



OPEN ACCESS

EDITED BY

Margaret Terry Orr,
Fordham University, United States

REVIEWED BY

Crystal Chambers,
East Carolina University, United States
Angel Yee-Lam Li,
Carnegie Foundation for the Advancement
of Teaching, United States

*CORRESPONDENCE

Sola Takahashi
✉ stakaha@wested.org
Carlos Sandoval
✉ csandov2@wested.org

RECEIVED 02 June 2024

ACCEPTED 10 September 2024

PUBLISHED 03 January 2025

CITATION

Takahashi S, Sandoval C, Jackson B,
Cunningham J and Taylor C (2025) Practical
measurement for equity and justice.
Front. Educ. 9:1442505.
doi: 10.3389/educ.2024.1442505

COPYRIGHT

© 2025 Takahashi, Sandoval, Jackson,
Cunningham and Taylor. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Practical measurement for equity and justice

Sola Takahashi^{1*}, Carlos Sandoval^{1*}, Brent Jackson¹,
Jahneille Cunningham¹ and Curtis Taylor²

¹WestEd, San Francisco, CA, United States, ²High Tech High Graduate School of Education, San Diego, CA, United States

What does it mean and how does it look for practical measurement - a cornerstone of improvement science methods - to support equity and justice in schools? Practical measures are not, in themselves, a “silver bullet” that will make school systems more equitable and just, but we propose that they may support ways of reflecting on and understanding the current system that lead to shifts in organizational routines. This author proposes 5 measurement routines that may position practical measurement as a resource for more equitable and just organizational processes. These are: (1) Attending to and reflect on day-to-day practices and processes for equity and justice; (2) Identifying instructional moves teachers can test to advance equity and justice; (3) Positioning students to take ownership over the learning environment; (4) Developing interpretations of how identity shapes/ is shaped the learning environment; (5) Surfacing and building on minoritized students’ strengths. These five measurement routines are examined in the context of a practical measurement development effort between WestEd and the CARE Network at the High Tech High Graduate School of Education. This collaboration led to the development of a practical measure focused on the experience of status hierarchies in math classrooms. The authors consider teachers’ uptake and use of these tools relative to the five routines of measurement for equity and justice.

KEYWORDS

practical measurement, equity & justice, improvement science, math education, organizational routines

Introduction

How might measurement and data catalyze and inform efforts to advance educational equity and justice in K-12 school systems in the United States? The standards-and-accountability movement that surged in the 1990’s and 2000’s in the U.S., codified in the No Child Left Behind Act of 2002, prompted a push for educators to engage in “data-driven decision-making” that was theorized to lead to more equitable schooling for students from historically underserved communities (Datnow et al., 2017). The data-use practices emerging from this push still typify educational data use in the U.S. today: data about educational outcomes - often standardized test results - are disaggregated by student demographics characteristics - often race - to highlight and then spur actions and interventions to reduce race- and class- based inequities.

Yet in practice, these equity aims of “data-driven decision making” were often unrealized (Datnow et al., 2017). Instead, a focus on racial “achievement gaps” has reinforced existing power differentials and further reified deficit and pathologizing beliefs

about student from Black, Latine, and Indigenous communities who are depicted at the lower end of a racial “achievement gap” (Bertrand and Marsh, 2021; Garner et al., 2017). An exclusive focus on standardized test results combined with the “achievement gap” narrative has been the context in which educators interpreted the data as evidence of what is wrong or lacking in students of color, their families, and communities. Rather than interrogating the schooling practices and critically reflecting on what is happening inside of the educational system in their data sensemaking conversations, educators have all too often located lower performance as a reflection of innate inabilities or assumed family and community dysfunction (Bertrand and Marsh, 2021). The focus on test scores and the “achievement gap” narrative has precluded educators from recognizing and celebrating the many assets brought by students of color, their families and their communities (Cunningham, 2019). When standardized test results are taken up to inform changes in practice, they can prompt changes that are remedial, focus on just the students who are thought to be able to score above a certain threshold (“bubble kids”), and prompt didactic instruction rather than the type of interactive instruction that provides opportunities for higher order thinking, especially for students of color and students from low-income backgrounds (Diamond, 2007).

