School Accountability Measures: Foundations of a Penal System

Martha Perez-Mugg
University of Illinois, Champaign-Urbana

> The examination combines the techniques of an observing hierarchy and those of a normalizing judgement. It is a normalizing gaze, a surveillance that makes it possible to qualify, to classify and to punish. — Foucault.[1]

In *Discipline and Punish*, Foucault lays out the systematic and intentional mechanisms present within our society to exert power and control over the populace.[2] Nearly fifty years after publication, and in the wake of national discussions around police reform and racialized violence, we have made little progress in the dismantling of systems that perpetuate systemic racism and disenfranchisement in our society and school systems. To begin the process of building a more equitable society, we must turn our gaze upon the structures and systems still present within our society that exact punitive measures on communities of color. We need look no further than within our school systems to uncover problematic, structural elements that punish students, teachers, and communities alike under the guise of "school improvement" and "state accountability." At the root of this punitive system, we find a thriving system of what Foucault calls "hierarchical surveillance" grounded in standardized assessment, graduation rates, and school attendance measures.[3] In this essay, I will explore the ways in which state systems of accountability perpetuate systems of inequality in communities through hierarchical surveillance; disproportionately impacting communities of color through a matrix of accountability measures.[4] In the first section, I will describe the historical context and evolution of state accountability in the United States. Next, I examine accountability frameworks as a form of penal system that aims to discipline schools and communities.[5] Then, I examine the state of Colorado as an example to illustrate how state accountability serves to penalize schools and communities (disproportionately

© 2022 Philosophy of Education Society

targeting communities of color), concluding that school accountability systems work to obscure the elements of structural inequality that have resulted in school performance differences rather than creating systems of support rooted in transformational change. Lastly, I offer some alternative means for school improvement that are not rooted in a penal system.

STATE ACCOUNTABILITY FRAMEWORKS

Noting gaps in testing scores, graduation rates, and college degree attainment, federal and state governments have proposed many solutions to providing schools the support and accountability needed to improve student outcomes. Before 1965, educational funding and oversight was largely provided by local and state governments. However, following Lyndon B. Johnson's enactment of the Elementary and Secondary Education Act, the federal government began providing funding for disadvantaged students throughout the country. This funding opened the door to additional oversight on the part of the federal government and greater accountability for schools. By the 1980s, a narrative of underperformance in education resulted in calls for voluntary standardized testing and national education goals. By the 1990s and 2000s, states were under pressure to formulate state standards and assessments. In 2001, George W. Bush signed "No Child Left Behind" into law, formalizing a national system for school accountability through standardized assessment and standards. With the evolution of "No Child Left Behind" into "Race to the Top" and finally the "Every Student Succeeds Act" (ESSA), school accountability has become a cornerstone of educational reform in the United States.[6] With the codification of roles for federal and state governments in school improvement, systems have evolved to both track schools' progress but also hold schools accountable for their performance. These systems largely revolve around the use of high stakes standardized testing and other quantitative measures like graduation and attendance rates. Following the enactment of ESSA in 2018, states are required to use at least one measure of "school quality and student success" to determine a school's performance, and they are permitted to use student growth as a measure of school performance. According to the Education Commission of the States, 33 states and the District of Columbia plan to use a College and Career Readiness metric and 35 states and the District of Columbia plan to use

an attendance measure.[7] These systems have grown into a state apparatus of control, accountability, and ultimately serve as a disciplinary system. This system of accountability operates through hierarchical power, the operationalization of ranking of schools, and the threat of punitive measures at the hands of the state.

STATE ACCOUNTABILITY AS A PENAL SYSTEM

In the interest of unpacking the punitive structural elements laden within school accountability frameworks, we must first lay out an exposition of Foucault's theory of hierarchical power and punishment. In *Discipline and Punish*, Foucault offers a historically contextualized rendering of the early prison complex, drawing a myriad of connections between the institutionalized prison system and other social institutions like schools and hospitals. The text illuminates the ways in which hierarchical structures of power share both underlying functions in society and many similar mechanisms for exerting social control. Foucault examines the similarities between the layout of buildings to offer greater opportunities for observation and surveillance as well as analogous methods of examination in hospitals and schools. Further, he outlines three foundational components that support the penal system which I outline in the diagram below.

