provide a larger quantity of examples, they commonly engage in a process termed “fine-tuning,” where the LLM undergoes an additional round of training with a smaller, task-specific training dataset. Even in fine-tuning, however, the number of examples provided is typically fewer than one hundred. Researchers may thus apply LLMs to text classification tasks while hand-labeling anything from a few documents to several dozen. However, using an LLM to automatically code a corpus without a larger hand-labeled dataset presents two important limitations. First, there is no guarantee that the chosen prompt wording and/or selected examples are the best approach for maximizing the quality of the LLM output. Second, without a trusted labeled sample, there is no way to know whether the LLM output matches the researchers' understanding of the concepts. For these reasons, we follow a more rigorous process in this paper. First, we hand-code a subset of our corpus. Second, we split this hand-labeled dataset into distinct training, development, and validation samples. Third, we use the training sample to provide examples via a few-shot approach—fine-tuning the LLM and updating the model parameters in a brief training round. Fourth, we use the development sample to select among prompts and between zero-shot, few-shot, and fine-tuned models. Fifth, we validate our final LLM-based classifiers using the hold-out validation sample.

Human Coding, Model Selection, and Validation

We began the modeling process with a simple random sample of 1,500, 150-word segments from our meeting minutes corpus. Because race-related segments in meeting minutes are rare, we supplemented this representative sample with a further sample of 1,500 segments, each of which contained at least one of 70 distinct race-related terms (see Appendix A for a description of how those terms were identified). A team of undergraduate research assistants then viewed these 3,000 segments and classified each segment as either race-related or not race-related—basing their choice solely on the segment content and using the key definitions discussed above. If a segment proved to be race-related, the research assistants further coded the segment's stance (affirmative/oppositional/both) and speaker (public comment/official). Inter-rater reliability for both race-related and affirmative/oppositional was estimated to be 0.75 (Cohen's Kappa), while inter-rater reliability for the speaker was 0.74. According to McHugh (2012), both of these figures indicate moderate reliability.

We randomly separated our hand-coded segments into training, development, and validation datasets, at a ratio of 20:20:60. We used the training and development datasets to engineer our prompt, to select examples for few-shot classification, and to choose between prompts and approaches (i.e., zero-shot, few shot, or fine-tuning). We used the validation data to assess the performance of the final model.

For model selection and validation, we assessed the relationship between *true positives*, *true negatives*, *false positives,* and *false negatives* for each of our concepts of interest. In these terms, positive and negative refer to whether the model identified a concept's presence in a segment (e.g., race-related or not, oppositional or not, public or official), while true and false indicate whether the human coder agreed. For example, a false positive for a race-related concept indicates that the model identified the segment as race-related while the human coder did not. Using these definitions, we formulate the following equations for accuracy, recall, and precision:

$$Accuracy = \frac{True\ Positives\ +\ True\ Negatives}{False\ Positives\ +False\ Negatives + True\ Positives\ +\ True\ Negatives};$$

Equation 1

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives};$$

Equation 2

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}.$$

Equation 3

Here, accuracy is simply the proportion of segments that the LLM classified correctly for a given concept (e.g., the proportion of segments correctly classified as race-related or not race-related). Recall is the proportion of correctly identified cases from among all true positive cases (e.g., the proportion of race-related segments, according to the human coder, that the model also identifies as race-related). Lastly, precision is the proportion of correctly identified positive cases among all LLM classified positive cases. In other words, of the segments the LLM identified as race-related, for example, what proportion did the human also classify as race-related? In our development dataset, our best performing model for race-related and stance classification was the fine-tuned model, and our best performing model for speaker classification was a zero-shot classification approach. These were the models we tested in validation data and applied to the remaining unlabeled excerpts.

Performance results for the validation sample (900 segments) are shown in Table 4. The model identified 91% of statements that a research assistant had coded as race-related (recall). Of the statements the model identified as race-related, a research assistant agreed 96% of the time (precision). In classifying statements as affirmative or not (where not affirmative includes both oppositional statements and statements that are not race-related) recall was 90% and precision was 94%. The model identified oppositional statements with somewhat lower performance, with recall at 78% and precision at 74%. To address this latter limitation, we supplemented our LLM classifications with a straightforward count of "critical race theory" mentions across all segments. Finally, among all race-related segments, the model could distinguish between public comments and official comments with an accuracy of 97%.

Aggregation Methods

We applied the LLMs to our full corpus of meeting minute segments, such that each segment was associated with a series of binary measures: race-related, affirmative, oppositional, public comment, and official comment. (The last four are only indicated if the segment was previously classified as race-related). We also generated indicators of affirmative-public comment, affirmative-not public comment, oppositional-public comment, and oppositional-not public comment. We aggregated these indicators to the meeting level, creating binary indicators to show whether there was any affirmative or oppositional race-related conversation from each speaker type within a given meeting.

Limitations of Meeting Minutes

Thanks to open meetings laws and our stratified approach, our collection of meeting minutes provided a large, generalizable sample. However, while school boards are required to publicly post meeting minutes, they are not required to take extensive notes on their discussions (although many do). We thus cannot routinely expect to find in-depth, detailed accounts of speakers' meeting contributions—raising questions about minutes' reliability as a source of information on meeting activities. To assess the scale of this potential drawback, we supplemented our primary sample with a non-representative sample of 118 U.S. school districts that upload video recordings of their school board meetings to YouTube (Appendix B). We then compared verbatim transcripts of these recorded meetings—which document almost every word spoken—to our primary meeting minutes data. While race-related discussions were more prevalent in the video transcript sample (25% in 2021) than the minutes sample (14% in 2021)—likely stemming from the difference in medium and respective sample characteristics—the two data sources nevertheless exhibited similar trends regarding the prevalence of race-related content over time and the relationship between prevalence and local politics. (Again, see Appendix B for further details on these analyses.)

Analysis

We began our analysis with a series of simple descriptive statistics, calculating averages and percentiles for our metrics (race-related, stance, and speaker) for all districts across the five-year study period. Then, to assess the relationship between race-related content and district characteristics, we calculated the proportion of race-related meetings in each year t for each district d (Y_{dt}) and estimated a series of regressions:

$$Y_{dt} = \beta_0 + \beta_1 X_d + \beta_2 BaselineWords_d + \lambda_t + + \varepsilon_{dt},$$

Equation 4

Here, X_d represents one of several district characteristics, including urbanicity, student demographics, geography (each measured in 2018), and political affiliation (measured in 2016). The variable $BaselineWords_d$ served as a proxy for the baseline level of detail in meeting minutes and is defined as the median number of words in a district's first three meetings. The term λ_t represents year fixed effects and controlled for the variability in minutes' availability across the study period exhibited in Table 2.