In this article, we explore a different model of utilizing data in school systems to advance equity and justice. We will use the phrase “equity and justice” to convey the goals that organize the work described in this article - a delineation of how educational systems should treat students from historically marginalized groups, particularly Black and Brown students, but also inclusive of other non-dominant student groups. While the elevation of the goal of “equity” in education reform over the past few decades has led to explicit and important attention to non-dominant student groups, this term has too often come to convey a narrow focus on equalizing specific student outcomes (Bensimon, 2018; Gutiérrez, 2018). While we recognize the impact of illuminating inequitable outcomes on the work of education reformers, “equity” in this usage is not sufficient. We also center the concept of “justice” - an understanding of the presence of whiteness in the structures and routines of an institution (Bensimon, 2018) that undercuts the whole humanity of students from non-dominant groups (Gutiérrez, 2018). “Justice” is about the freedom to live full lives with dignity (Nussbaum, 2000). By calling out equity and justice, we intend to reach beyond a conception of closing “gaps” in outcomes to focusing on how educational institutions can better “rehumanize” students (Gutiérrez, 2018) - to give legitimacy to students’ full participation in school and honor the assets with which they enter the schooling space - and to focus on their “comfort, dignity, and agency” (Sandoval & Neri, same journal issue) in their daily schooling experiences.

We draw on the use of measurement and data from improvement science that focuses on system (re)design as the avenue toward improved outcomes. Improvement science - also known as “quality improvement” or more generally “continuous improvement” in education - is an approach to improving organizations that engages various members of that organization to enact changes in their work toward common goals, taking an inquiry stance toward change-making (Bryk et al., 2011; Grunow et al., 2024). The approach includes organizational members learning about the system they aim to improve as a set of processes

that unfold in the day-to-day of schooling, articulating their theories of improvement, and critically reflecting on the changes they make, while utilizing data to inform their learning about how to achieve their aims.

Practical measurement plays a critical role in an improvement science approach, as a way for improvers to learn about how key organizational processes are working, as a way to check improvers’ hypotheses of how to achieve the aim, and as a way to get feedback about if changes that are being made appear to be heading in the right direction. Practical measurement is defined as “the deliberate and routine gathering, analysis, and interpretation of information with the distinct purpose of enhancing the learning of system actors as they test changes and improve processes that are at the heart of their work” (Takahashi et al., 2022). Practical measures illuminate specific processes, norms, or structures that are theorized to be key areas of work toward a common goal - they are signals of practice. They also require minimal time and effort for educators to collect, and/or the collection is embedded in daily flows of work, to enable regular instances of collection and analysis. These measures provide data as the improvement work unfolds - not after the fact.

In the context of an equity-focused improvement effort, practical measures can serve the purpose of illuminating and focusing attention on the key organizational routines where equity is advanced - or inequity is produced and perpetuated - in the day-to-day of schooling. Diamond and Gomez (2023), in critically examining how improvement science tenets apply to tackling the pernicious marginalization experienced by Black students in U.S. schools, argue that “Anti-Black racism and white supremacy are perpetuated through their deep embeddedness in organizational processes” (p.2).

While intentional actions that disadvantage, harm, and dehumanize Black students perpetuate white supremacy and anti-Black racism, taken-for-granted organizational routines...are also crucial to creating hostile, anti-Black educational environments. ...[W]e argue that organizational routines are foundational mechanisms that generate and reproduce white supremacy and anti-Black racism but can also be used to disrupt these forces. (p.2)

As an example, in their deep analysis of the perpetuation of racial inequity in one suburban high school, Lewis and Diamond (2015) describe how the performed processes of student behavioral discipline and the placement of students in academic tracks unfold in ways that marginalize Black and Brown students and privilege white, middle class students. They demonstrate that educational inequities are not simply the result of “raving racists” (p. 179) in the school system, but they are rather woven into mundane and normalized organizational routines.

