SPECIAL ISSUE
Teachers and Educational Policy: Markets, Populism, and Im/Possibilities for Resistance

education policy analysis archives

A peer-reviewed, independent, open access, multilingual journal

Arizona State University

Volume 30 Number 101 July 19, 2022 ISSN 1068-2341

"It's the Teachers' Fault": Personalizing the Gaps with School-Based Value-Added Measures[1]

Rachel Garver
Montclair State University
United States

Citation: Garver, R. (2022). "It's the teachers' fault": Personalizing the gaps with school-based value-added measures. *Education Policy Analysis Archives, 30*(101). https://doi.org/10.14507/epaa.30.6186 This article is part of the special issue *Teachers and Educational Policy: Markets, Populism, and Im/Possibilities for Resistance,* guest edited by Meghan Stacey, Mihajla Gavin, Jessica Gerrard, Anna Hogan and Jessica Holloway.

Abstract: Efforts to measure teacher effectiveness have intensified across the globe over the last two decades. In the United States, educational reforms that attempt to further educational equity have concentrated on aspects of schooling, such as teacher quality, that are more easily manipulated and monitored than powerful out-of-school factors, such as economic and racial segregation. Teachers in high-poverty, racially segregated schools are subject to strengthened accountability policies that seek to precisely evaluate teachers as they face student needs that research shows are difficult to address fully within the classroom. How does this disconnect between causes and remedies shape equity work in high-poverty, segregated schools? I examine how the administration at a public middle

[1] This article is based on research that was made possible in part by financial assistance from the Ruth Landes Memorial Research Fund, a program of the Reed Foundation.

Journal website: http://epaa.asu.edu/ojs/
Facebook: /EPAAA
Twitter: @epaa_aape

Manuscript received: 10/2/2020
Revisions received: 11/16/2021
Accepted: 12/7/2021

school in the United States held teachers independently responsible for the comparative performance of each student subgroup. This project of "personalizing the gaps" involved developing a theory of action that linked performance gaps to classroom practice and creating a value-added data set that traced the growth of each subgroup in teachers' classrooms. In turn, teachers 'personalized the gaps' when under direct administrative supervision and at other times resisted by proposing alternative explanations for performance gaps or questioning the validity of the administration's data practices.
Keywords: teacher evaluation; value added models; data use; middle schools; ethnography; neoliberalism

"La culpa es de los docentes": Personalizando las brechas con medidas de valor agregado en las escuelas

Resumen: Los esfuerzos para medir la eficacia docente se han intensificado en todo el mundo en las últimas dos décadas. En los Estados Unidos, las reformas educativas que intentan promover la equidad educativa se han concentrado en aspectos de la educación, como la calidad de los maestros, que son más fáciles de manipular y controlar que los factores externos, como la segregación económica y racial. Los maestros de escuelas segregadas racialmente y con un alto nivel de pobreza están sujetos a políticas que buscan evaluar con precisión a los maestros que enfrentan necesidades de los estudiantes que, según muestran las investigaciones, son difíciles de abordar por completo dentro del salón de clases. ¿De qué manera esta desconexión entre las causas y los remedios da forma al trabajo de equidad en las escuelas segregadas de alta pobreza? Examino cómo la administración de una escuela secundaria pública en los Estados Unidos responsabilizó a los maestros de manera independiente por el desempeño comparativo de cada subgrupo de estudiantes. Este proyecto de "personalizar las brechas" involucró el desarrollo de una teoría de acción que vinculaba las brechas de desempeño con la práctica en el aula y creaba un conjunto de datos de valor agregado que rastreaba el crecimiento de cada subgrupo en las aulas de los maestros. A su vez, los docentes "personalizaron las brechas" cuando estaban bajo supervisión administrativa directa y en otras ocasiones se resistieron proponiendo explicaciones alternativas para las brechas de desempeño o cuestionando la validez de las prácticas de datos de la administración.
Palabras clave: evaluación docente; modelos de valor agregado; uso de datos; escuelas intermedias; etnografía; neoliberalismo

"A culpa é dos professores": Personalizando as lacunas com medidas de valor agregado nas escolas

Resumo: Os esforços para medir a eficácia dos professores se intensificaram em todo o mundo nas últimas duas décadas. Nos Estados Unidos, as reformas educacionais que buscam promover a equidade educacional concentraram-se em aspectos da escolarização, como a qualidade do professor, que são mais facilmente manipulados e monitorados do que fatores externos, como a segregação econômica e racial. Professores em escolas de alta pobreza e segregadas racialmente estão sujeitos a políticas que buscam avaliar com precisão os professores cujas necessidades dos alunos que a pesquisa mostra são difíceis de abordar plenamente dentro da sala de aula. Como essa desconexão entre causas e remédios molda o trabalho de equidade em escolas segregadas e de alta pobreza? Examino como a administração de uma escola pública de ensino médio nos Estados Unidos responsabilizou os professores independentemente pelo desempenho comparativo de cada subgrupo de alunos. Esse projeto de "personalizar as lacunas" envolveu o desenvolvimento de uma

teoria de ação que vinculava as lacunas de desempenho à prática em sala de aula e criava um conjunto de dados de valor agregado que acompanhava o crescimento de cada subgrupo nas salas de aula dos professores. Por sua vez, os professores 'personalizaram as lacunas' quando sob supervisão administrativa direta e outras vezes resistiram ao propor explicações alternativas para as lacunas de desempenho ou questionar a validade das práticas de dados da administração.
Palavras-chave: avaliação de professores; modelos de valor agregado; uso de dados; escolas de ensino médio; etnografia; neoliberalismo

"It's the Teachers' Fault": Personalizing the Gaps with School-Based Value-Added Measures

Efforts to measure, monitor, and incentivize teachers have intensified across the globe over the last two decades (Echávarri & Peraza, 2017; Flores & Derrington, 2017; La Londe, 2017; Straubhaar, 2017; Verger & Curran, 2014). International organizations and policy entrepreneurs have disseminated new technologies of teacher accountability, and their work has been fueled by international assessments that foster country-to-country comparisons in an increasingly competitive worldwide market (Ball, 2016; Echávarri & Peraza, 2017; Flores & Derrington, 2017; Tatto, 2006; Verger & Curran, 2014). However, the way that teacher supervision and evaluation have been implemented varies by country, depending in part on local conceptions of the teaching profession and of good teaching (Tatto, 2006).

In the United States (US), the rise of teacher accountability ran parallel to a greater national focus on equitable academic outcomes (Superfine, 2013), both of which were facilitated by the vast amounts of student performance data made available with mandated standardized testing. Educational reforms that attempt to further educational equity have concentrated on aspects of schooling, such as teacher quality, that are more easily manipulated and monitored than powerful out-of-school factors, such as economic and racial segregation (Berliner, 2009, 2013; Downey & Condron, 2016). Teachers and administrators in high-poverty, racially segregated schools are caught at the crux of the contradiction between the causes and remedies for performance gaps.[2] They are subject to strengthened accountability policies that seek to carefully measure and evaluate teachers as they face student needs that are difficult to address fully within the classroom. How does this disconnect in high-poverty, segregated schools between causes and remedies shape equity work—broadly defined as the pursuit of a specified equitable educational outcome through policy or practice? How do teachers and administrators in segregated schools manage the contradiction between the causes and remedies for educational inequities?

In this paper, I examine how the administration at an urban, public middle school in the United States, Lytle Middle School (LMS)[3], held teachers independently responsible for the comparative performance of each student subgroup in their effort to promote educational equity. This project of "personalizing the gaps" involved developing a theory of action that linked student performance gaps to teachers' classroom practices, creating a data set to measure each teacher's value-added for each student subgroup, and shaping teachers' mindsets so that they took ownership of subgroup performance gaps and ignored alternative causes for inequities. In turn, the teachers personalized the gaps when under direct administrative supervision and at other times resisted the project by proposing different approaches to equity work or questioning the validity of the administration's data practices. The extensive efforts made to implicate teachers in the school's performance gaps indicate how administrators draw upon symbolic and technical tools within the

policy environment to manage working at the crux of education policy's contradiction. This paper illustrates how marrying equity work with teacher accountability associates teacher professionalism with the full ownership of inequalities that are beyond educators' sole control. The case further shows how this marriage is an unpromising approach to meaningfully investing teachers in addressing educational inequity.

Policy Versus Research on What Matters for Student Performance

Since the Coleman Report (1966), scholars of education have attempted to measure to what extent performance gaps are due to school-based, neighborhood, family, or student-level factors. At stake in this line of inquiry is an understanding of what schools can do, if anything, to improve students' outcomes. Although different modes of inquiry and statistical models have led to mixed findings, scholars have consistently found that out-of-school factors matter more in determining a student's academic performance and trajectory than school factors (Card & Rothstein, 2006; Egalite, 2016; Jencks, 2016; Rothstein, 2004). Poverty and its effects, such as poor medical care, exposure to pollutants, and food insecurity, are the most significant obstacles to student achievement (Berliner, 2009; Chetty et al., 2015; Egalite, 2016; Rothstein, 2004), and concentrated poverty—reflecting residential economic segregation—limits a school's ability to meet the needs of its student body (Berliner, 2009; Owens et al., 2016).