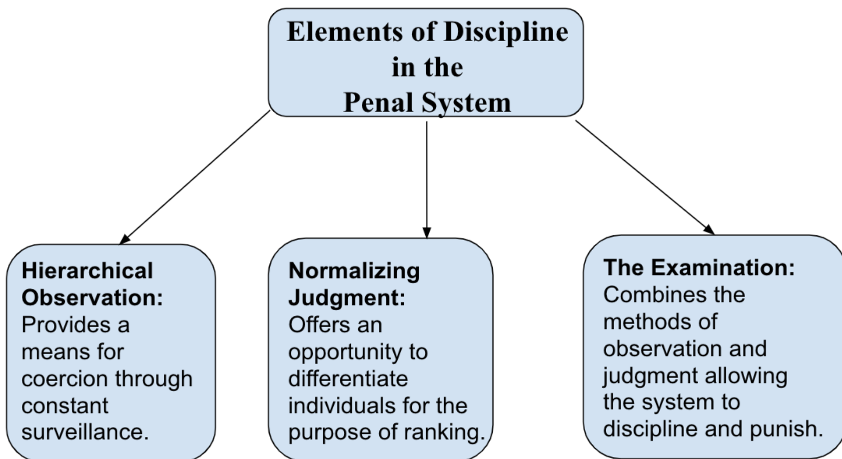

figure 1: *Elements of discipline that support the penal system. All these elements exist in current school accountability frameworks.*

DOI: 10.47925/78.3.001

First, hierarchical observation supports the power structure by offering a means of coercion for those in power. Individuals are observed regularly to both increase alignment to the norms and offer opportunities for normalizing judgment. As Foucault puts it: "The exercise of discipline presupposes a mechanism that coerces by means of observation; an apparatus in which the techniques that make it possible to see induced effects of power, and in which, conversely, the means of coercion make those on whom they are applied clearly visible."[8] Within state accountability systems for schools, regular observations occur at many different levels; data is collected regularly on a variety of different components for each school including standardized test scores, graduation rates, and attendance measures. Schools who are targeted based on previous low performance may experience additional surveillance in several ways: (1) The state monitors standardized test performance either at a yearly interval or potentially even more frequently through normed benchmark testing like MAPs or STAR. (2) Agents of the state and support staff from state offices conduct classroom observations at regular intervals. (3) Schools are required to prepare regular status updates for the state that include data on specific elements of interest. Consequently, schools are constantly aware that their performance in these categories is being monitored and experience the anxiety that comes along with hierarchical surveillance.

Next, states exact normalizing judgment on schools through their public ranking in the form of report cards and/or other reports that reveal their status in comparison to other schools. This systematic ranking of schools for comparative purposes impacts schools and communities in several ways. First, by ranking, the state can shape perceptions of schools and communities as a means of influencing enrollment in particular schools. Parents often consult school ratings to select which schools are the best fit for their children. This public ranking has the punitive effect of stripping enrollment from schools that are low performing, which in turn impacts their per pupil funding and exacerbates issues in already challenged school communities. Moreover, public ranking incentivizes schools to focus their improvement efforts on specific elements that the state surveils for the purpose of improving their status in the eyes of the community. Normalizing judgment establishes which specific elements are

worth focusing on, regardless of the needs of the specific school community. Consequently, normalizing judgment ensures that schools comply with the state framework and focus their efforts on the elements of improvement that the state deems important, not those elements that matter to the teachers, students, and school community.

Lastly, the examination system serves as a venue for the state to bring their data and judgment to bear on a community. The examination system allows the surveillance and judgment to take place through a standardized method that holds schools accountable across multiple community contexts. The examination itself is problematic, with a history of bias and unfair treatment of communities of color. Standardized testing persistently reveals gaps between the performance of students of color and white students, as well as students from lower income households and their wealthy peers. These gaps have been categorized as "achievement gaps" and have served as a call to action for educators to make drastic improvement to the educational outcomes of students of color and students from lower income communities. However, recent calls exist in the literature to recategorize these gaps as "opportunity gaps" or even "educational debt."[9] This draws attention to the persistent fact that standardized tests do not necessarily measure the learning of students alone, but rather capture inputs from teachers, communities, and a variety of factors present in a students' home life. Moreover, when standardized tests are used to both surveil schools and hold them in normalizing judgment, it unfairly holds schools in lower income communities accountable for gaps that they alone are not responsible for. This logic reveals a faulty assumption that it is the immediate actions of the school and the performance of individual teachers that result in achievement gaps and not the historical and social forces of structural inequality that result in lower standardized testing scores. This is not to say that schools have no impact on test scores, they do. However, as a national and state level policy, it calls into question why we would focus on penalizing schools for performance that is rooted in a social, historical, and contextual inequity regardless of the impact that schools do have on these test scores. Is it morally just to place the burden of student performance on schools when this performance has been impacted by decades of poor investments into schools and communities them-

doi: 10.47925/78.3.001

selves? Moreover, how does this surveillance and accountability system further perpetuate structural inequalities through normalizing judgment?