Because racial (and gendered, and classed, and other) marginalization occurs in the day-to-day processes of schooling, practical measurement has the potential to call attention to where this marginalization occurs, illuminate what is happening in practice, and provide the opportunity for critical reflection and informed next steps and actions where inequities are enacted (Takahashi et al., 2022). Practical measurement could play a role in shifting the conversation about racial inequity in schools from the sphere of educational outcomes (e.g., standardized test results) and

where students are at the end of their educational experiences, to a focus on educational routines and the lived experiences of students of color day-in and day-out. The latter focus is more proximate to organizational practices and routines - more clearly in the realm of responsibilities of educators in the system, and therefore less likely (although not impossible) to be attributed to students' innate abilities and dispositions. Measures operationalize what we hope to (or hope not to) see and hear in educational processes - they operationalize what quality looks like, and practical measures can point to the kinds of processes that are affirming, humanizing, and anti-racist. [Cunningham et al. \(2024\)](#) illustrate the potential of measures to do equity work in this way by highlighting the Equity Quantified in Participation (EQUIP) tool, a classroom observation tool focused on the equity of student participation in mathematics classroom. Protocols used as part of this tool focused on combating deficit orientations of students by shifting the onus to improve learning environments onto teachers rather than on students. In this vein, the potential for practical measures to advance equity in education mirror the possibilities that improvement science more broadly presents for advancing equity. [Hinnant-Crawford \(2020\)](#), for instance, provides insight into this potential at all stages of the improvement process; in particular, she illustrates the potential of practical measurement to advance equity and justice by surfacing how one might apply it to the work of culturally relevant and sustaining pedagogy.

While we recognize that there are affordances and possibilities, there is nothing inherent in practical measurement that would cause it to be a tool for advancing educational equity and justice. In fact, just as the application of improvement science methods do not, on their own, disrupt the pernicious practices of educational marginalization ([Diamond and Gomez, 2023](#); [Valdez et al., 2020](#)), so too can practical measurement be used to reinforce the *status quo* of systemic inequities, if they are not intentionally used to disrupt inequitable and unjust processes. They can be used to strengthen processes that are at their core reproducing the very inequities and injustices we aim to combat. In this paper, we propose a set of features of practical measurement that is more conducive to advancing equity and justice, not as a set of static criteria that inhere in measurement tools, but as an emergent set of "measurement routines" that can, in various combination, prompt the kinds of uptake and use that connects to critical reflection and informed changes toward more equitable practice. We discuss the idea of "measurement routines" further in the next section.

In this paper, we describe the theoretical underpinnings of our approach to practical measurement as a tool for changed practice. We draw on theoretical work that seeks to complicate the relationship between tools and action, viewing the two as inseparable in practice, in order to motivate an attention to routines. We then discuss a proposed set of five aspirational measurement routines for centering equity and justice. Our exploratory example in this paper is a partnership between WestEd and High Tech High. A team at WestEd embarked on a project with members of the CARE Network at the High Tech High Graduate School of Education to co-develop a measurement tool that addresses one of the challenging equity issues in math classrooms - the issue of status hierarchies among students. We will discuss the measurement development process, the learnings we have had from use in math classrooms, and our reflections on practical measurement to advance equity.

Technology, objects, and patterns of action

At the center of our concern are school organizational routines that perpetuate social inequities and injustices by their repeated, taken-for-granted daily enactment ([Diamond and Gomez, 2023](#); [Lewis and Diamond, 2015](#)). Organizational routines "are generative systems that produce repetitive, recognizable patterns of interdependent action carried out by multiple participants," and they are the foundation for how organizations carry out work ([Pentland and Feldman, 2008](#), p.236). These routines are not rigid, static, or documented - they are alive and mutable in their daily enactment. But how do we change these routines, when they reflect white-normative and anti-Black racial ideologies, to be more equitable and just? We see a role for practical measurement in these kinds of change efforts, though we are mindful of the "folly of designing artifacts, while hoping for patterns of action" ([Pentland and Feldman, 2008](#), p.235). Decades of research has made visible the ways in which tools come to shape, constrain, enable, and mediate actions and interactions between and among actors (e.g., [Engeström, 1999](#); [Latour, 2007](#)). Tools do not, by their mere presence, change action - and no practical measure is a "silver bullet" that will lead to more equitable and just schooling. To that end, we focus this paper on the kinds of actions and insights that move schools closer to equity and justice, and how we envision practical measurement engendering those actions and insights.