At the same time, the simplistic and widespread interpretation of the Coleman Report's findings that 'schools don't matter' has been dismissed by a more precise distinction between the reproductive and compensatory features of schooling (Downey & Condron, 2016; Jennings, et al., 2015). One finding that captures the importance of out-of-school and in-school factors is that children enter kindergarten with a large socioeconomic performance gap, however this gap does not grow when school is in session (Downey & Condron, 2016). On the other hand, research has shown that schools exacerbate racial performance gaps and that their compensatory impact on socioeconomic gaps may not hold beyond elementary school (Gamoran, 2016; Palardy, 2015), reinforcing Downey and Condron's (2016) call for appreciating schools' refractive effects.

Scholars have been dedicated to distilling which school-level factors and teacher practices have the most impact on student performance (e.g., Hattie & Anderman, 2020). Among the school-level factors important for student learning, teachers and tracking account for the most variance in achievement (Chetty et al., 2015; Gamoran, 2016; Goldhaber, 2016; Hanushek, 2016; Kalogrides & Loeb, 2013; Palardy, 2015; Raudenbush, 2015). Value-added growth models show that teachers matter, but we do not know which of their characteristics or practices specifically are responsible for their influence (Hanushek, 2016; Raudenbush, 2015). Easily measurable aspects of teaching, such as years of experience and educational preparation, have a weak relationship to student achievement (Downey & Condron, 2016). Despite the power of schools to impact student achievement, they cannot overcome the effects of out-of-school factors (Huang & Sebastian, 2015). Schools may compensate for out-of-school factors that disadvantage students to a certain point beyond which they are no longer effective (Downey & Condron, 2016; Egalite, 2016).

Policy approaches to ameliorating socioeconomic and racial performance gaps have not reflected the research on school effects. Based on the scholarship, policy interventions that seek to promote income equality would be most effective (Berliner, 2013). Education policy has been too narrowly focused on schools; arguably, the most important education policies would increase the minimum wage, expand free medical insurance, reduce environmental pollutants, and minimize unemployment (Dobbie & Fryer, 2009; Egalite, 2016). Within the education system, student assignment policies that integrate students by socioeconomic status and race would help attend to the negative impact of concentrated poverty (Card & Rothstein, 2006; Chetty et al., 2015). Yet,

public support for desegregation has waned, and legal obstacles to desegregation programs have compounded (Owens et al., 2016). Unsurprisingly, the Black-white performance gap and the socioeconomic performance gap have not improved over the last 50 years (Hanushek, 2016). Despite the preoccupation with increasingly regulating and measuring teacher quality to improve schools, teachers' abilities to attend to inequalities are limited (Blazer et al., 2016; Palardy, 2010).

Teacher Accountability and Value-Added Measures

The accountability movement's initial focus on measuring student performance subsequently turned its attention to measuring teacher effectiveness and the relationship between the two. The intensified attention paid to teacher quality has promoted the assumption that teachers are both the problem and the solution behind the failings of U.S. public schools, including racial and economic inequalities in student outcomes (Amrein-Beardsley & Holloway, 2017; Berliner, 2009). Teacher accountability attempts to precisely assess teacher quality and includes evaluative tools such as detailed rubrics that dissect teacher practice, regular administrator observations, student surveys, and value-added measures.

In the United States, value-added measures were lauded by the Obama administration and incentivized by their Race to the Top initiative and their No Child Left Behind waivers as a way to ensure teacher accountability. Value-added measures link teacher ratings to students' growth on standardized tests, which is understood as an improvement over measuring teacher performance according to whether students reach an established threshold (AERA Council, 2015). A large body of literature has examined the reliability and validity of value-added measures, and numerous concerns have been raised by professional organizations and researchers, including: the difficulty of isolating the impact of one teacher; the exclusion of peer effects, school composition, and student attitudes about testing; missing data; small sample sizes; inherent bias against teachers whose student populations show little growth (e.g., very low- and very high- performing students); and the reliance on a single test (AERA Council, 2015; Amrein-Beardsley, 2014; Amrein-Beardsley & Holloway, 2017; Berliner, 2013; Shifrer, 2020). While value-added measures carry the veneer of objectively assessing individual teachers' effectiveness, student growth measures typically indicate a teacher's percentile in relation to their colleagues and thereby dictate that some teachers will be below average, irrelevant of their students' actual performance (Amrein-Beardsley & Holloway, 2017). Aside from the concerns about validity and reliability, evidence suggests that value-added measures may demoralize and demotivate teachers (Jennings & Pallas, 2016), and it is not clear that pay for performance programs that link teachers' value-added measures to financial incentives operate as intended (Amrein-Beardsley & Holloway, 2017). Lacireno-Paquet, Bocala, and Bailey (2016) surveyed a nationally representative sample of teachers and found that teachers whose ratings included student test scores in some manner were two and a half times less likely to be satisfied with their evaluation system. Despite these concerns, value-added measures have been widely adopted as one component of a teacher's evaluation across the United States. Although teacher evaluation guidelines became more flexible under the Trump administration, states and districts had already adopted and invested in value-added measures and many have continued to use them.

School leaders play an important role in buffering or brokering connections between contested teacher accountability policies, such as value-added measures, and school-level practices (Honig & Hatch, 2004; LeChasseur et al., 2018). Teacher perceptions of evaluation policies are based largely on their experiences with implementation in their school—within their local professional culture (Braun, 2015).

Theoretical Framework

Teacher accountability and value-added measures specifically are reflective of the rise of neoliberal reforms to education. Neoliberal approaches to school improvement have centered private sector, free-market values such as competition and choice and are characterized by extensive monitoring systems to track the individual performance of students, teachers, and administrators (Ball, 2000, 2003, 2016). Building on Lyotard (1984), Ball (2003) refers to this practice of constant assessment and inspection as performativity, defined as "a technology, a culture and a mode of regulation that employs judgements, comparisons and displays as means of incentive, control, attrition and change—based on rewards and sanctions (both material and symbolic)" (p. 216). Performance measures claim to "encapsulate or represent the worth, quality or value of an individual or organization within a field of judgement," (Ball, 2003, p. 216) such as the field of education. Accordingly, neoliberal reforms have introduced a multitude of metrics for school and teacher quality, impact, and effectiveness.

In response to the pressures of performativity, schools and teachers expend considerable energy on fabricating a representation of themselves that will favorably communicate their value according to given neoliberal measures (Ball, 2003; Lewis & Holloway, 2019). These neoliberal measures are a reflection of the state's definition of 'good teaching' and provide teachers with the language and opportunity to demonstrate excellence (Ball, 2016; Holloway, 2021; Lewis & Holloway, 2019). They create a "vocabulary of performance" to "talk about oneself and others" (Ball, 2003, p. 218), often with numbers.

Evaluating individual quality and comparing workers are facilitated by the widespread use of quantitative performance measures (Piattoeva, 2021) that seem "misleadingly objective and hyper-rational" (Ball, 2003, p. 217). Faith in quantitative data and data-driven-decision-making is central to neoliberal education reform (Spillane, 2012). The culture of performativity purports a linear, rational connection between measurement, increased data, and improved outcomes. According to a neoliberal logic, teacher quality will improve as the work of teachers is more regularly surveilled and precisely measured.

Often performance metrics are created by reconfiguring and individualizing measures of collective and complex phenomena. In her study of teacher accountability in Russia, Piattoeva (2021) explains how large-scale educational assessments are translated into measures of individual teachers' performances: "the measures invented at the federal and regional levels as synoptic overviews, become, in the school context, individualizing instruments evaluating individual employees" (p. 525). In the same way, wicked problems, such as racial and socioeconomic inequality in schools, may become simplified and attributed to individual teachers.

Neoliberal reforms and the culture of performativity have reconstituted what it means to be a teacher (Anderson & Cohen, 2015; Ball, 2000, 2003, 2016; Holloway, 2021; Weiner, 2020). Some have claimed that neoliberal reforms have deprofessionalized teaching, which previously enjoyed a level of autonomy and self-regulation representative of professionalism, while others are wary of idealizing past conceptions of teaching and characterize the shift as a reprofessionalization (Holloway, 2021; Weiner, 2020). Given that the profession of teaching has been redefined and reconstituted repeatedly, what is important to critically examine is who has the power and authority to create and implement new performative measures that indicate what good teaching is and who is a quality teacher (Ball, 2000; Holloway, 2021). Although accountability policies may come from the federal, state, or district levels, the way that they are implemented and experienced by teachers is largely due to the practices of school administrators (Hall & McGinty, 2015; Honig & Hatch, 2004; Spillane et al., 2002), and, in some cases, neoliberal reforms have given local school administrators

more authority and discretion to manage and evaluate teachers (Tatto, 2006; Tuytens & Devos, 2017).