Next, we estimated a conditional model that included controls for the full set of analyzed district characteristics. This allowed us to assess which characteristics explained the greatest amount of variation in racial discourse, after adjusting for other characteristics (such as attempting to separate the variation in urbanicity from the variation in vote share). Despite the inclusion of these controls, we were not aiming to identify a causal relationship here, given likely unobserved confounders between characteristics and race-related discourse. Rather, we wished to identify the most robust and consistent relationships. In both the individual and conditional models, standard errors were clustered so as to reflect the hierarchical nature of the analysis (i.e., the inclusion of multiple minutes for each district).

To assess the relationship between the prevalence (and character) of race-related meeting conversations and national events, we first plotted a series of graphs for all districts demonstrating changes in meeting content over time, disaggregated by district politics. We also estimated a series of equations:

$$Y_d = \beta_0 + \beta_1 X_d + \delta_1 BaselineWords_d + \varepsilon_e,$$

Equation 5

where Y_d is one of three measures of interest: (1) $Y_{2018,d}$ the proportion of 2018 meetings within a district featuring affirmative race-related content from an official speaker (school board member, administrator, or invited guest); (2) $Y_{2021-2018,d}$ the change in this proportion between 2018 and 2021; and (3) $Y_{2021,d}$, the proportion of meetings within a district with oppositional race-related public comments in 2021. Again, $BaselineWords_d$ is defined as above and serves as a proxy for reporting quality,X_d is a district characteristic of interest—including student and community demographics, geography, student achievement and community political leanings—and β_1 captures the relationship between the given district characteristic and our measures of interest. Once again, after estimating a series of individual regressions, we estimated a conditional model following the process described above.

Results

The Prevalence and Nature of Race-Related Discussions

Over the five-year study period, we found that race-related discussions were relatively rare in U.S. school board meetings, were driven by a minority of districts, and were most commonly affirmative in nature. In the median district, only 4% of meetings mentioned race, while in a quarter of districts, race was never mentioned (see Table 5). Only a small subset of districts frequently engaged with race-related topics. In the 95th percentile district, for example, nearly half of all meetings addressed race. Overall, when racial discussions occurred, they were overwhelmingly affirmative; 12% of meetings contained affirmative race-related content, whereas only 1% featured oppositional content (see Table 6). Notably, affirmative content was most commonly introduced by officials, with public contributions accounting for only 17% of such discussions. Oppositional content, however, was nearly equally likely to be voiced by officials and members of the public.

To contextualize the frequency of race-related discussions, we also examined the prevalence of various non-race-related topics. To do so, we selected approximately 500 meeting minutes from 50 randomly sampled districts (ten randomly sampled meetings from each district). These topics included academics and curriculum, student attendance and enrollment, pandemic-related policies, facilities and technology, finance and budget, and hiring and personnel. To assess the frequency of these topics, we used LLMs in a manner comparable to the approach described above, though now employing zero-shot classification with minimal prompt engineering. This analysis thus provided a useful, albeit less rigorous, point of comparison for our race-related study. As might be expected, financial and budgetary matters dominated school board discussions, appearing in 84% of meeting minutes, followed closely by personnel and human resources at 79%. Academics and curriculum were also a common topic, discussed in 38% of the minutes. Notably, every non-racial topic analyzed appeared more frequently than race-related discussions, further underscoring the relative rarity of race as a discussion point in school board meetings (Table 7).

District Characteristics and Race-Related Discussions

The frequency of race-related discussions in school board meetings varied significantly by region, student demographics, and community characteristics. Columns 1, 3, and 5 in Table 8 present the regression coefficients for the proportion of meetings with race-related content, the proportion of meetings with affirmative race-related content, and the proportion of meetings with oppositional race-related content for district characteristics of interest—controlling for year fixed effects and baseline meeting minutes length (Equation 4). We found that race-related discussions are most common in the West and Northeast regions and occurred far less frequently in the South (which serves as the omitted category in the region model, such that the coefficients of the remaining regions should be interpreted as differences from the Southern comparison). Likewise, urban and suburban districts engaged with race-related topics at higher rates than rural districts (the omitted category in the urbanicity model). For instance, districts in the West discussed race 13 percentage points more frequently than those in the South (where race was discussed in 6% of meetings), and urban districts discussed race 14 percentage points more frequently than rural districts (where race was discussed in 5% of meetings). These same patterns tended to hold when distinguishing by stance; both affirmative and oppositional content were most common in Western and suburban districts, and least common in rural, Southern, and Midwestern districts.

Student demographic characteristics were similarly associated with the prevalence of race-related discussions. Larger school districts, and districts with larger proportions of Black and Hispanic residents, exhibited significantly greater race-related discourse, while smaller districts, and districts with higher proportions of white residents, exhibited less. However, these demographic associations tended to be driven by affirmative content rather than oppositional content.

We also found that the prevalence of race-related discussion varied according to the characteristics of adults within the district area. Communities with higher levels of educational attainment and income exhibited significantly more race-related discourse, encompassing both affirmative and oppositional content. Finally, we found that the 2016 Republican presidential vote share was negatively associated with race-related content. While this was true for both affirmative and oppositional content, the pattern was strongest for affirmative content.

In Columns 2, 4, and 6 of Table 8, we present the modeling results that control for all characteristics. Using this approach, the characteristics that best predicted race-related conversations were district geography, school enrollment, and local political affiliation.

Trends and National Events

Examining changes over time, we found that the prevalence and nature of race-based discussions both tended to align with relevant national events. Figure 1 plots the average percentage of district meetings containing race-related content across the study period. The graph shows the prevalence of race-related content beginning to climb in the summer of 2020, concurrent with George Floyd's murder and subsequent protests. The vast majority of this content was affirmative and reached a peak in the spring of 2021. After this date, affirmative content declined, while oppositional content increased, peaking in the summer of 2021.