Furthermore, standardized testing has received significant attention in the literature as a problematic measure of student academic success. Wayne Au unpacks the historical legacy of standardized testing, tracing its history as a tool rooted in the ideology of eugenics and social engineering (not to mention explicitly racist notions of intelligence).[10] Au notes that, originally, standardized testing was proposed as a "value free" judgment of individual intelligence (and has continued to be portrayed in this way) and supported the sorting of students in the most fitting pathways for their future roles in society. Au argues:

> The historical roots of high-stakes, standardized testing in racism, nativism, and eugenics raises a critical question: why is it that, now over 100 years after the first standardized tests were administered in the United States, we have virtually the same test-based achievement gaps along the lines of race and economic class? Given the historical origins of standardized testing…there is no reason to believe that these testing systems could shake off their racist and classist legacies so easily.[11]

With a legacy of inequity and racial discrimination, it raises the question of whether standardized testing is a valid measure of school performance at all.

When we step back and consider the implications of utilizing standardizing testing and other forms of quantitative data to surveil and judge schools as a public policy, it raises questions about the impact of state surveillance and intervention on communities of color. Standardized testing is a problematic measure of student achievement and an even more problematic measure of school performance, with evidence suggesting that standardized tests reveal more about the context of a school community than the educational opportunities present within a school building. As Au argues: "As a racial project, high-stakes, standardized testing constructs which children (and communities) are identified as "failures" by the tests… and, subsequently, what policies and practices are to be enacted on those children and communities identified by the tests as 'failures.'"[12] This construction of a penal system centering on the use of high-stakes testing does a disservice to communities by enacting policies on

them, ultimately stripping them unfairly of the autonomy needed to improve outcomes for their students. In the next section, I will use the state of Colorado as a case study to explore the impact of this penal system on schools.

COLORADO: A CASE STUDY

The state of Colorado uses its school accountability system to publicly rank schools, hold schools accountable to the state board, and dissolve schools that are unable to make it out of "turnaround" status. Schools are measured quantitatively, earning points for different targeted school success factors based on the states perception of which factors reveal persistent problems regarding student outcomes. The state framework measures academic achievement, academic growth, and postsecondary readiness in all schools to provide school ratings and district accreditation. In the high school framework, schools are rated based on student achievement scores, student growth scores, student participation in standardized assessments, graduation rates, dropout rates, and matriculation to postsecondary programs as well as a variety of student population level data (see detailed scoring guide in figure 2). Throughout the course of the year the state collects formal data on every school, assigning specific points for different items based on a system of cut scores. Based on this point system schools are assigned a label and color (performance, improvement, priority improvement, or turnaround); low scores and assignment into the turnaround category begins the turnaround clock for a school and eventually triggers state intervention. Districts can also be held accountable through changes to their accreditation status (accredited with distinction, accredited, improvement, priority improvement, turnaround, or insufficient state data).