School administrators broker connections or create buffers to policies from the district, state and federal levels, crafting (in)coherence between external mandates and internal initiatives (Honig & Hatch, 2004) and influencing the extent to which school policies, organizational routines, and instruction are tightly or loosely coupled with external pressures (Meyer & Rowan, 1977). Spillane (2012) explains how school leaders often draw upon policy logics in the field of education to design their own projects locally such that "the macro informs the micro" (Spillane, 2012, p. 7). Through this process, school actors reshape the 'raw materials' found in the institutional environment to create unique local initiatives (Spillane, 2012). Closely surveilled by the state, schools that serve low-income communities of color are likely to see less buffering from administrators and a greater influence of neoliberal reforms (Scott, 2011). The culture of performativity is intensified in contexts where past performance on assessments has been low and the avoidance of sanctions, such as administrative turnover, teacher excessing, or organizational restructuring, relies on a favorable rating, whereas comparatively high-ranking institutions can mitigate performance pressures (Ball, 2003).

Teachers have responded to neoliberal reforms as mediated by school administrators in several ways that demand further study and conceptualization beyond a dichotomous scale of compliance and resistance (Holloway, 2021; Moore & Clarke, 2016; Weiner, 2020). Some teachers experience an internal conflict between the culture of performativity and their personal, professional commitments to foster the academic and social development of youth (Ball, 2003; Moore & Clarke, 2016). Teachers critically engage with the quantitative measures of their quality employed by the state, demonstrating that the stories numbers communicate are subjective, contextual, and emotional (Piattoeva, 2021). However, as neoliberal reforms have persisted over the last decades, some teachers have come to embrace and find pleasure in performing professionalism as defined by accountability metrics (Holloway & Brass, 2017; Weiner, 2020). Even teachers who remain critical of accountability policies may conform, as professionalism is tied up with compliance, the imaginary for alternatives is limited by internalized neoliberal logics, and teacher resistance is perceived as unprofessional and backward-looking (Anderson & Cohen, 2015; Ball, 2003; Hall & McGinty, 2015). Teachers may be torn between a desire to be seen as valuable through accountability metrics while they remain critical of the metrics as valid reflections of their practice.

How teachers take-up neoliberal accountability policies and practices is highly influenced by school-level implementation (Hall & McGinty, 2015), and therefore by school administrators. As local managers working with teachers on a daily basis, being an administrator "involves instilling the attitude and culture within which workers feel themselves accountable and at the same time committed or personally invested in the organization" (Ball, 2003, p. 219). The relationship between the practices of school administrators and teacher responses to neoliberal reforms requires further study.

Methodology

Ethnographic Case Selection

This paper draws from a one-year ethnographic study at Lytle Middle School (LMS), a public sixth-eighth grade school in a large urban school district with approximately 750 students. In order to examine how educational equity is understood and pursued in an economically and racially segregated school, I selected LMS for its diverse racial demographics and for its citation under No Child Left Behind (NCLB; the 2001 reauthorization of the federal Elementary and Secondary

Education Act) for failing to demonstrate adequate yearly progress for three student subgroups. I reasoned that the school's racial and ethnic diversity as well as the high rate of poverty in the neighborhood would create increased ambiguity about comparative student need for educators and therefore about how to pursue equity work. Asian, Black, and Latinx students each constituted more than 20% but less than 50% of the study body, and approximately 15% of students were identified as English learners (ELs) and 15% as students with disabilities. The entire student body qualified for free lunch. Fewer than 30% of students at LMS tested proficient on the state math or the state English language arts (ELA) exams the year prior to my study. Based on ranking amongst the lowest performing schools in the state as measured by subgroup performance on state standardized assessments, LMS had been designated as a "Focus" school under NCLB (a category that ceased to exist with the introduction of a new set of categories to hierarchize and compare schools under the Every Student Succeeds Act). Accordingly, LMS had been deemed by the state as an underperforming school and the administration was under pressure to fabricate an image of the school as improving according to measures intelligible within the regime of neoliberal accountability policies (Ball, 2003). After I identified LMS as an ideal case and acquired the principal's tentative agreement for the study, I obtained permission for the study from my university and the district Institutional Review Board.

An ethnographic approach allowed me to discover inductively how LMS approached equity work. With the initial intention of understanding how LMS responded to its citation under NCLB's subgroup accountability policy, I found that the administration invested in the project of personalizing the gaps, which consumed the work of the data specialist and a significant amount of time at administrative meetings, school leadership team meetings, and whole-school staff meetings over the 2014-2015 academic year. At this time, formal teacher evaluation at LMS involved 1) classroom observations guided by a detailed rubric and 2) a value-added measure based on student growth on state standardized tests and on locally-selected assessments from a state-approved list. The value-added measure had only two years prior become a factor in teachers' overall performance ratings after a contentious negotiation between the teachers' union and the district, and it accounted for forty percent of teachers' evaluations. A teacher with a low overall rating was assigned an individualized professional development plan and was supervised more closely. These teacher evaluation policies contextualize the local efforts around teacher accountability initiated by LMS's administration.

Data Collection

Data collection included more than 120 days of participant observation spaced over one academic year (August to June). All students at LMS were considered participants in my observations unless they (or their guardians on their behalf) opted out. At the beginning of the year, I visited each classroom to introduce myself and to explain my study. I spoke in Spanish in the bilingual Spanish classrooms, and relied on teachers to translate in the Bangla bilingual classrooms. I distributed information forms in English, Bangla, and Spanish that included instructions on how students or their parents could opt out of the study. No students or parents contacted me to remove themselves from the study. Staff members had to provide consent for me to observe in their classrooms or offices.

I intentionally distributed my observations across official organizational units (e.g., content areas, grade levels) and followed connected series of events, such as consecutive meetings between the principal and the data specialist where they discussed plans for how to organize and analyze data. Particularly important for this paper were administrative meetings, school leadership meetings, meetings involving the data specialist, and faculty meetings where the project of personalizing the gaps was presented and implemented. LMS's administrative team, which consisted of a principal

(Richard) and three assistant principals (APs; Janice, Shaun, and John), met formally in administrative meetings and informally throughout the school week. The school leadership team included the administration and teacher leaders, educators at LMS who taught and also took on programmatic or teacher-support roles. Teacher leaders included content-area department leaders, special education program leaders, bilingual education program leaders, the data specialist, and the test coordinator. The administration relied on the school leadership team to monitor compliance with policies, to plan and oversee teacher professional development, and to help implement new initiatives.

I also conducted 73 semi-structured interviews with teachers, administrators, parents, and students, which included all school administrators and the majority of teachers. All adults interviewed provided their active consent, while students had to acquire their guardian's consent and offer their own assent to be interviewed. I conducted interviews after spending several months in the field, so that I could adapt the interview protocols to LMS's context and inquire about practices and policies that I had observed. Interviews with administrators and teachers addressed a range of topics that aimed to capture educators' understandings of student needs and how the school does and should support students through organizational structures, policies, and practices. After discovering in administrative meetings that teacher accountability was central to how leaders at LMS wanted to pursue educational equity, I added questions to the administrator and teacher interview protocols about perceptions of the teacher evaluation system and of the school's data practices. The interviews were typically one hour with teachers and administrators and were approximately half an hour with students and family members.

Ethnographic observations and interviews helped me to capture "data use from the perspective of practice, that is, the practice of those who policy makers hope will use data to make decisions about improving classroom instruction" (Spillane, 2012, p. 2). Policymakers and district administrators intend for data to be leveraged by schools for improvement in particular ways, but field-based research methods are needed to understand how the data produced are actually utilized in order to understand the range of effects.

Data Analysis

Data analysis began with open coding the observation and interview data and generating an extensive list of emergent themes with the assistance of the qualitative analysis program MAXQDA (Glaser & Strauss, 1967). An example of an emergent theme I identified is "my gaps," which was defined as educators using language that suggested they accepted responsibility for performance gaps between student subgroups. The emergent themes were added to a set of a priori themes that framed my initial inquiry, such as "subgroup accountability policy," to create a codebook that was subsequently used to recode the entire data set (Miles et al., 2014). With their illustrative excerpts of data attached, the codes were grouped in MAXQDA into topical "sets" that became the subject of analytical memos. Three of the sets particularly salient for this paper were "understandings and explanations for performance gaps," "teacher evaluation and accountability," and "data practices." The codes included in "teacher evaluation and accountability" were: "definitions of good teaching," "perceptions of teachers and teaching," "teacher contract and teacher union," "teacher evaluation," "teacher resistance," and "value-added measures." The analytical memos I wrote about each set helped me to articulate the various relationships among codes, and they served as the foundation for my findings once additional validity was established through the consideration of potentially disconfirming evidence and peer debriefing with colleagues who were concurrently engaged in ethnographic projects in schools (Creswell & Miller, 2000).