The Relationship between Race-Related Discussions, National Events, and District Characteristics

The relationship between national events and race-related discussion in school board meetings is not homogenous. Following George Floyd's murder and associated protests, we found that the greatest increase in affirmative race-related content occurred in left-leaning districts. Following the national CRT debates, the greatest increase in oppositional content occurred in politically competitive districts. Figure 2 disaggregates these relative frequencies by local political affiliation. Panel A shows the proportion of meetings featuring affirmative statements within blue districts (defined as having a <.45 Republican 2016 presidential vote share), purple districts (politically competitive district with a Republican 2016 presidential vote share between 0.45 and 0.55), and red districts (with a >.45 Republican 2016 presidential vote share). The graph demonstrates that blue districts included affirmative statements most often, that they were the first to experience increases in discussion frequency following George Floyd's murder and subsequent protests, and that they experienced the greatest increase in frequency. Panel B plots a similar graph for oppositional race-related statements, showing that purple districts experienced the greatest increase in statements following CRT's entry into the national conversation. Formal tests of these and other patterns revealed further details about the relationship between district characteristics and national trends (Table 9, Columns 1, 3, and 5). Using a district level dataset, and a series of regressions described in Equation 5, we examined the following three variables: (1) the proportion of 2018 meetings within a district featuring affirmative race-related content from official speakers; (2) the change in the previous variable between 2018 and 2021; (whereby officials might be responding to the moment of heightened racial awareness); and (3) the proportion of 2021 meetings within a district with oppositional public comments (whereby the public might be responding to national debates surrounding CRT). For 2018, after controlling for baseline meeting length, we found that board officials in highly educated, suburban districts in the Northeast and West tended to discuss affirmative race-related content most frequently. These same characteristics also tended to predict an increase in affirmative statements from officials following racial equity protests in 2020, as well as oppositional comments from the public in 2021. However, as Figure 2 indicates, we identified an additional characteristic with a significant relationship to oppositional public comments: whether the district was located in an area where the 2016 presidential vote share was split relatively evenly between Republicans and Democrats. On average, these politically competitive districts experienced almost 2% more meetings featuring oppositional public comments in 2021—a 110% increase in frequency. This finding remains consistent in Column 6, which shows the coefficients for the conditional model.

Discussion

Theoretically, the decentralized nature of the United States K-12 education system is designed to encourage "the greatest participation by those most directly concerned" (*San Antonio*

Independent School District v. Rodriguez, 1973). In this sense, elected school board members and public citizens who participate in open school board meetings are expected to discuss substantive issues so that school district decision-making reflects local values and preferences (White, 2023). On the other hand, such localized control permits substantial variation regarding whether and how school districts consider key social issues, including racial equity and diversity initiatives.

One long-standing critique of school boards in the academic literature is that boards spend too much time on so-called “administrivia” (Nowakowski & First, 1989). However, recent reports of strife and anger at school board meetings suggest that school boards are spending much of their time and energy on controversial social issues. In this paper, we do not address the normative question of whether school boards should engage more deeply in policy-oriented and/or value-based discussions. Rather, we pose a question that is currently under-addressed in the field: to what degree do school boards, and members of the public, engage in such conversations, and under what circumstances?

We find that school board discussions relating to race are relatively uncommon compared to other topics. For example, while 84% of meeting minutes during our study period discussed finance and budgets in this year, 13% discussed race in some manner. Moreover, racial discussions are not equally distributed among school districts; 26% of districts generated zero identifiable race-related content during the study period, while the 5% of districts that discussed race most frequently did so in nearly half of their meetings. Our findings also illuminate the extent to which changes in school board practice and public participation coincide with national events. Despite national reporting on CRT debates, just 3% of meetings in 2021 contained any oppositional critiques, and only 2% contained the term CRT or “critical race theory.” These figures are remarkably similar in the smaller dataset containing verbatim transcripts of LocalView meeting discussions—3% and 2%, respectively. These figures are perhaps much lower than the journalistic reporting at the time might have suggested. However, we can identify a few school districts where CRT has become a prominent issue. In 2021, for example, three districts featured oppositional race-related content in over 40% of their meetings.

Variations in overall race-related content, and especially in the frequency of oppositional critiques, raise the important question of which U.S. school board districts are most likely to discuss these issues. We found that race-related discussions (including affirmative statements and oppositional critiques) most often occur in suburban districts and in districts with a more highly educated population. Previous research has demonstrated that highly educated individuals are more likely to participate in board meetings (Campbell, 2005), and our findings suggest that more highly educated communities are perhaps also more likely to elect school board members who engage in policy and value-based issues like racial equity and diversity. Previous research also suggests that demographic and political heterogeneity encourages competitive and conflictual engagement (in this case, for example, debates over CRT). As Campbell (2005, p. 4) notes, for example, "a social context that triggers conflict over differing political preferences will result in higher levels of political activism, while one characterized by people sharing common characteristics—and thus, by implication, preferences—will instead foster collective action motivated by feelings of social solidarity." Our findings largely support this hypothesis, as oppositional critiques appear most likely in districts with mixed political affiliations. Finally, we also observe that race-based conversations, both affirmative and oppositional, happen most often on the East and West coasts. Midwestern districts and Southern districts were less likely to have discussed race in 2018, less likely to feature affirmative race-related content in 2020, and less likely to feature oppositional content in 2021. This suggests that racial equity initiatives and oppositional critiques at school board meetings are not uniformly distributed across the country but rather vary significantly by region and local context.

Limitations

Our findings underscore the value of meeting minutes as a data source for examining the dynamics of school board governance, particularly when coupled with scalable methods of analysis like LLMs. However, this choice of data source and analysis both present limitations. First, meetings minutes are an incomplete record of events, ultimately subject to the discretion of the minute-taker. We addressed this shortcoming by supplementing the minutes dataset with the smaller sample of meeting transcripts—finding that our results remained largely robust. Second, as the performance results in Table 4 indicate, our LLM-based meeting content measures are imperfect; we do not capture every instance of race-related content, nor does every LLM-identified instance of race-related content meet our definition of "race-related." (This same imperfection also applies to categorizations of affirmative, oppositional, public comment, or official comment). While the use of LLMs thus certainly introduces some noise into these measures, we believe this shortcoming is compensated for by LLMs' ability to analyze a much larger number of documents than would be feasible by hand. Finally, our study only explains the role school boards play in shaping school districts' handling of racial issues, focusing specifically on the frequency with which these topics arise. We hope that future research will address both the variable content of these discussions—the specific issues discussed, and any actions taken—and, ultimately, their impact on student experiences.

Conclusion

Given the decentralized education system in the United States and heterogeneous local cultural contexts, there is the potential for great variability in the nature and extent of school boards' engagement with key social issues like racial equity and diversity. Recent media reports suggest that racial issues have become increasingly prominent in school board discussions. This paper has revealed substantial variability concerning the extent to which districts address these issues and illuminates the influence of local politics and demographics on the conversation. In so doing, it brings large-scale, empirical evidence to bear on important questions which have thus far only been explored in anecdotal media reports and small-scale, qualitative studies.

Bibliography

Anderson, S. E. (2003). The school district role in educational change: A review of the literature. *International Centre for Educational Change*, *34*(2), 25–45.

Barari, S., & Simko, T. (2023). LocalView, a database of public meetings for the study of local politics and policy-making in the United States. *Scientific Data*, *10*(1), 135.

Billings, S. B., Macartney, H., Park, G., & Singleton, J. D. (2022). *Self-interest in public service: Evidence from school board elections*. National Bureau of Economic Research.