To better understand how practical measures can surface inequities and injustices in classroom processes, we conceptualize practical measures as tools that can exert agency over and shape (inter)actions. Rather than viewing these measures and their resulting data as static, disembodied, and separate from the actors who use and engage them, we instead position them as parts of assemblages of people, objects, artifacts, and technology that come to be performed in the work of schools and schooling, what [Orlikowski and Scott \(2008\)](#) call sociomateriality. In this paper, we are concerned less with what might be colloquially called the "design features" inherent to measurement tools, instead conceptualizing the tools in terms of how they are taken up and used. We focus on the kinds of organizational routines ([Feldman, 2000](#)) that we believe measures should support and play a role in if they were to focus on advancing equity and justice. To better conceptualize the role of objects and artifacts within the production of organizational routines, we turn to work on conceptualizing resources and, importantly, *resourcing*. We end with a discussion of what we gain from viewing practical measures from these lenses.

Resourcing

To better position practical measures as objects that come to be performed to produce equity- and justice-focused patterns of action, we draw on [Feldman \(2004\)](#) conceptualization of *resourcing*. Feldman argues that dominant notions of "resources" view them as static "things" that exist independent of context. Instead, resources are created through actions that, in turn, enable people to enact rules and schemas that create more resources. In this frame, resources only become resources when they are created and drawn

on through action; structured and structuring patterns of actions—or organizational routines (Feldman, 2000)—and resources do not exist separate from one another but instead are mutually constitutive and generative. In this frame, “resources” are not resources until they have become *resourced* to do something.

Pentland and Feldman (2008) highlight the challenges with treating resources and routines as separate. In their illustrative example of a failed effort to implement a new software program within an organization, they make visible the ways in which the software program, intended to be a resource that changed day-to-day organizational routines of labor and human resource management, came to have much of its intended functionality ignored. Pentland and Feldman’s analysis of their example revealed that much of the intended functions and actions were eschewed in favor of upholding and maintaining existing organizational routines. Their example highlights the distinction between what they call “live” and “dead” routines. Where live routines are people-driven generative systems where actions are dynamic and contingent upon the circumstances, dead routines are strictly made of artifacts; dead routines are “rigid, mindless, and can be explicitly stored” (p. 236). The authors articulate an example of a dead routine as a checklist “developed by people who do not enact the routine” that is then “largely if not totally ignored by those who do enact the routine” (p. 240). These routines are mindless and easily changed but require little to no thought for actors, generate no new learning, and rarely result in changes to existing routines. While potentially useful for limited circumstances, dead routines are often not the transformative changes that leaders within organizations seek. Live routines, on the other hand, generate actions and performances that are constantly changing and shifting depending on the circumstances.

We use the work on resources and resourcing to focus our attention away from designing for practical measures should look like as artifacts independent of context, and toward designing for aspirational patterns of actions and the measurement tools that can enable them. In particular, we aim to articulate a working, aspirational set of measurement routines that are intended to enact equity and justice in schools, focusing specifically on classroom processes and practices, and the ways that practical measurement tools ought to support these measurement routines.

Measurement routines for equity and justice

We argue that, in order for practical measurement tools to be recruited for surfacing and redressing inequities and injustices in classroom practices and processes, they ought to support a set of measurement routines among educators in schools. We articulate five measurement routines that practical measures ought to enable teams of teachers and instructional leaders to engage in: a) attending to and reflecting on the day-to-day practices and processes that constrain and/or enable equity and justice in classrooms; b) identifying and generating instructional moves they can test to advance equity and justice; c) positioning students to take ownership over the learning environment; d) engaging in conversations and develop interpretations of how race, gender, and identity more broadly shape the learning environment; and e)

surfacing minoritized students’ strengths and generating ideas for how to build on them.

We note that these are not contradictory to the tenets of practical measurement (e.g., Takahashi et al., 2022) more broadly, but instead overlay on top of them. While intended to be useful for practitioners, the work of practical measurement does not, on its own, focus attention on and aim to improve either classroom practices and processes or inequities and injustices. We offer these patterns of action, and how measures ought to support them, as ways for researchers, practitioners, and designers of practical measurement tools to orient their work toward equity, justice, and the moment-to-moment work of the classroom.

We turn to describe each of these measurement routines, grounding them in literature on teacher professional learning and teachers’ discourse about students.