Findings

"It's the Teachers' Fault": Personalizing the Gaps

The administration's project of "personalizing the gaps"—holding individual teachers responsible for the performance gaps between student subgroups—coupled the commitment to educational equity with teacher accountability. The principal, Richard, proposed his idea of linking performance gaps to teachers during an administrative meeting with the three assistant principals (APs) in late August before the first official day of school: "We need to personalize it, individualize it with professional development, it's the teachers' fault." The APs built on the idea as the principal searched for a file on his computer. AP Janice suggested that they tie subgroup performance gaps directly to teachers' evaluations, "We need to have a discussion of how are you targeting these kids…it needs to be a part of the rating." By that point, the principal had the teacher evaluation rubric displayed on the monitor along one wall of his office and was scrolling through each component to see where they could link teacher ratings to the performance gaps between student subgroups.

Although the idea was generated in August, the project of personalizing the gaps involved several initiatives put to practice throughout the academic year. First, the administration proposed a theory of action that attributed performance gaps to teachers' differential engagement of student subgroups. Second, the administration used student performance data to attempt to measure individual teachers' value-added with each student subgroup. Third, the administration sought to shape teacher mindsets so that teachers would take ownership of the data set and accept their individual responsibility for performance gaps. And, finally, the administration discouraged alternative explanations for inequities in student outcomes that potentially attenuated the force of personalizing the gaps. These efforts to personalize the gaps, which I describe at length below, demonstrate how LMS's administrators assembled symbolic and technical tools in the policy environment, such as the intensive focus on teacher accountability and value-added models, to individualize with seeming precision the complex and collective factors that contribute to educational inequities, a process of simplification and individualization that is characteristic of neoliberalism (Ball, 2003; Piattoeva, 2021).

"Giving the Right Medicine to Everybody": Explaining Performance Gaps Through Differential Engagement

At a school leadership team meeting with administrators and teacher leaders early in the school year, Richard presented his 'theory of action'—a hypothesis that explains the connections between inputs and outcomes in a process—that teachers' differential engagement of students accounted for the variation in subgroup performance. With a focus on the relationship between inputs and outputs that characterizes a culture of performativity (Lyotard, 1984), Richard explained: "what a theory of action is, is 'if we do this, then this will happen,' it is something in school leadership practice right now…right now my theory of action is if you address engagement, we will address the achievement gap." According to the principal's theory of action, inequitable engagement explained performance gaps based on race, gender, EL, and special education status, and teachers' differentiated engagement of different student subgroups would serve as an instructional antidote to the disparities.

Richard's theory of action aligned with the school's annual instructional focus—one area of growth selected to drive professional development for the school year (a practice required by the district). 'Engagement' was chosen the previous Spring to be the focus of the 2014-2015 school year, and subsequently over the summer Richard resolved to address the performance gaps at LMS after

being influenced by research and the media. He folded the focus on performance gaps into the professional development focus on engagement: "we use engagement to hone in on every population where there is a gap…that's going to look different with each of those populations...this for this group, this for this group...giving the right medicine to everybody." Richard posited that teachers needed to differentiate—subgroup by subgroup—the ways in which they engaged students. By integrating the new commitment to performance gaps into the focus on engagement, Richard married the school's equity work into teacher professional development as well as teacher evaluation. The instructional focus not only served as the core of professional learning for the year, but also set a performance standard that all teachers must improve in the realm of engagement on the teacher evaluation rubric.

The attribution of performance gaps to the previously established instructional focus highlights that the explanation for performance gaps between student subgroups and the efforts to address them were not based on a root cause analysis, but were rather solutions-focused. A solutions-oriented approach is common in business consulting (e.g., Bain & Company, a global management consulting firm, refers to the approach as "answer-first"). Richard's reliance on management vocabulary and business practices reflected his own background and professional training (Spillane, 2012; Spillane et al., 2002). For a previous career, he received a MBA, and he participated in a school leadership program that framed principals as CEOs of their schools and that focused on using test score data to drive decision-making. This adoption of private market management strategies in public organizations is illustrative of neoliberal reforms in schools that have changed school leadership and the supervision of teachers (Anderson & Cohen, 2015; Ball, 2016).

"The Poor Man's Regression Analysis": School-Level Value-Added Modeling

The theory of action driving LMS' equity work developed from a concept into a targeted initiative to measure each teacher's impact on the progress of each student subgroup. Richard was excited about parsing student performance data by teacher and worked with his data specialist, who was also a teacher in the school, to measure individual teachers' value-added for student subgroups (female, male, Black, Latinx, Asian, White, general education students, students with disabilities, non-limited English proficient (LEP), LEP, and formerly LEP). In an early September meeting with the data specialist, Richard explained that he wanted her to present the growth in performance for each student subgroup by whole school and grade level, but "the ultimate is by class." I noted that the principal described the most important unit of disaggregation as 'classroom' with the data specialist, but when I had met with him privately a few days earlier, he had described it to me as finding "value-added from individual teachers—how did each teacher do with the subgroups." During this private discussion, he had attempted to enlist me to support the data specialist in creating the value-added data set, but I gently resisted and raised some concerns. I noted that his method of measuring the change in a teacher's class performance did not follow the same students from one year to the next. Richard was unfazed and responded that he was just trying to get at trends of how teachers performed with each subgroup. It was not "statistically sound, but it is conceptually sound," he reasoned, and it was "not about making value judgments," but rather finding strong teachers. Richard wished to downplay the ways in which this new performance measure would facilitate comparison and judgement of teacher productivity and effectiveness, as technologies of performativity do (Ball, 2003), and to emphasize the data set as a "means of incentive…and change" (Ball, 2003, p. 216).

The analysis of student subgroup growth using the 2013 and 2014 annual state test data was the principal's first attempt to measure individual teachers' value added by student subgroup. Yet,

Richard hoped that the data set would motivate immediate action and would therefore ideally not rely on the state exams given at the end of the year. In an effort to analyze student subgroup growth more quickly, the principal required teachers to administer mini-assessments around one concept from the Common Core State Standards that would take students about ten minutes to complete and could be quickly graded. In contrast to the state exams, the principal reasoned, this data set would inform teachers in a timely manner so that they could adjust their practice. These mini-assessments were introduced midway through the year, such that the value-added data set was drawn from assessments with vastly different formats in the Fall and Winter; the Fall assessment was a comprehensive unit test aligned to the classroom curriculum and the Winter assessment was a single task focused on one Common Core standard. (Although the principal aimed to collect another round of mini-assessment data, this never occurred.) Richard reasoned that the variations in test format and teacher grading practices were unimportant because they were at least consistent across all students, and he was interested in capturing comparative growth: "Progress gets evened out…if you are grading hard you are doing that to all of your kids…the other thing that gets evened out is gaps in subgroups…grading hard appl[ies] to all their subgroups." Focusing on student subgroup trends, according to Richard, mitigated any issues related to grading that may have undermined the data set's validity. If teachers were particularly hard or easy graders or if they were prejudiced in their grading against particular subgroups, the focus on growth would eliminate these biases, he reasoned. When I commented in a meeting with the principal and the data specialist midway through the year that research has shown value added modeling is complex and hard to make valid (e.g., AERA Council, 2015), the principal admitted the data's imperfection but was not dissuaded: "I am actually only looking for trends...it's the poor man's regression analysis...I am going to see which populations are moving more than other populations." Richard felt confident that the data would tell a clear story at the aggregate, subgroup level—that looking at patterns would protect against methods that may be statistically unsound.

The principal was excited by the unique potential of the data set to provide feedback to the teachers that they could not ignore and to improve teacher practice. He was hopeful that the data's story would resonate with teachers, motivate them, and lead to instructional growth. The data, he reasoned, could serve as the basis for identifying strengths among the teachers and peer-to-peer learning; if the data showed a teacher had a gap between ELs and non-ELs, they could find in the data set another teacher who did not have this gap, go to them for suggestions, and observe in their classroom. This approach to teacher learning insinuated that there were specific strategies that were more or less promising with particular student subgroups, and that these strategies could be distilled through observation or sharing. Richard circumvented the suggestion that all students in a subgroup respond similarly to the same pedagogical practices, which would reinforce racial, ethnic, and ableist stereotypes, and emphasized that the empirical value of the data set is that it objectively reveals the teachers, and thereby the practices, that work with certain student subgroups—a focus on 'what works' that is characteristic of neoliberal management (Ball, 2003, 2016). In an interview, he explained at length how the value-added data set would help teachers address diverse student needs disaggregated by official subgroup categories:

> It is for teachers to understand, why are you getting different outcomes with different groups, and it says something about the teacher…*The reason you want to break it down by these categories is because as much as possible, they indicate to you, somewhat discrete learning needs. Not so much that you can make pre-assumptions about what the learning needs of those kids are going to be as that you can look at outcomes and see trends and make retrospective analysis of what your strategies did for different populations.* [emphasis added] If your outcomes say your girls got different outcomes than your boys, then you did something that allowed girls to make one kind of progress and [boys] to make

> another kind of progress, or allowed students of a certain ethnicity to make one kind of progress and a different ethnicity to make a different kind of progress, or gender or disability status, or whatever it is. To me, it allows you to learn and understand that you are not meeting the needs of all students. *You can't really ultimately say, this is what you're supposed to do for all ESL students or this is what you're supposed to do for all girls, or this is what you're supposed to do for all Black kids, or whatever the case, but you can say, I got very disparate progress outcomes, let me look at what I did do and understand that and understand why.* [emphasis added]

Richard sought to avoid the problem of stereotyping by making the approach a largely behavioral one in which teachers who were more effective would be emulated without considering what linked their practice and their students' performance. Richard was confident that the value-added data set would ameliorate LMS's performance gaps if teachers accepted the 'objective' lessons it revealed.