Campbell, D. E. (2005). Contextual Influences on Participation in School Governance. In W. G. Howell (Ed.), *Besieged* (pp. 288–307). Brookings Institution Press; JSTOR. http://www.jstor.org/stable/10.7864/j.ctt127xjr.15

Cobb, J., Alvarez, M., Reese, E., & Bright, E. (2024, May 25). *'No Justice, No Streets': 4 years after murder, George Floyd Square stands in protest*. MPR News.

Cohodes, S. R., & Parham, K. S. (2021). *Charter schools' effectiveness, mechanisms, and competitive influence*.

Collins, J. (2021). Should School Boards Be in Charge? The Effects of Exposure to Participatory and Deliberative School Board Meetings. *Peabody Journal of Education*, *96*(3), 341–355. https://doi.org/10.1080/0161956x.2021.1943239

Collins, J. E. (2021). Does the Meeting Style Matter? The Effects of Exposure to Participatory and Deliberative School Board Meetings. *American Political Science Review*, *115*(3), 790–804. https://doi.org/10.1017/s0003055421000320

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv Preprint arXiv:1810.04805*. https://doi.org/10.48550/arXiv.1810.04805

Diem, S., Frankenberg, E., & Cleary, C. (2015). Factors That Influence School Board Policy Making. *Educational Administration Quarterly*, *51*(5), 712–752. https://doi.org/10.1177/0013161x15589367

Ehrensal, P. A., & First, P. F. (2008). Understanding school board politics: Balancing public voice and professional power. In *Handbook of education politics and policy* (pp. 87–102). Routledge.

Exec. Order No. 13950. (2020, September 22). *Executive Order on Combating Race and Sex Stereotyping*. https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-combating-race-sex-stereotyping/

Graves, M. V. (2024). *The New Culture War: Critical Race Theory, Gender Politics, K-12 School Board Meetings, Founding Myths, and the Religious Right*. Bowling Green State University.

Hess, F. M., & Meeks, O. (2010). School Boards Circa 2010: Governance in the Accountability Era. *Thomas B. Fordham Institute*.

Hochschild, J. L. (2005). What school boards can and cannot (or will not) accomplish. *Besieged: School Boards and the Future of Education Politics*, 324–338.

Holman, M., Johnson, R., & Simko, T. (2024). *Measuring Conflict in Local Politics*. OSF PrePrint. URl: https://osf. io/vst9g.

Jacob, B. (2017). The Changing Federal Role in School Accountability. *Journal of Policy Analysis and Management*, *36*(2), 469–477. https://doi.org/10.1002/pam.21975

Kitchens, K., & Goldberg, M. (2024). Partisanship and Professionalization: School Board Decision-Making in the Midst of a Pandemic. *Urban Affairs Review*, *60*(5), 1439–1475.

Klein, A. (2024, November 14). *Can Trump Force Schools to Change Their Curricula?* EdWeek. https://www.edweek.org/policy-politics/can-trump-force-schools-to-change-their-curricula/2024/11

Kogan, V. (2022). Locally elected school boards are failing. *Education Next*, *22*(3).

Kogan, V., Lavertu, S., & Peskowitz, Z. (2021). How does minority political representation affect school district administration and student outcomes? *American Journal of Political Science*, *65*(3), 699–716.

Kokal, K. (2023, February 2). *Leave signs and insults at home: Palm Beach County School Board could clamp down on speakers*. The Palm Beach Post. https://www.palmbeachpost.com/story/news/education/2023/02/02/palm-beach-county-school-board-meeting-speakers-mics-may-be-cut-signs-banned/69858185007/

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://hrcak.srce.hr/89395

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. https://doi.org/10.48550/arXiv.1301.3781

National School Board Association. (2018). *Today's School Boards and Their Priorities for Tomorrow*. https://cdn-files.nsba.org/s3fs-public/reports/K12_National_Survey.pdf

National School Board Association. (2021). *Reimagining school board leadership: Actions for equity*. DIRE Initiative and Center for Safe Schools. https://www. nsba. org/-/media ….

National School Board Association. (2023). *About School Board and Local Governance*. https://www.nsba.org:443/About/About-School-Board-and-Local-Governance

Nowakowski, J., & First, P. F. (1989). A Study of School Board Minutes: Records of Reform. *Educational Evaluation and Policy Analysis*, *11*(4), 389–404. https://doi.org/10.3102/01623737011004389

NSBA. (2021, October). *National School Boards Association letter to Biden*. https://www.documentcloud.org/documents/21094557-national-school-boards-association-letter-to-biden

Olson, L., & Bradley, A. (1992). Boards of contention: Introduction. *Education Week*, *11*(32), 2–3.

OpenAI. (2024). *Embeddings*. https://platform.openai.com

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. OpenAI.

Reardon, S. F., Kalogrides, D., & Shores, K. (2016). Stanford Education Data Archive (SEDA) Summary of data cleaning, estimation, and scaling procedures. *Stanford Center for Education Policy Analysis Https://Cepa Stanford Edu/Sites/Default/Files/SEDA% 20data% 20construction% 20documentation*.

Reuters. (2020, June 25). *Timeline: Key events in the month since George Floyd's death*. https://www.reuters.com/article/world/us/timeline-key-events-in-the-month-since-george-floyds-death-idUSKBN23W1NQ/

San Antonio Independent School District v. Rodriguez (U.S. Supreme Court 1973).

Schwartz, S. (2021, June). *Map: Where Critical Race Theory Is Under Attack*. https://www.edweek.org/policy-politics/map-where-critical-race-theory-is-under-attack/2021/06

Shah, P., Weinschenk, A., & Yiannias, Z. (2024). Schoolhouse rocked: Pandemic politics and the nationalization of school board elections. *State Politics & Policy Quarterly*, *24*(2), 207–217.

Sutherland, D. H. (2022). School board sensemaking of federal and state accountability policies. *Educational Policy*, *36*(5), 981–1010.

Tracy, K., & Durfy, M. (2007). Speaking out in public: Citizen participation in contentious school board meetings. *Discourse & Communication*, *1*(2), 223–249.