DOI: 10.47925/78.3.001

Scoring Guide for 2019 District/School Performance Frameworks					
Performance Indicator	*Measure/Metric*	*Rating*	*Point Value*		
Academic Achievement & ELP On Track Growth	*The district or school's mean scale score (or percent On Track) was*: see tables below for actual values*		*All Students*	*Each Disaggregated Group*	*ELP On Track Growth*
	• at or above the 85th percentile	Exceeds	8	1.00	2.0
	• at or above the 50th percentile but below the 85th percentile	Meets	6	0.75	1.5
	• at or above the 15th percentile but below the 50th percentile	Approaching	4	0.50	1.0
	• below the 15th percentile	Does Not Meet	2	0.25	0.5
	Students Previously Identified for a READ Plan (bonus point)				
	• CMAS ELA Mean scale score at or above 725 (Approaching Expectations cut-score)		1 bonus point		
Academic Growth	*Median Growth Percentile was:*		*All Students*	*Each Disaggregated Group*	*ELP*
	• at or above 65	Exceeds	8	1.00	2.0
	• at or above 50 but below 65	Meets	6	0.75	1.5
	• at or above 35 but below 50	Approaching	4	0.50	1.0
	• below 35	Does Not Meet	2	0.25	0.5
Postsecondary and Workforce Readiness	*Mean CO SAT Evidence-Based Reading and Writing (EBRW) scale score was**:*		*All Students*	*Each Disaggregated Group*	
	• at or above 554.7	Exceeds	4	1.00	
	• at or above 501.3 but below 554.7	Meets	3	0.75	
	• at or above 458.0 but below 501.3	Approaching	2	0.50	
	• below 458.0	Does Not Meet	1	0.25	
	*Mean CO SAT Math scale score was**:*		*All Students*	*Each Disaggregated Group*	
	• at or above 544.6	Exceeds	4	1.00	
	• at or above 488.0 but below 544.6	Meets	3	0.75	
	• at or above 439.9 but below 488.0	Approaching	2	0.50	
	• below 439.9	Does Not Meet	1	0.25	
	Dropout Rate: The district or school dropout rate was (of all schools in 2017):		*All Students*	*Each Disaggregated Group*	
	• at or below 0.5%	Exceeds	8	2.0	
	• at or below 2.0% but above 0.5%	Meets	6	1.5	
	• at or below 5.0% but above 2.0%	Approaching	4	1.0	
	• above 5.0%	Does Not Meet	2	0.5	
	Matriculation Rate (of all schools in 2018):		*All Students*		
	• at or above the 75.8%	Exceeds	4		
	• at or above 61.1% but below 75.8%	Meets	3		
	• at or above 46.8% but below 61.1%	Approaching	2		
	• below 46.8%	Does Not Meet	1		
	Graduation Rate and Disaggregated Graduation Rate (Best of 4-, 5-, 6-, or 7-year):		*All Students*	*Each Disaggregated Group*	
	• at or above 95.0%	Exceeds	8	2.0	
	• at or above 85.0% but below 95.0%	Meets	6	1.5	
	• at or above 75.0% but below 85.0%	Approaching	4	1.0	
	• below 75.0%	Does Not Meet	2	0.5	

figure 2: *The Scoring Guide for School Performance Frameworks used in the state of Colorado.*[13]

This system is founded on the assumption that schools who do not meet the quantitative expectations laid out by the state deserve both public scrutiny as well as punishment exacted through the loss of school autonomy. This framework aligns closely to the punitive model that Foucault outlines in *Discipline and Punish*. As you can see, standardized assessments serve as "the examination," offering the state an opportunity to cast normalizing judgment and rank schools. In Colorado's case, Colorado Measures of Academic Success (CMAS) and the PSAT/SAT provide data on student performance and growth for the state to assign schools points and color-coded categories. Once "the examination" takes place, the state casts normalizing judgment, formally assigning schools points and publicly ranking schools based on their performance. Finally, the constant monitoring by the state and requirement for schools to submit paperwork documenting their efforts to improve their standing in the rankings serves as the "hierarchical observation." School leaders are constantly aware that their standing within the framework could cost them school autonomy and even their state accreditation.

When states intervene there are many possible forms of disciplinary action that exist, but in the case of Colorado, state intervention often includes: additional surveillance and meetings with the state, state mandated supports that include the loss of autonomy for the district and school staff, demotion in accreditation status, and ultimately the hiring of an "external manager" that takes over day-to-day operations at the school or district level. Enactment of the 2009 Education Accountability Act allows for schools or districts to remain on the "turnaround clock" for five years before state intervention takes place. After the fifth year, the State Board of Education can convert the school or district into a charter school, grant the school or district "innovation status" to circumvent certain state requirements and rules, require the school or district to use an external manager within the district, require a charter management organization to take over the operations of the school, close the school, or force a district reorganization process.[14] Schools enter into a special relationship of surveillance called "performance watch" when they receive a rating of "priority improvement" or "turnaround" for two consecutive years and need to receive an improved rating for two consecutive years to exit "performance watch."