The administration envisioned that the value-added data set would deliver incontestable evidence of the teachers' practices more effectively than conferencing and conversation, which were, to the principal's frustration, often characterized by teacher defensiveness and resistance. The administration wanted the teachers to personalize the gaps, but not to take them personally. Richard celebrated how the value-added data would allow the administrators to externalize and depersonalize their feedback to teachers, to make it appear matter-of-fact rather than judgmental. As he explained to me one winter day as the two of us sat in his office, "All I want to do is say to a teacher… 'you have a problem teaching boys, your Black kids, or your Hispanic students'…The data will create leverage." The principal understood the data as an unemotional foundation for initiating conversations about how a teacher's practice leads to performance gaps between student subgroups.

For teachers to utilize the data set as LMS's leadership envisioned, they needed a particular orientation to it, one that gave credence to the objectivity and validity of quantitative data. Teachers were expected to receive the data with gravity and take responsibility for the gaps the data revealed. In an effort to "[instill] the attitude and culture within which workers feel themselves accountable" (Ball, 2003, p. 219), the principal implored the teachers to "let the data speak" during a whole staff professional development session in February. In order to "let the data speak," teachers had to accept the story it told, as if it were objective rather than curated by the principal, impacted by variations in teacher grading and test format, and affected by school context and student characteristics. Richard exerted tight control over the organization, interpretation, and use of the data set in order to shape teachers' mindsets, exercising his institutional power to shape data practices (Coburn & Turner, 2011).

Let's "Not Have an Extended Conversation on Causes": Marginalization of Alternative Explanations for Performance Gaps

The project of personalizing the gaps relied on the principal's answer-first theory that teachers could be held responsible for LMS's performance gaps. When educators offered alternative explanations or sought to explore the root causes behind the school's performance gaps, their comments were quickly challenged, dismissed, or marginalized.

Beginning early in the year, suggestions to investigate the reasons for LMS's performance gaps were deflected by the administration and some instructional leaders. At a teacher leadership meeting in September, a teacher leader from the science department asked whether they would be exploring the root causes of LMS's disparities, but her question was quickly buried by warnings about teachers making excuses for low student performance and about the complexity of the problem:

Science Department Leader / Teacher:
Are we going to have a discussion about what are the reasons for the achievement gap—

Principal: Should there be?

Science Department Leader / Teacher:
…a discussion with teachers, with students…The solution starts with finding out why we have, why these large groups are [lagging]

Testing Coordinator / Teacher: [I'm] worried about having a discussion that devolves into…pointing fingers.

…

Principal: It's a loaded issue…It would require doing some reading…not creating excuses or distractions…[Engagement's] connection to the achievement gaps is natural, looking at [the problem] as the disengaged learner…[We need to] keep [working at making] those two related.

Richard redirected the conversation to teachers' engagement of students and then proceeded to recruit the science department leader to a committee that would be in charge of aligning the school's equity work with supporting disengaged subgroups.

Similarly, explanations for LMS's performance gaps that strayed from the principal's theory of action were marginalized. After the principal proposed his theory of action at a school leadership team meeting, Nina, an English language arts and bilingual education teacher who also served as a coordinator of the school's bilingual program, suggested that student grouping and class assignment could explain and address the performance gaps. Nina's proposal aligns with research that shows that tracking is the most significant factor contributing to within-school performance gaps (Gamoran, 2016). This organizational explanation would have placed more responsibility with school leadership, who determined student placement within the constraints set by district policy. Quick to respond, Richard briefly recognized that school organization and student assignment were detrimental to equity at LMS and then redirected the conversation back to his theory of action that implicated teachers.

LMS's leadership sought to keep the staff's attention on teacher practice and on solutions. As AP John explained in an administrative planning meeting, their approach was to "focus on how are we going to fix this, and not have an extended conversation on causes." Richard was determined to avoid conversations that explored the reasons behind performance gaps, as he explained in a school leadership team meeting in the Spring: "'Why' can be used as an embedded excuse…'why is this the data,' then you start looking for things that don't have to do with teacher impact." The administration and some instructional leaders worried in particular that an exploration of causes would create an opportunity for teachers to introduce student characteristics or factors from the school's social context that minimized their responsibility for student performance. In an interview at the end of the school year, AP Shaun reflected that the value-added data were "eye-opening" for

some teachers, while other teachers were not able to see their role in the disparate student outcomes and blamed students instead:

> I've had conversations with people who said that it's just [the students] don't want to work hard enough…it becomes a perception of like, "They just don't want to work hard enough, and if I made it, and I could do it, then they could do the same thing too." As opposed to understanding that your one-size-fits-all approach, your in-the-face approach, your talking-down approach, your not-building-a-relationship approach, probably has something to do with that.

The administration's tight control of the explanation for performance gaps and the marginalization of alternative approaches to equity work at LMS was embedded in the concern that the teachers would blame students, who were primarily students of color from low-income communities. To avoid this risk, the administration forcefully promoted the notion that performance gaps were the responsibility of teachers and materialized this idea by generating a measure of equitableness in each teacher's practice.

Teacher Responses

Rather than finding that teachers either complied with or resisted the administration's efforts to personalize the gaps, I found evidence of conditional compliance. When under direct administrative supervision, teachers fabricated (Ball, 2003) an image of consensual buy-in to personalizing the gaps. At other times, teachers expressed resistance or indifference to the administration's efforts to motivate them by a new quantitative measure of their quality.

My Gaps

The few instances in which I found teachers at LMS personalizing the gaps, they were under direct administrative supervision, such as during professional development sessions or one-on-one meetings with the principal. Teachers fabricated an image of professionalism by performing ownership of the value-added metrics. One day midway through the school year as the principal and the data specialist discussed how to display the value-added data to teachers, Richard noted that the Latinx students had made progress. Perhaps in reference to her Latina identity or her role as a math teacher in the Spanish bilingual program, the data specialist responded, "Because we rule!" Richard applauded her: "This is what we mean about personalizing!" The data specialist made the comment jokingly, but the principal praised how her use of language reflected ownership of the performance gaps at LMS. At the whole-staff professional development session in March where Richard taught teachers how he wanted them to use the value-added data set, he modeled similar language, guiding teachers as to how they can demonstrate their compliance and, according to the administration's definition, professionalism. Richard modeled for teachers how to talk about their own and other's data by referring to the gaps with possessive pronouns (e.g., 'Marie's EL gap') as opposed to other possibilities, such as naming gaps with the passive voice or referring to them by class (e.g., 'the gaps in 7F'). By providing the "vocabulary of performance" (Ball, 2003, p. 218) to discuss the data, Richard encouraged teachers to accept that there was an inextricable link between their own practice and inequitable student growth, even for teachers who may not consciously agree with the administration's theory of action or data practices. Following the principal's presentation, the teachers were given time to identify their gaps in the data distributed. One group of 8th grade math teachers discussed what they found, using language that performed ownership of the data. In the following excerpt from my fieldnotes, the teachers are engaged in discussion when the principal comes over to check-in with them:

> Math Teacher 1: Males and Hispanics, I'm terrible with.

Math Teacher 2: Mine's Asians…my female Black is 88%...My male Black is down seven.
Math Teacher 1: My male Black is still down, but my male Spanish—
Math Teacher 2: ESL is my lowest thing….
The principal walks over to the table and stands behind the teachers.
Principal: I was told I have to get in on this [conversation]…
Math Teacher 2: I'm not good with Asian.
Math Teacher 1: [Teacher 3] went up in the males.
Principal: You guys are killing it with the data conversation!

The teachers' use of possessive language (e.g.,"my male Black") in their self-assessments suggested that the data set encouraged them to understand themselves as strong instructors for some student subgroups and weak instructors for others and validated the theory that particular teachers could be traced to and held responsible for performance gaps. Richard praised the teachers for performing ownership of the performance gaps in their classrooms. Of note, the principal's high regard for these teachers was not related to how equitable or inequitable the data suggested their practice was, but instead related to their alignment with the neoliberal project of personalizing the gaps. For the teachers, taking up this discourse provided an opportunity to receive positive feedback from the administration and to perform excellence and professionalism, even if they may harbor criticisms of the new metric (Lewis & Holloway, 2019; Moore & Clark, 2016).