White, R. S. (2023). Are Locally Elected School Boards Really Failing? *Education Next*, *23*(1). Coronavirus Research Database; Publicly Available Content Database. https://www.proquest.com/scholarly-journals/are-locally-elected-school-boards-really-failing/docview/2758544550/se-2?accountid=30699

Wilder, D., Bensen, J., & Swalec, A. (2021, June 23). *The Meeting Has Degenerated': 1 Arrest, 1 Injury at Loudoun Schools Meeting on Equity*. NBC Washington. https://www.nbcwashington.com/news/local/northern-virginia/loudoun-school-board-transgender-student-policy-race-equity/2708185/

Tables and Figures

Table 1

School District Sample Characteristics

	National Population	Minutes Sample
Prop. Districts in Northeast	0.19	0.23
	[0.4]	[0.42]
Prop. Districts in Midwest	0.37	0.37
	[0.48]	[0.48]
Prop. Districts in South	0.24	0.24
	[0.43]	[0.43]
Prop. Districts in West	0.2	0.17
	[0.4]	[0.37]
Prop. Districts Urban	0.06	0.08
	[0.24]	[0.26]
Prop. Districts Suburban	0.23	0.29
	[0.42]	[0.46]
Prop. Districts Town	0.18	0.22
	[0.39]	[0.41]
Prop. Districts Rural	0.52	0.41
	[0.5]	[0.49]
Enrollment (1,000) in District	3.55	3.95
	[13.94]	[9.49]
Prop Black or Hispanic/Latino Students in District	0.23	0.25
	[0.26]	[0.27]
Prop Other, including Asian, American Indian, Alaska Native, Native Hawaiian, and Two or More Races Students, in District	0.05	0.05
	[0.13]	[0.1]
Prop. White Students in District	0.71	0.71
	[0.28]	[0.28]
Prop. FRPL Students in District	0.51	0.48
	[0.23]	[0.23]
Aggregate district mean test score	0.01	0.05
	[0.35]	[0.36]
Prop adults in district area with BA+	0.25	0.26
	[0.13]	[0.14]
N	13071	500

Note: These data come from the Common Core of Data (CCD) and the Stanford Education Data Archive (SEDA). Standard deviations are in brackets.

Table 2

Number of Districts and Identified Meeting Minutes Per District by Year

Year	# Unique Districts	Min. # of Meetings	Ave. # of Meetings	Median # of Meetings	Max. # of Meetings
2018	348	1	19.36	18	72
2019	388	1	18.97	18	93
2020	428	1	19.79	18	81
2021	479	1	19.18	18	81
2022	482	1	17.51	16	79

Note. 334 districts have at least one meeting across all years in the timeframe.

Table 3

Classification of Example Excerpts According to Four Key Definitions: Not Race-Related, Race-Related Affirmative, Race-Related Oppositional, and Race-Related Both

Example Excerpt	Classification
Services for students with disabilities equity requirement…students must not be excluded from career technical or academic programs, courses, services, or activities due to equipment barriers or because necessary related aids and services or auxiliary aids are not available.	Not Race-Related
Technology Academy received the College Board AP Computer Science Female Diversity Award	Not Race-Related
It was also noted that our Equity and Diversity presentation to the NSAA went very well. The NSAA directors were impressed by the efforts we are taking to ensure that everyone within the district feels safe, welcome, and valued.	Race-Related, Affirmative
This fits into the school's long-term plan of closing the achievement gap for all students [including]: reading disparity for students entering 2 or more grade levels below when entering 7th grade, ethnic disparity, economic disparity, SPED student disparity	Race-Related, Affirmative
Expand institutional awareness by incorporating content and methodologies that promote justice and equity for all into the school curriculum, establish a district diversity council, …develop a strategic response to assessing and developing appropriate interventions for students when they return from distance learning to close the achievement gaps.	Race-Related, Affirmative
This model curriculum…focuses on skin color instead of focusing on what's important and that is content and character.	Race-Related, Oppositional
He has received questions about implementing Critical Race Theory. The district will continue to align itself with the WY Board of Education standards.	Race-Related, Oppositional
Comments by: …Leonard B - critical race theory - we're creating the problem by bringing it here, no right to teach it Jason M - Merry Christmas issue, critical race theory - don't need in this community	Race-Related, Oppositional
[Participant] applauded the district's equity work and asked the district to continue its dedication to the equity goals, [Participant] commented on Critical Race Theory citing a course from last summer and asked the district for greater transparency in how Critical Race Theory is being applied in the district.	Race-Related, Both
I'm a supporter of DEI and am committed to working within the bounds of my Board of Ed role to improving any inequities that are occurring, but I do not agree with the tenets of Critical Race Theory	Race-Related, Both

Note. Excerpts have been shortened and anonymized where appropriate.

Table 4

Validation of Classifier

	Accuracy	Precision	Recall
Race-Related	0.91	0.96	0.91
	(0.005)	(0.004)	(0.0)
Affirmative Statements	0.9	0.94	0.9
	(0.005)	(0.005)	(0.0)
Oppositional Statements	0.98	0.74	0.78
	(0.002)	(0.037)	(0.003)
Public Comment	0.97	0.9	0.96
	(0.01)	(0.02)	(0.01)

Note. Standard errors (in parentheses) are estimated using $\frac{\sqrt{\rho(1-\rho)}}{\sqrt{n}}$ where p is the performance metric and n is the number of observations in the denominator of Formulas 1-3. Segments which were hand labeled and/or classified as both affirmative and oppositional are included in the calculation for each respective performance metric.

Table 5

Proportion of Meetings within District with Race-Related Content, by Percentile and Stance

District Percentile	Any Race-Related	Any Affirmative	Any Oppositional
0.25	0	0	0
0.5	0.04	0.04	0
0.75	0.15	0.14	0.01
0.9	0.33	0.33	0.04
0.95	0.49	0.47	0.07

Note. The sum of the proportion of meetings with affirmative content and the proportion of meetings with oppositional content may be greater than the proportion of meetings with any race-related content. This is because statements may qualify as having both affirmative and oppositional components.

Table 6

Proportion of Meetings with Measure of Interest

Proportion of Meetings with…	
Any Race-Related Content	0.13
Affirmative Race Content	0.12
Oppositional Race Content	0.01
Affirmative Race Content from Officials	0.12
Affirmative Race Content from Public	0.02
Oppositional Race Content from Officials	0.01
Oppositional Race Content from Public	0.01

Note. Proportions estimated across all collected meetings within the study period (2018-2022).

Table 7

Frequency of Various Topics in a Random Sample of 50 Districts

Topic	Proportion of Meetings in Sample
Academics & Curriculum	0.38
Attendance & Enrollment	0.39
Pandemic Policies & Reports	0.33
Facilities & Technology	0.73
Finance & Budget	0.84
Personnel & HR	0.79

Note. Proportions of topics in a random sample of (up to) 10 meetings from 50 districts. 499 meetings total, as one district only had nine meeting minutes posted online.