This heightened surveillance coupled with the expectation of improved performance can place significant strain on a school or district. Moreover, the expectation that schools can make drastic improvements within the scope of a few years seems to contradict what we know about which factors align to standardized test performance. Because standardized testing scores are reflective of the greater social and historical context of the community, it is problematic to expect dramatic changes to academic achievement within the scope of a few years. Yet, this system does serve a disciplinary purpose in penalizing schools for their performance. When we consider the implications of disciplinary actions for factors outside of the school's control it calls into question the validity of the penal system itself and whether the system serves in a just and fair way. When schools face closure or are converted into charter schools, it sends a clear message to the community that their institutions are inadequate. It labels the community and its residents as "in need of improvement." With the system of normalizing judgment, surveillance, and ultimately discipline it is the residents of the community that are penalized in the process. This disciplinary process

DOI: 10.47925/78.3.001

ultimately results in the uprooting of an institution from a community and the subsequent loss of voice that comes from having a communal institution stripped away or changed.

Hierarchical surveillance, normalizing judgment, and assessment exert power over schools, and subsequently, the communities that they serve. These measures illuminate the gaps among students and schools, bringing to light the ways in which students are not served well by the institutions in their communities. However, the faulty underlying assumption within this system is that schools alone face the burden of poor performance and that the path to success lies in school accountability rather than community support and interventions to address the underlying structural inequalities that result in testing gaps. In the case of Colorado, the state intervention system serves as a penal system with a disciplinary function. This system harms the community when it unfairly uses socially and historically produced differences to make changes to community institutions, especially when these changes come from the state body rather than community members themselves.

CONCLUSION

State governments spend a significant amount of time, energy, and money monitoring the performance of schools and enacting the previously described penal system. This surveillance of school activities is problematic for several reasons, and communities of color often face the greatest scrutiny resulting in loss of autonomy and other penalties due to school performance. I have outlined the ways in which this system is unfair to communities of color and ultimately unjust because it penalizes differences in school performance that are a result of social and historical factors. Yet, the question becomes: what alternative exists for improving student outcomes and school operations? Can we just leave schools to operate without surveillance and consequences for poor performance?

I think that the primary issue with the way we formulate "school accountability" lies in the hierarchical and disciplinary function derived from the penal system itself. Of course, schools want to improve their outcomes for students, and they likely appreciate all the support provided by governments to meet that end. Whether this support needs to happen within a penal system that

targets low-income communities and communities of color is another question. I am left with a few important questions with regard to the role of states in school improvement: How do communities measure the effectiveness of their schools? What voice should stakeholders have in determining the priorities for school improvements? Moreover, what role does a state government have in supporting communities to enact change? And why are state governments the body responsible for determining the course of change for a school community?

State governments have a role to play in the improvement of student outcomes, but this role may not need to be the one it currently holds. Student outcomes will improve when communities have the resources they need to thrive; when fair housing practices are enacted, when the disenfranchised have food, clothing, and utilities, and when the roots of structural inequalities are addressed head on. Poverty, neighborhood violence, and structural inequality all play a role in determining the outcomes for students in communities where the schools are deemed "low-performing." State governments have an enormous role to play in addressing these issues of poverty and structural inequality without reliance on coercive penal systems. Yet, these reforms may not have much to do with schools at all, but rather center on investment in underserved communities. I believe that there are better, more just alternatives to the penal system that currently exists for schools. First, funding that goes into the surveillance of school systems could be used to increase per pupil funding in historically underserved communities. Next, community voice could be leveraged to determine which shifts in practice will best serve the children in the community. And lastly, investment in community-based programming that combats the impacts of poverty and neighborhood violence might be used to improve student outcomes.

School funding is often a source of tension and rife with issues that result in tangible challenges for schools. According to the United Negro College Fund, schools with a concentration of 90% or more of students of color spend on average $733 dollars less per student than schools with 90% or more white students.[15] This inequity is even more problematic when considering that these schools are often the target of school accountability measures. Changes to school funding might not guarantee better student outcomes, but these shifts

doi: 10.47925/78.3.001

would likely result in the reduction of problematic gaps that are the result of structural inequality. This might serve as one small step towards addressing historical social inequalities. Moreover, communities deserve a voice in how schools serve as partners in the education of children. Through direct community outreach and action, communities can work to improve schools in ways that might result in greater belonging for students, more equitable curricular choices, and more alignment between the values of the community and the school itself. Rather than further disenfranchising communities of color through the loss of autonomy, which is the current state of school accountability, states should facilitate greater community involvement and ownership. Lastly, there are many investments that can be made to alleviate poverty and violence in communities that will do more to benefit students than school improvement alone. This direct action might serve to correct the underlying inequalities that lead to testing gaps in the first place. These community-based changes may make a lasting impact on students and their families by addressing a root cause for poor academic achievement, rather than addressing perceived failures within the school itself. State governments have tremendous power over the implementation of social safety nets and the benefits they confer to society; rather than utilizing accountability mechanisms, states can invest in communities in ways that will result in tangible benefits for students.