Despite the extensive time spent on the project of personalizing the gaps at administrative and school leadership team meetings throughout the school year, I only found a handful of instances in which teachers demonstrated ownership of the performance gaps in their classroom as evidenced by LMS's value-added data set. All cases were characterized by the close proximity to administrative supervision. By late Spring when I conducted many interviews, most teachers had forgotten about the value-added data set and were consumed with preparing their students for 'testing season,' a period of weeks when state and district standardized tests across content areas were administered. In none of my interviews were teachers able to remember what the data set said about the specific performance gaps in their classroom. A science teacher explained to me in an interview that the data were irrelevant to him, except when he was forced to utilize it: "I don't focus on data at all, I only look at it when we're asked to."

"The Problem with Data": Teacher Resistance

Several teachers at LMS expressed that components of the project to personalize the gaps were objectionable. Below, I describe how teachers questioned the theory of action attributing performance gaps to teachers' engagement of each student subgroup as well as the validity of LMS's value-added data set.

Resistance to the Administration's Theory of Action. For some teachers, the principal's theory of action attributing performance gaps to differential engagement did not align with their experiences in the classroom. When Richard asked the school leadership team to reflect on the performance gaps in their classrooms, Marie, an ELA department leader, noted that her male students were participating and behaviorally engaged, even though the state test data showed that they were lagging behind her female students. "It is not lining up with the data," she noted. Marie's comment contradicted the notion that performance gaps were due to student subgroups' disparate levels of engagement. She 'tried-on' the principal's theory of action, but it did not fit her experience.

Some staff members felt that the theory of action was too one-sided and placed an unreasonable amount of responsibility on the teachers' shoulders. Matt, a math teacher, pointed to

student absences and family responsibilities as two factors outside his control that affected the gaps in his classroom data:

> The problem with data is that there's so many outside factors…[In one class,] I think there's five or six Black students, and they all perform low, but [one is] never here, [another] misses 40 days out of the year, [a third] has missed a grade or been held back before and her parents are never home. She's taking care of her family…For me, closing the gap, it's got to be a balance between…your school setting and how the educational experience is being laid out for you, but it's also such a social, economic issue that there's no way you can solely close the gap.

Matt recognized the importance of his classroom practice, but hoped that the administration would promote a more balanced approach to closing the performance gaps that acknowledged out-of-school factors. Based on Matt's comment, the administration's worry that recognizing factors in the school context would lead teachers to avoid responsibility and to make excuses for unequal student outcomes may have been misguided and underestimated teachers' willingness to act upon a more nuanced perspective. A more balanced theory that emphasized the school's role despite external factors may have lent the administration's efforts more legitimacy.

On the other hand, some teachers did draw on cultural and racial stereotypes in their resistance to the administration's theory of action and deemphasized their role in the performance gaps at LMS. Echoing the myth of the model minority (Kim, 1999; Lee, 1996) to explain the academic performance of the school's Bengali students, a science teacher reasoned that teacher practice is less significant than parents' emphasis on education: "The way I see it, the only [group] that's always rising is the Bengali students, that has nothing to do with the classroom…in my opinion, what it has to do [with] is those kids come here from other countries and their parents put more pressure on education than anything else…that's what's going to show in the data." Accordingly, the Bengali EL students were often cast by teachers as outperforming the Latinx EL students at LMS, despite what the value-added data set showed and even though, according to AP Janice, the state test scores from the previous Spring did not reveal this disparity.

Resistance to the Administration's Data Practices. Teachers questioned the validity of LMS's value-added data set, expressing concerns with the composition and interpretation of the data. Some teachers argued that more data were necessary to make the data set valid. The equitableness of teachers' practices was being assessed by the difference between two data points that tested one Common Core concept. In an interview, an ELA instructional leader wished that the administration had collected data more frequently and extensively: "There needs to be more data points…I think we need to collect more data and look at some different aspects and not just one component." The limited number of data points and their narrow focus created uncertainty about the conclusions that could be drawn about teacher practice from LMS's value-added data set.

Other teachers took issue with the reliance on official subgroup categories to measure inequality in the classroom. Some teachers suggested that official subgroups were not differentiated enough to capture the disparities relevant to LMS's student body. At a school leadership team meeting in Winter where the principal sought to ensure that all the teacher leaders were clear about how to use the data to personalize the gaps, a teacher in the bilingual Bangla program suggested that the bilingual teachers should compare their students based on EL level: "[we should]…compare her beginners with my beginners and [her] intermediates [with my intermediates] because it wouldn't be fair to compare them to the rest of the class." Subgroup categories, in these cases, were perceived as too monolithic and inadequate to illuminate sub-subgroups of students that may be more significant for student performance, such as the difference between a beginner EL student who may have been

in the US for only a few days and an advanced EL student about to enter a general education classroom.

Teachers also questioned the meaningfulness of the gaps identified in their classrooms when subgroups were very small in size. Some teachers only had a few students belonging to a particular subgroup, but because the data measured comparative growth, the actual number of students was masked. One math teacher in the bilingual Bangla program gave no credence to the racial performance gaps in his data because he believed all of his students were Asian. He explained, "the [racial] comparison didn't make sense to [my data]," and he proceeded to focus on the gender gap, which seemed, based on his experience, more relevant to his classroom despite what the data 'said' was most important.

Discussion

Policymaking for Teacher Accountability at the School Level

Personalizing the gaps was a local incarnation of a national trend in education policy to hold teachers accountable for complex phenomena that research suggests is not solely within their capacity to address. The relationship between the unofficial value-added measure at LMS and the official value-added measure illustrates how policies are reappropriated in unintended ways through the process of implementation. In this case, the administration reapplied the logic of value-added accountability policies to further the local priority of addressing inequitable academic outcomes between student subgroups. The growing acceptance and use of value-added measures throughout the US lent an aura of legitimacy to the project of personalizing the gaps, demonstrating the symbolic importance of policies for authorizing and rationalizing local practices.

Personalizing the gaps as a school policy raises several concerns. For one, the administration integrated its problematic value-added measure into their understandings of teacher quality and effectiveness, which may have had implications for teachers' formal evaluation ratings and personnel decisions, such as recommendations for tenure. Moreover, personalizing the gaps may have shaped how teachers oriented to district and state teacher accountability policies and practices, including their understandings, responses to, and framings of the formal value-added measure, even though the two were independent. Teachers' doubts about locally forged value-added metrics may disengage them from all methods of teacher supervision, despite their varying validity and foundations in research. Unpredictable reappropriations of the 'raw materials' (Spillane, 2012) of teacher accountability, which require field-based inquiries to be uncovered, can have significant implications for teachers' experiences of supervision.

Willful Ignorance and Reproductive Teacher Resistance

The case of LMS suggests that the contradiction between policy interventions, such as increased teacher accountability, and the roots of educational inequity may be managed day-to-day through willful ignorance with an active downplaying of social context and a committed disinterest in the causes of inequality. The administration at LMS discouraged their staff from exploring the root causes behind LMS' performance gaps, and deflected theories of change that would place the responsibility for performance gaps outside of the teachers' locus of control. Not only was the socio-demographic context of the school dismissed by LMS's solutions-focused approach to equity work, but organizational critiques related to program design and student placement, which would place greater responsibility in the hands of administrators and the district, were marginalized by the reigning theory of action. Teachers' efforts to raise factors external to the school were framed as making excuses and shirking their responsibilities. Professionalism and being equity-minded were

defined in part at LMS as accepting responsibility for performance gaps that were not clearly tied to one's own practice and as refraining from independent, critical analysis of the roots of inequity. The significant efforts taken to "[instill] the attitude and culture within which workers feel themselves accountable" (Ball, 2003, p. 219) reflects how this neoliberal approach to teacher supervision worked to produce un-agentic teachers who cannot be trusted to invest in equity work that gives credence to the socio-economic context of schools. An alternative approach to equity work would require administrators with the capacity to lead educators in complex discussions about racial, economic, and social (in)justice and who cast teachers as professionals capable of concurrently understanding their role and the role of out-of-school factors in producing educational inequities.

On the other hand, the ways that teachers at LMS resisted the project of personalizing the gaps complicate calls to increase teacher autonomy to counter neoliberal reforms. In questioning the principal's theory of action, several teachers did not rely on deep professional knowledge or an established practice of inquiry, but rather drew upon cultural and racial stereotypes to propose alternative explanations for the performance gaps. Holloway (2021) emphasizes that "teachers' subjectivities, practices and knowledge are themselves based on deeply problematic logics and discourses" (p. 414) and for this reason, teacher behavior—as situated within the problematic policies, discourses, and institutions in which they have been socialized—must be open to critique. Devolving authority to teachers who reproduce racist beliefs and practices does not portend a better approach to equity work; indeed, earlier forms of teacher professionalism that gave teachers significant autonomy did not bring equitable educational outcomes (Weiner, 2020). Teachers' resistance also focused on their concerns with the value-added data set's validity, but teachers did not question the quantification of teacher quality in principle. Often teacher resistance to accountability policies retains the practice of defining the work of teaching through numbers (Piattoeva, 2021). A focus on the validity of LMS's value-added data set may have consumed teachers' energy for resistance, while the policies of teacher accountability that fostered such practices remained uncontested (Piattoeva, 2021). In short, teacher resistance must be differentiated (Anderson & Cohen, 2015; Holloway, 2021). Some forms of resistance to neoliberal policies and practices reproduce inequities, are focused on individual advancement, and insulate problematic practices, while others are collective and transformative (Anderson & Cohen, 2015).