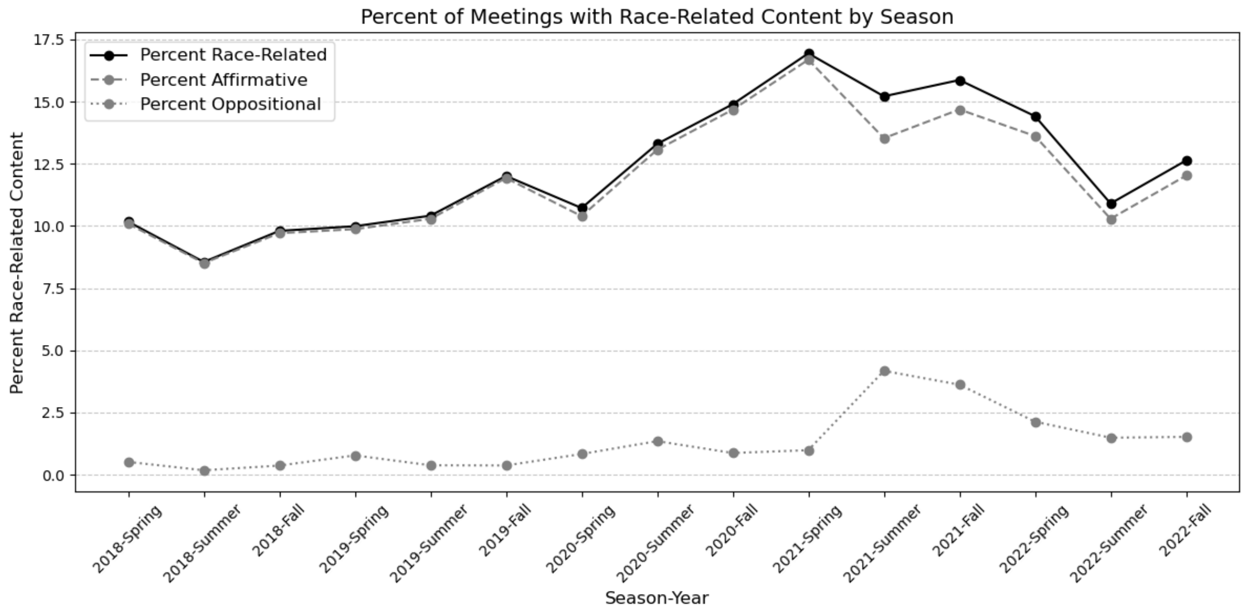
Table 8

Associations Between Race-Related Discussions in School Board Meetings and Regional, Demographic, and Community Characteristics (2018-2022)

	Race-Related Content		Affirmative		Oppositional	
	Individual Models (1)	Conditional Model (2)	Individual Models (3)	Conditional Model (4)	Individual Models (5)	Conditional Model (6)
Northeast	0.093***	0.042†	0.088***	0.036†	0.009*	0.006
	(0.02)	(0.02)	(0.02)	(0.02)	(0.0)	(0.01)
Midwest	0.023†	0.008	0.024*	0.009	-0.003	-0.003
	(0.01)	(0.02)	(0.01)	(0.02)	(0.0)	(0.0)
West	0.125***	0.062**	0.121***	0.056*	0.011*	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.0)	(0.01)
Urban	0.14***	0.027	0.139***	0.026	0.011*	0.005
	(0.03)	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)
Suburb	0.117***	0.017	0.117***	0.018	0.01**	0.005
	(0.02)	(0.02)	(0.02)	(0.02)	(0.0)	(0.0)
Town	0.012	-0.001	0.014	0.001	0.001	0.0
	(0.01)	(0.01)	(0.01)	(0.01)	(0.0)	(0.0)
Enrollment (1,000)	0.004**	0.002*	0.004**	0.002*	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Prop Black/Hispanic	0.001***	-0.0	0.001***	-0.0	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Prop White	-0.002***	-0.0	-0.002***	-0.0	-0.0†	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Prop Free/Reduced Lunch	-0.0	0.0	-0.0	0.0	-0.0	-0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Aggregate District Mean Test Score (weighted)	0.029	0.005	0.027	0.004	0.005	0.005
	(0.02)	(0.03)	(0.02)	(0.03)	(0.0)	(0.01)
Adults with BA	0.319***	-0.013	0.315***	-0.011	0.022*	-0.025
	(0.05)	(0.1)	(0.05)	(0.1)	(0.01)	(0.02)
Median Income in Thousands	0.002***	0.001†	0.002***	0.001†	0.0**	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Proportion Republican Presidential Vote Share	-0.004***	-0.003***	-0.004***	-0.003***	-0.0**	-0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)

Note. In individual models, Southern districts serve as the omitted category for geographic variables and rural districts serve as the omitted category for urbanicity variables. Standard errors are in parentheses and reflect clustering at the district level. (***p<0.001, **p<0.01, *p<0.05, †p<0.10.)