Our current state and national school accountability policies place unfair and unjust responsibility on low-performing schools without acknowledging the social and historical structural inequalities that produce educational gaps in communities of color. These penal systems result in disciplinary actions for schools who are placed on the accountability clock. Moreover, use of problematic measures like standardized testing scores call into question the validity of school accountability frameworks in identifying which schools are truly low performing, rather than merely identifying communities impacted by structural inequality. More equitable and just systems would involve equitable school funding, greater community autonomy and involvement, and the expansion of programming to address the root causes of poor school performance.

1 Michel Foucault, *Discipline and Punish: The Birth of the Prison* (New York: Pantheon Books, 1977): 184.

2 While Foucault's work offers a myriad of ways to understand state power as a penal system for schools, the framework provided by *Discipline and Punish* offers the clearest opportunity for examining the mechanisms at work in penalizing communities of color disproportionately under the guise of school improvement. Due to restrictions on time and space, I confine my discussion of Foucault to *Discipline and Punish*.

3 According to the Education Commission of the States, following the enactment of ESSA in 2018, 45 states plus the District of Colombia have enacted some form of summative rating systems to rank school performance. State rating systems fall into five categories: (1) Those that use an A-F report card style (used by 13 states). (2) Those that use a descriptive categorization- for example, labels like "needs improvement" (used by 11 states). (3) Those that use a numerical score- for example, 1-100 (used by 12 states). (4) Those that use a "support system" where they indicate the level of support required for the school (used by 6 states). (5) Those that use a system of stars to indicate quality (used by 4 states and the District of Columbia). These frameworks are nearly ubiquitous in the United States and serve as the foundation of what I describe as a "school accountability system." Ben Erwin et al., "50-State Comparison: States' School Accountability Systems," *Education Commission of the States*, 2021. accessed September 15, 2021, https://www.ecs.org/50-state-comparison-states-school-accountability-systems/

4 Throughout this paper, I use the word "state" to refer to the broad state apparatus of accountability but also the specific states in the United States. While this might be confusing at times, individual states do have power within the educational system to enact specific disciplinary measures and these measures contribute to the overarching mechanisms of state accountability.

5 In the framework of this paper, I use the term "penal system" to refer to the disciplinary elements upheld by state accountability frameworks in the United States. While these systems have a positive intent, aimed at school support and

DOI: 10.47925/78.3.001

improvement, the disproportionate impact of these systems acts to penalize communities of color. Although the intent of these systems may be disciplinary in nature, the overall effect is one of punishment.

6 Alex Spurrier et al., "The Historical Roots and Theory of Change of Modern School Accountability*," Bellwether Education Partners,* 2020. accessed October 21, 2021, https://files.eric.ed.gov/fulltext/ED606415.pdf

7 Erwin et al., "50-State Comparison: States' School Accountability Systems."

8 Foucault, *Discipline and Punish*, 170-171.

9 Gloria Ladson-Billings, "From the Achievement Gap to the Education Debt: Understanding Achievement in U.S. Schools," *Educational Researcher* 35, no. 7 (2006): 3-12.

10 Wayne Au, "Hiding Behind High-Stakes Testing: Meritocracy, Objectivity and Inequality in U.S. Education," *International Education Journal: Comparative Perspectives*, 12, no. 2 (2013): 7-20.

11 Wayne, ""Hiding Behind High-Stakes Testing," 12.

12 Wayne Au, "Meritocracy 2.0: High-Stakes, Standardized Testing as a Racial Project of Neoliberal Multiculturalism," *Educational Policy* 30, no. 1 (2016): 42-43.

13 Colorado Department of Education, "Scoring Guide for 2019 District/School Performance Frameworks," *CDE*, 2019. accessed October 1, 2020, http://cde.state.co.us/accountability/2019-framework-scoring-guide_080319

14 Colorado Department of Education, "Accountability Clock," *CDE*, accessed October 25, 2021, https://www.cde.state.co.us/communications/accountabilityclockfactsheet

15 United Negro College Fund, "K-12 Disparity Facts and Statistics," *UNCF*, accessed September 15, 2021, https://uncf.org/pages/k-12-disparity-facts-and-stats