Equity Work as Measuring Teacher Equitableness

In LMS, equity work was constructed by the administration as the elimination of performance gaps in test data through holding teachers increasingly accountable for the outcomes of student subgroups. Particular teachers may indeed contribute to educational inequities, yet the assumption that teachers' contributions to performance gaps within one school year can be quantified and addressed by a ready acceptance of these measures is indicative of how equity work has been shaped by a culture of performativity that continually produces new simplistic measures to assess teacher worth with little attention to developing practices. The consequence is not only "the translation of complex social processes" such as the production of educational inequity "into simple figures or categories of judgement" (Ball, 2003, p. 217), but it is also a new categorization of teachers according to their equitableness. Teachers resistant to this new measure were framed as resistant to equity work.

The case of LMS shows how equity work was folded into neoliberal educational reforms. Another example of neoliberal equity work that preceded the emphasis on teacher accountability is educational triage, in which educators focus on students immediately below a given performance threshold and direct resources away from those far below or above the threshold in order to demonstrate growth (Booher-Jennings, 2005; Gillborn & Youdell, 2000). Introduced with NCLB,

the collection of performance data by student subgroups was an important achievement for civil rights advocates to ensure that equity would be a priority in the accountability movement, however the production of data for monitoring equity should not be confused with reforms to policy and practice that mitigate disparities illuminated in data. Further research is needed to explore how equity work is shaped by the policy environment and by the neoliberal policy regime specifically.

The promise of neoliberal efforts to engage teachers in equity work is bleak if it is singularly focused on the production of new performative measures and fails to engage teachers in the critical examination of their classroom practices and beliefs that may further educational inequities. As Ball explains, in a culture of performativity extensive time spent on monitoring and measuring takes the place of actual reform or changes to practice: "we are required to spend increasing amounts of our time in making ourselves accountable, reporting on what we do, rather than doing it" (Ball, 2016, p. 1053; see also Lyotard, 1984, on the law of contradiction). The use of data to measure teacher effectiveness without direct links to implications for practice, may demoralize teachers and undermine improvement (Jennings & Pallas, 2016). The cost of reducing equity work to the production and acceptance of a new teacher quality measure was the alienation of teachers from meaningful efforts to address educational inequities.

References

AERA Council. (2015). AERA Statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher, 44*(8), 448-452. https://doi.org/10.3102/0013189X15618385

Amrein-Beardsley, A. (2014). *Rethinking value-added models in education*. Routledge. https://doi.org/10.4324/9780203409909

Amrein-Beardsley, A., & Holloway, J. (2017). Value-added models (VAMs) for teacher evaluation and accountability: Commonsense assumptions. *Educational Policy, 33*(3), 516-542. https://doi.org/10.1177/0895904817719519

Anderson, G., & Cohen, M.I. (2015). Redesigning the identities of teachers and leaders: A framework for studying new professionalism and educator resistance. *Education Policy Analysis Archives, 23*(85). https://doi.org/10.14507/epaa.v23.2086

Ball, S. J. (2000). Performativities and fabrications in the education economy: Towards the performative society? *Australian Educational Researcher, 27*(2), 1-23. https://doi.org/10.1007/BF03219719

Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy, 18*(2), 215–228. https://doi.org/10.1080/0268093022000043065

Ball, S. J. (2016). Neoliberal education? Confronting the slouching beast. *Policy Futures in Education, 14*(8), 1046–1059. https://doi.org/10.1177/1478210316664259

Berliner, D. C. (2006). Our impoverished view of educational research. *Teachers College Record, 108*(6), 949-995.

Berliner, D. C. (2009). *Poverty and potential: Out-of-school factors and school success*. Education and the Public Interest Center & Education Policy Research Unit.

Berliner, D. C. (2013). Effects of inequality and poverty vs. teachers and schooling on America's youth. *Teachers College Record, 115*(12), 1-26. https://doi.org/10.1177/016146811311501203

Blazer, D., Litke, E., & Barmore, J. (2016). What does it mean to be ranked a "high" or "low" value-added teacher? Observing differences in instructional quality across districts. *American Educational Research Journal, 53*(2), 324-359. https://doi.org/10.3102/0002831216630407

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal, 42*(2), 231–268. https://doi.org/10.3102/00028312042002231

Braun, H. (2015). The value in value added depends on the ecology. *Educational Researcher, 44*(2), 127–131. https://doi.org/10.3102/0013189X15576341

Card, D., & Rothstein, J. (2006). *Racial segregation and the Black-White test score gap.* (NBER Working Paper No. 12078). National Bureau of Economic Research. https://doi.org/10.3386/w12078

Chetty, R., Hendren, N., & Katz, L. F. (2015). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. (NBER Working Paper No. 21 156). National Bureau of Economic Research. https://doi.org/10.3386/w21156

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research and Perspectives, 9*(4), 173–206. https://doi.org/10.1080/15366367.2011.626729

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., & Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. U.S. Government Printing Office.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124-130. https://doi.org/10.1207/s15430421tip3903_2

Downey, D. B., & Condron, D. J. (2016). Fifty years since the Coleman Report: Rethinking the relationship between schools and inequality. *Sociology of Education, 89*(3), 207-220. https://doi.org/10.1177/0038040716651676

Dobbie, W., & Fryer, Jr., R. G. (2009). *Are high-quality schools enough to close the achievement gap? Evidence from a social experiment in Harlem*. (NBER Working Paper No. 15473). National Bureau of Economic Research. https://doi.org/10.3386/w15473

Echávarri, J., & Peraza, C. (2017). Modernizing schools in Mexico: The rise of teacher assessment and school-based management policies. *Education Policy Analysis Archives*, *25*(90). https://doi.org/10.14507/epaa.25.2771

Egalite, A. J. (2016). How family background influences student achievement: Can schools narrow the gap? *Education Next*, *16*(2), 70-78.

Flores, M. A., & Derrington, M. L. (2017). School principals' views of teacher evaluation policy: Lessons learned from two empirical studies. *International Journal of Leadership in Education: Theory and Practice*, *20*(4), 416–431. https://doi.org/10.1080/13603124.2015.1094144

Franco, M. S., & Seidel, K. (2014). Evidence for the need to more closely examine school effects in value-added modeling and related accountability policies. *Education and Urban Society, 46*(1), 30-58. https://doi.org/10.1177/0013124511432306

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Gamoran, A. (2016). Gamoran comment on Downey and Condron. *Sociology of Education, 89*(3), 231-233. https://doi.org/10.1177/0038040716651931

Gillborn, D., & Youdell, D. (2000). *Rationing education: Policy, practice, reform, and equity*. Open University Press.

Goldhaber, D. (2016). In schools, teacher quality matters most. *Education Next, 16*(2), 56-62.

Hall, D., & McGinity, R. (2015). Conceptualizing teacher professional identity in neoliberal times: Resistance, compliance and reform. *Education Policy Analysis Archives, 23*(88). https://doi.org/10.14507/epaa.v23.2092

Hanushek, E. A. (2016). What matters most for student achievement. *Education Next, 16*(2), 19–26.

Hattie, J., & Anderman, E. M. (2020). *Visible learning: Guide to student achievement*. Routledge. https://doi.org/10.4324/9781351257848

Holloway, J. (2021). Teachers and teaching: (Re)thinking professionalism, subjectivity and critical inquiry. *Critical Studies in Education, 62*(4), 411-421. https://doi.org/10.1080/17508487.2021.1966065

Holloway, J., & Brass, J. (2017). Making accountable teachers: The terrors and pleasures of performativity. *Journal of Education Policy, 33*(3), 361–382. https://doi.org/10.1080/02680939.2017.1372636

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher, 33*(8), 16-30. https://doi.org/10.3102/0013189X033008016

Huang, H., & Sebastian, J. (2015). The role of schools in bridging within-school achievement gaps based on socioeconomic status: A cross-national comparative study. *Compare: A Journal of Comparative and International Education, 45*(4), 501-525. https://doi.org/10.1080/03057925.2014.905103

Jencks, C. (2016). Jencks comment on Downey and Condron. *Sociology of Education, 89*(3) 221-222. https://doi.org/10.1177/0038040716651677

Jennings, J. L., Deming, D., Jencks, C., Lopuch, M., & Schueler, B. E. (2015). Do differences in school quality matter more than we thought? New evidence on educational opportunity in the twenty-first century. *Sociology of Education, 88*(1), 56-82. https://doi.org/10.1177/0038040714562006

Jennings, J. L., & Pallas, A. M. (2016). How does value-added data affect teachers? *Educational Leadership, 73*(8).