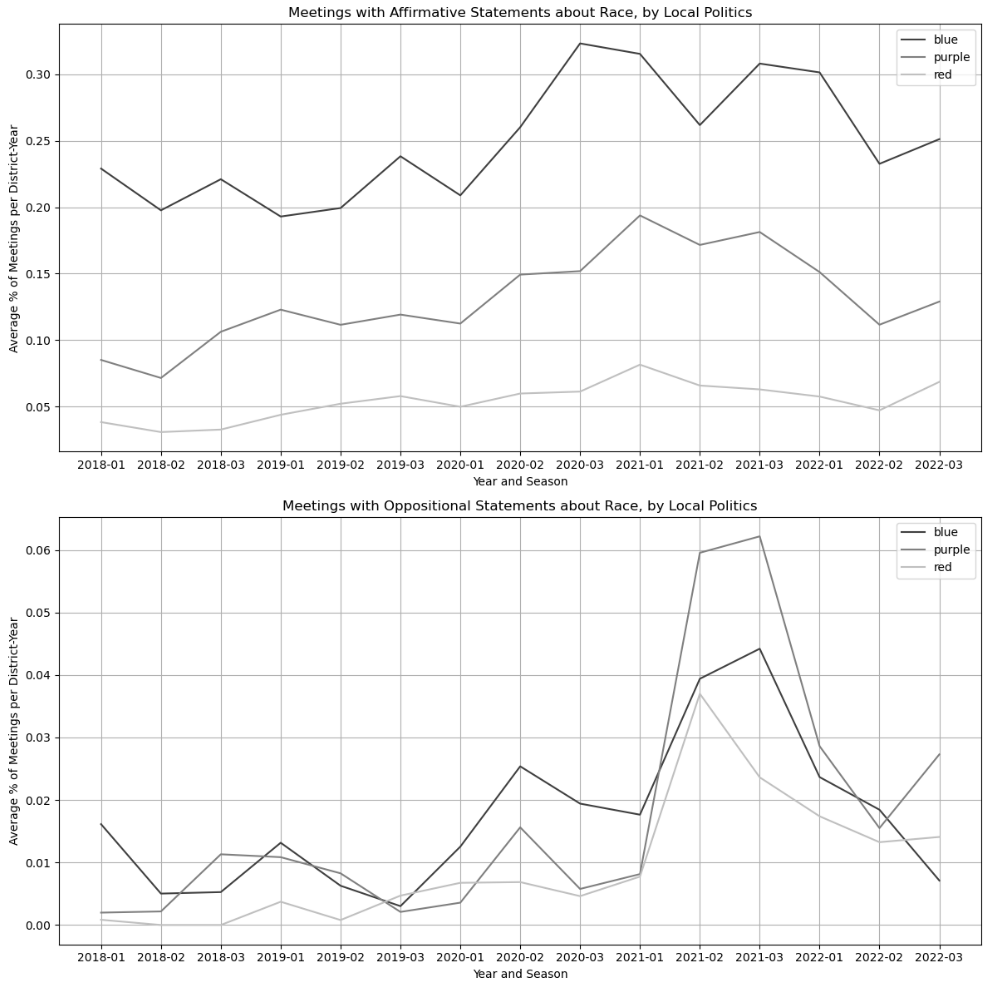
Figure 1



Note: For each year, spring refers to January through April, summer to May through August, and fall to September through December.

Figure 2

Affirmative and Oppositional Race-Related Content, by District Politics (2018-2022)



Note. Graph presents the proportion of meetings for each year, where 1 = January through April (spring), 2 = May through August (summer), and 3 = September through December (fall). Blue districts are defined as <.45 Republican 2016 presidential vote share, purple districts are defined as a Republican 2016 presidential vote share between 0.45 and 0.55, and red districts are defined as >.45 Republican 2016 presidential vote share.

Table 9

Relationship Between District Characteristics and Proportion of Meetings with Affirmative and Oppositional Race-Related Content

	Official Affirmative Content in 2018		Change in Official Affirmative Content 2018 to 2021		Oppositional Public Statements in 2021	
	Individual Models (1)	Conditional Model (2)	Individual Models (3)	Conditional Model (4)	Individual Models (5)	Conditional Model (6)
Northeast	0.053*	0.035	0.059*	-0.018	0.011	-0.002
	(0.02)	(0.03)	(0.02)	(0.03)	(0.01)	(0.01)
Midwest	0.013	0.018	0.002	-0.036	-0.012*	-0.017*
	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
West	0.084***	0.047†	0.057*	0.011	0.012	0.009
	(0.02)	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)
Urban	0.055†	-0.052	0.066*	0.06	0.008	0.003
	(0.03)	(0.03)	(0.03)	(0.04)	(0.01)	(0.01)
Suburb	0.059***	-0.019	0.09***	0.068**	0.017**	0.01
	(0.02)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
Town	-0.02	-0.03	0.022	0.03	0.004	0.004
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Enrollment (1,000)	0.003***	0.002***	-0.0	-0.001	0.0†	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Prop Black/Hispanic	0.001***	-0.0	-0.0	-0.001	0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Prop White	-0.001***	-0.0	-0.0	0.0	-0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Prop Free/Reduced Lunch	-0.0	-0.001	-0.001*	0.002*	-0.0*	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Aggregate District Mean Test Score (weighted)	-0.015	-0.078*	0.094***	0.123**	0.017**	0.009
	(0.02)	(0.03)	(0.02)	(0.04)	(0.01)	(0.01)
Adults with BA	0.169***	-0.05	0.232***	-0.115	0.059***	-0.002
	(0.05)	(0.12)	(0.05)	(0.14)	(0.02)	(0.04)
Median Income in Thousands	0.001***	0.001	0.002***	0.001	0.0***	0.0

	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Pct. Republican Vote Share	-0.003***	-0.003***	-0.001**	-0.001	-0.0	0.0
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Politically Competitive	0.017	-0.007	0.04†	0.019	0.017**	0.013*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)

Note. Standard errors are in parentheses. ***p<0.001, **p<0.01, *p<0.05, †p<0.10.

Online Appendix A

Expanding the Labeled Sample using Race-Related Terms

Because race-related segments in meeting minutes are rare—involving less than 3% of our segments—it is difficult to validate our LLM classifications by just using a representative sample of meeting segments. Therefore, we supplemented this representative sample with another sample of 1500 segments, each of which contained at least one of 70 race-related terms. We used a predominantly deductive approach, using word embeddings, to identify race-related terms. We began with an initial seed list of four terms: "race," "racism," "racial equity," and "critical race theory." This initial seed list was designed to invoke multiple perspectives on race, and was drawn from a review of academic literature, school board equity policies, and school board anti-CRT policies. (The term "racial equity," for example, is commonly used in diversity and equity initiatives, while "critical race theory" is commonly used in critiques of race-based equity initiatives.) Next, we expanded this list via an empirical analysis of word embeddings. Word embeddings map individual terms (i.e., words and phrases) to numerical vectors. The meaning of a given word, then, is represented by the vector coordinates, and the vector is optimized so that words appearing in similar contexts will be close together in vector-space (Mikolov et al., 2013). For example, the words "sickness" and "illness" will likely have a higher vector similarity than "garden" and "violin." Word vectors are commonly trained on very large volumes of online text, and the latest versions can elegantly handle homonyms and ambiguous meanings by considering a word's context (Devlin et al., 2019; Radford et al., 2018). In this paper, we used transformer-based embeddings from OpenAI, which were last pre-trained in September 2021 (OpenAI, 2024)—after the murder of George Floyd and the initial uptick in debates surrounding critical race theory. This timing is important because, for an embedding to adequately capture the meaning behind a phrase like "critical race theory," it must be trained on multiple documents containing that term. We used this seed list of embeddings to find additional race-related terms, identifying a further 50 terms within the corpus which were close in vector space to each of the terms in our seed list. For each of these terms, we chose a random selection of 10 transcripts containing the term and categorized the relevant excerpt as either: (a) clearly about race; (b) clearly not about race; or (c) unclear. To remain in our final list of terms, at least half of the excerpts containing the term needed to be clearly about race, and no more than 20% could be identified as unrelated to race.

Our final list included the following 70 terms (case insensitive): African, Asian, biases, bigoted, bigotry, biracial, black history, black lives matter (BLM), black male, black man, black student, black woman, Caucasian, civil rights, critical consciousness, critical race theory (CRT), culturally relevant, culturally responsive, desegregation, disparities, diversity, ethnic, ethnicity, European, Haitian, hatred, Hispanic, Indian, indigenous, inequality, inequitable, inequity, injustices, intersectional, intersectionality, Jewish, Latina, Latino, Marxist, microaggressions, multicultural, multiculturalism, multiracial, NAACP, Native American, people of color, racial, racism, racist, slavery, social justice, stereotype, unconscious bias, white fragility, white privilege, white supremacy, whiteness, and young black.

Online Appendix B

Analysis of LocalView Transcripts

Transcript Sample

We selected our sample of meeting transcripts from the LocalView database (see Barari & Simko, 2023). The LocalView database extends back to 2006, but we focused our analysis on meetings occurring between 2018 and 2022 in order to align with our principal analysis of meeting minutes data. During this period, LocalView collected text captions from 3739 probable school board meetings, from 150 unique entities that had uploaded meetings to YouTube. Of these, we excluded meetings from entities which could not be linked to a school district in the National Center for Education Statistics dataset, meetings which were shorter than ten minutes, and meetings lacking English captions. This generated a final sample of 2953 transcripts from 118 unique school districts. Of these, 47 are located within blue communities (<0.45 Republican presidential vote share in 2016), 45 are within red communities (> 0.55 Republican presidential vote share), and 24 are within purple communities (politically competitive districts with a Republican presidential vote share between 0.45 and 0.55). For comparison, Table B1 provides additional description of the sample characteristics alongside those of the primary sample. It shows, for instance, that the LocalView sample contains substantially more urban districts, with higher enrollment numbers, and more students of color.

Table B1

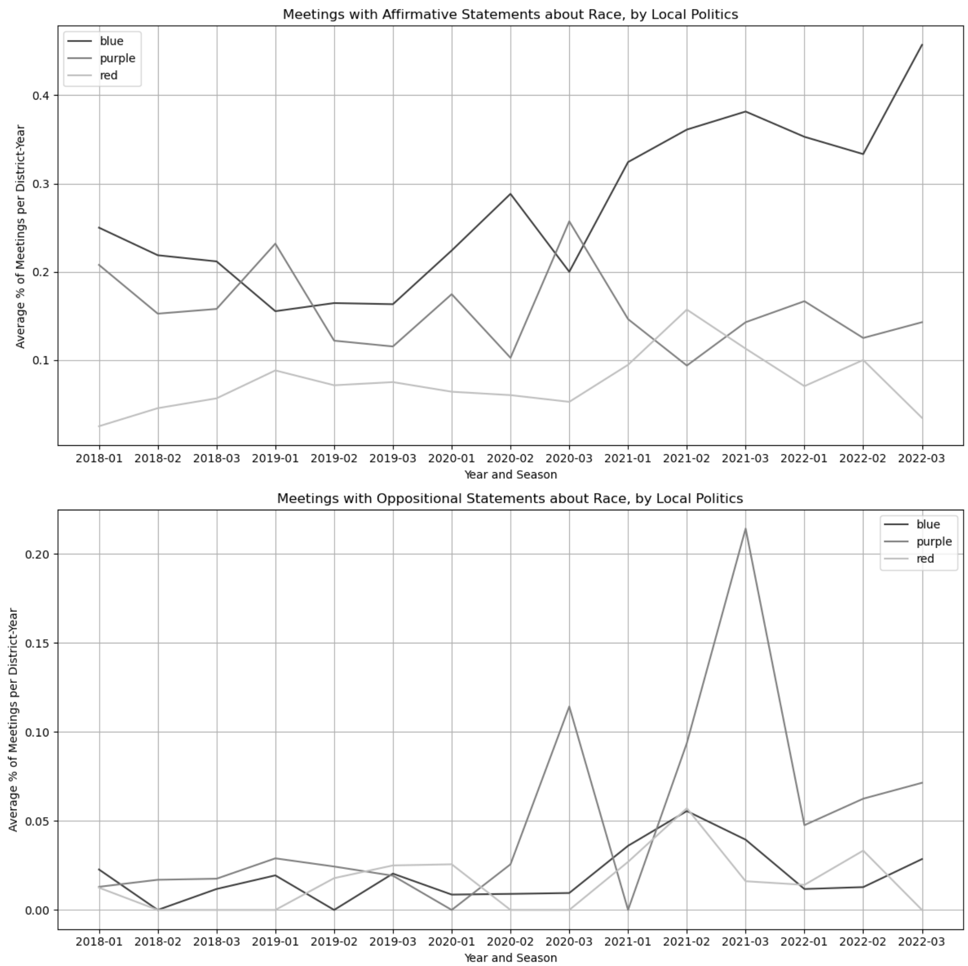
Comparison of Minutes Sample and LocalView Sample Characteristics

	Minutes Sample	LocalView Sample
Northeast	0.23	0.25
	[0.42]	[0.43]
Midwest	0.37	0.28
	[0.48]	[0.45]
South	0.24	0.27
	[0.43]	[0.44]
West	0.17	0.14
	[0.37]	[0.35]
Urban	0.08	0.22
	[0.26]	[0.42]
Suburb	0.29	0.26
	[0.46]	[0.44]
Town	0.22	0.25
	[0.41]	[0.43]
Rural	0.41	0.21
	[0.49]	[0.41]
Enrollment (1,000)	3.95	9.29
	[9.49]	[16.75]
Prop Black/Hispanic	0.25	0.3
	[0.27]	[0.28]
Prop Other	0.05	0.05
	[0.1]	[0.09]
Prop White	0.71	0.65
	[0.28]	[0.29]
Prop FRPL	0.48	0.49
	[0.23]	[0.24]
Aggregate district mean test score	0.05	0.05
	[0.36]	[0.39]
Prop adults in district with BA+	0.26	0.32
	[0.14]	[0.16]
District area-weighted Republican Presidential 2016 vote prop	0.58	0.49
	[0.18]	[0.21]
N	500	116

Trends in Race-Related Content over Time

We applied the same models to the segmented transcripts as were used to classify the minutes sample. The segments were classified as race-related or not, and, if race-related, as affirmative, oppositional, or both. We then aggregated to the meeting level, creating binary indicators for whether a meeting contained any race-related content.

Figure B1



Generally, we found race-related content to be more common in the LocalView transcripts. For example, in 2021, 25% of LocalView transcripts contained race-related content, compared to 14% of meeting minutes. Such differences in frequency likely reflect the transcripts' more comprehensive content and the samples' different composition; districts with greater proportions of students of color are perhaps more likely to discuss race, for example, and, as noted, these districts are disproportionately represented in the LocalView sample.

However, the trends observed in the LocalView sample were quite similar to those observed in the primary sample—featuring an increase in race-related prevalence between 2018 and 2021 (albeit less linear), followed by a decline in 2022. As with the primary sample, much of the observed increase is driven by a subset of districts. In any given year, 25% of LocalView districts have no identifiable race-related content. However, among the top quartile of districts (in terms of prevalence), race-related content featured in up to 40% of meetings in 2021. Notably, as in the primary sample, just 3% of meetings in 2021 contained oppositional content, and just 2% contained the phrase "critical race theory." Similarly, too, transcripts from blue districts were most likely to contain affirmative content, while those from purple districts were most likely to contain oppositional content.