Kalogrides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher, 42*(6), 304-316. https://doi.org/10.3102/0013189X13495087

Kim, C. J. (1999). The racial triangulation of Asian Americans. *Politics and Society, 27*(1), 105-138. https://doi.org/10.1177/0032329299027001005

Lacireno-Paquet, N., Bocala, C., & Bailey, J. (2016). *Relationship between school professional climate and teachers' satisfaction with the evaluation process.* Regional Educational Laboratory Northeast & Islands.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in US schools. *Educational Researcher, 35*(7), 3–12. https://doi.org/10.3102/0013189X035007003

La Londe, P. G. (2017). A failed marriage between standardization and incentivism: Divergent perspectives on the aims of performance-based compensation in Shanghai, China. *Education Policy Analysis Archives, 25*(88). https://doi.org/10.14507/epaa.25.2891

LeChasseur, K., Donaldson, M., Fernandez, E., & Femc-Bagwell, M. (2018). Brokering, buffering, and the rationalities of principal work. *Journal of Educational Administration, 56*(3), 262-276. https://doi.org/10.1108/JEA-10-2016-0129

Lee, S. J. (1996). *Unraveling the "model minority" stereotype: Listening to Asian American youth.* Teachers College Press.

Lewis, S., & Holloway, J. (2019). Datafying the teaching 'profession': Remaking the professional teacher in the image of data. *Cambridge Journal of Education, 49*(1), 35-51. https://doi.org/10.1080/0305764X.2018.1441373

Lyotard, J.-F. (1984). *The postmodern condition: A report on knowledge.* Manchester University Press.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83*(2), 340-363. https://doi.org/10.1086/226550
Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.
Moore, A., & Clarke, M. (2016). 'Cruel optimism': Teacher attachment to professionalism in an era of performativity. *Journal of Education Policy, 31*(5), 666-677. https://doi.org/10.1080/02680939.2016.1160293
Murphy, J., Hallinger, P., & Heck, R. H. (2013). Leading via teacher evaluation: The case of the missing clothes? *Educational Researcher, 42*(6), 349–354. https://doi.org/10.3102/0013189X13499625
Owens, A., Reardon, S.F., & Jencks, C. (2016). Income segregation between schools and school districts. *American Educational Research Journal, 53*(4), 1159-1197. https://doi.org/10.3102/0002831216652722
Piattoeva, N. (2021). Numbers and their contexts: How quantified actors narrate numbers and decontextualization. *Educational Assessment, Evaluation and Accountability, 33*, 511–533. https://doi.org/10.1007/s11092-021-09363-x
Palardy, G. J. (2010). The multilevel crossed random effects growth model for estimating teacher and school effects: Issues and extensions. *Educational and Psychological Measurement, 70*(3), 401-419. https://doi.org/10.1177/0013164409355693
Palardy, G. J. (2015). Classroom-based inequalities and achievement gaps in first grade: The role of classroom context and access to qualified and effective teachers. *Teachers College Record, 117*(2), 1-48. https://doi.org/10.1177/016146811511700203
Raudenbush, S. W. (2015). Value added: A case study in the mismatch between education research and policy. *Educational Researcher, 44*(2), 138–141. https://doi.org/10.3102/0013189X15575345
Rothstein, R. (2004). *Class and schools*. Economic Policy Institute.
Scott, J. (2011). Market-driven education reform and the racial politics of advocacy. *Peabody Journal of Education, 86*(5), 580-599. https://doi.org/10.1080/0161956X.2011.616445
Shifrer, D. (2020). Contextualizing educational disparities and the evaluation of teacher quality. *Social Problems,* 1-46. https://doi.org/10.1093/socpro/spaa044
Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education, 118*(2), 1-29. https://doi.org/10.1086/663283
Spillane, J. P., Diamond, J. B., Burch, P., Hallett, T., Jita, L., & Zoltners, A. (2002). Managing in the middle: School leaders and the enactment of accountability policy. *Educational Policy, 16*(5), 731–762. https://doi.org/10.1177/089590402237311
Straubhaar, R. (2017). The 'power' of value-added thinking: Exploring the implementation of high-stakes teacher accountability policies in Rio De Janeiro. *Education Policy Analysis Archives*, *25*(91). https://doi.org/10.14507/epaa.25.3034
Superfine, B. M. (2013). *Equality in education law and policy: 1954–2010*. Cambridge University Press. https://doi.org/10.1017/CBO9781139061797
Tatto, M. T. (2006). Education reform and the global regulation of teachers' education, development and work: A cross-cultural analysis. *International Journal of Educational Research, 45*(4–5), 231–241. https://doi.org/10.1016/j.ijer.2007.02.003
Tuytens, M., & Devos, G. (2018). Teacher evaluation policy as perceived by school principals: The case of Flanders (Belgium). *Teachers and Teaching, 24*(3), 209-222. https://doi.org/10.1080/13540602.2017.1397508

Verger, A., & Curran, M. (2014). New Public Management as a global education policy: Its adoption and re-contextualization in a southern European setting." *Critical Studies in Education, 55*(3), 253-271. https://doi.org/10.1080/17508487.2014.913531

Weiner, J. M. (2020). From new to nuanced: (Re)considering educator professionalism and its impacts. *Journal of Educational Change, 21*, 443–454. https://doi.org/10.1007/s10833-020-09371-6

About the Author

Rachel Garver
Montclair State University
mailto:garverr@montclair.edu
https://orcid.org/0000-0001-7504-8103
Rachel Garver is an assistant professor in the Department of Educational Leadership at Montclair State University. Her research examines the implementation of equity-oriented policies by teachers and administrators, school discipline and safety, and the preparation of educational leaders committed to racial and social justice.

About the Editors

Meghan Stacey
UNSW Sydney
m.stacey@unsw.edu.au
https:// orcid.org/0000-0003-2192-9030
Meghan Stacey is senior lecturer in the UNSW School of Education, researching in the fields of the sociology of education and education policy. Taking a particular interest in teachers, her research considers how teachers' work is framed by policy, as well as the effects of such policy for those who work with, within and against it.

Mihajla Gavin
University of Technology Sydney
mihajla.gavin@uts.edu.au
https:// orcid.org/0000-0001-6796-5198
Mihajla Gavin is lecturer at UTS Business School. Her PhD, completed in 2019, examined how teacher trade unions have responded to neoliberal education reform. Her current research focuses on the restructuring of teachers' work and conditions of work, worker voice, and women and employment relations.

Jessica Gerrard
University of Melbourne
jessica.gerrard@unimelb.edu.au
https:// orcid.org/0000-0001-9011-6055
Jessica Gerrard is associate professor at the Melbourne Graduate School of Education. Jessica researches the changing formations, and lived experiences, of social inequalities in relation to education, activism, work and unemployment. She works across the disciplines of sociology, history and policy studies with an interest in critical methodologies and theories.

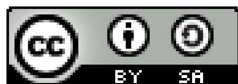
Anna Hogan
Queensland University of Technology
ar.hogan@qut.edu.au
https:// orcid.org/0000-0003-1934-2548
Anna Hogan is senior research fellow in the School of Teacher Education and Leadership at the Queensland University of Technology. Her research focuses on education privatisation and commercialisation. She currently works on a number of research projects, including investigating philanthropy in Australian public schooling, the privatisation of global school provision, and the intensification of teachers' work.

Jessica Holloway
Australian Catholic University
jessica.holloway@acu.edu.au
https:// orcid.org/0000-0001-9267-3197
Jessica Holloway is senior research fellow and ARC DECRA Fellow at the Australian Catholic University. Her research draws on political theory and policy sociology to investigate: (1) how metrics, data and digital tools produce new conditions, practices and subjectivities, especially as this relates to teachers and schools, and (2) how teachers and schools are positioned to respond to the evolving and emerging needs of their communities.

SPECIAL ISSUE
Teachers and Educational Policy: Markets, Populism, and Im/Possibilities for Resistance

education policy analysis archives

Volume 30 Number 101 July 19, 2022 ISSN 1068-2341



Readers are free to copy, display, distribute, and adapt this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, the changes are identified, and the same license applies to the derivative work. More details of this Creative Commons license are available at https://creativecommons.org/licenses/by-sa/4.0/. **EPAA** is published by the Mary Lou Fulton Teachers College at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), Directory of Open Access Journals, EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A1 (Brazil), SCImago Journal Rank, SCOPUS, SOCOLAR (China).

About the Editorial Team: https://epaa.asu.edu/ojs/index.php/epaa/about/editorialTeam

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join EPAA's Facebook community at https://www.facebook.com/EPAAAAPE and **Twitter feed** @epaa_aape.