Attending to and reflecting on day-to-day practices and processes for equity & justice

Decades of research has highlighted the wide array of objects, people, and relations that teachers could possibly attend to in classrooms and the ways in which teachers filter and focus their attention (van Es and Sherin, 2002, 2021; Sherin and van Es, 2005). In recent years, work on teachers’ noticing has illustrated how teachers engage in noticing for equity (e.g., Williams et al., 2020; Benak, 2022). For instance, van Es et al. (2022) use the notions of stretch and expanse to highlight how teachers’ noticing for equity includes the history, present, and possible futures of their students and communities (“stretch”) and the breadth and range of interactions and activities that happen in classrooms (“expanse”).

Researchers have highlighted how teachers reflect on day-to-day practices (e.g., Cobb et al., 2020; Horn, 2005, 2007) and processes and the ways in which they interrogate them to better center equity (e.g., Jilk and Erickson, 2017; McDuffie et al., 2014). For example, with regard to classroom discourse, teachers can collect observational data about the types of interactions and opportunities made available to students. Teachers are then encouraged to try on racial and/or gender lenses to understand different students’ experiences—and then purposefully plan ways to create new interactional patterns that are more equitable and just (Busby et al., 2017).

What is included within “day-to-day practices and processes for equity and justice” is shaped by our definition of equity and justice. The way in which we define these terms calls for us to attend not only to what may be inequitable experiences happening between students with differing identity characteristics within a single classroom or school (what the term “equity” tends to point to), but also to the daily practices and experiences that feature a just classroom experience. For example, a math classroom where all students are treated as doers of mathematics with opportunities to reflect on their mathematical conceptions is a more just learning space than a classroom where students are categorized by their perceived “math ability” and where the speed of solving a problem is valued rather than the reflections on various math conceptions (Gutiérrez, 2018). Just math classrooms are places where students can flourish as their whole selves and where they are

“rehumanized” as mathematicians. Teachers’ noticings of practices in this measurement routine therefore include not only those practices that give students of differing identity markers equitable experiences and opportunities, but also those that promote the dignity of all students as their full selves in classroom spaces.

We draw on work on teachers’ noticing and on teachers’ reflection of day-to-day practices to highlight the vast array of people, objects, and relations to which teachers can attend, and to argue that focusing on the intersection of classroom interactions and equity and justice is effortful. We contend that practical measurement tools can, if designed explicitly to do so, play a central role in focusing teachers’ and other instructional leaders’ attention on the specific ways that (in)equity and (in)justice emerge in classrooms. Additionally, practical measures can contribute to organizational routines where teachers are engaged in conversations about moment-to-moment interactions and the ways in which those interactions can be more equitable.

Identifying instructional moves teachers can test to advance equity & justice

Secondly, a practical measure for equity should enable teachers to generate instructional routines they can try to advance equity and justice. Consistent with and building on the design principle of practical measures focused on providing actionable data to practitioners (Takahashi et al., 2022), we argue that equity-focused practical measures ought to support routines around generating ideas for instructional moves that teachers can test in their classrooms to advance equity. This second measurement routine is important because it is within the instructional practices that students’ agency and authority can either be constrained or expanded. To advance equity the tool must create a response that affects the opportunities that are made available to students in the classroom. As Horn (2005) argues, classroom artifacts stimulate and can be a valuable source for teachers’ learning. When classroom artifacts—here the data from a practical measurement tool—are used in teachers’ professional conversations, they promote teachers to make explicit their goals and assumptions about teaching, learning, and students—and colleagues then have opportunities to provide a different perspective/interpretation on a classroom artifact, while offering feedback and advice. It is in these generative collegial conversations that teachers may share new instructional routines that are likely to advance equity. We argue that equity- and justice-focused practical measures ought to enable educators in generating new or different actions to better center equity and justice.

Positioning students to take ownership over the learning environment

Third, a practical measurement for equity and justice should support teachers to position students to take ownership over the learning environment. For instance, it may enable teachers to grant students agency in the sensemaking of data generated by the measurement tool. Generally, educational data about students are kept from students themselves, for others to make sense of and

make decisions on behalf of students’ interests. Classrooms are rife with authority structures, many of which position the teacher as the authority which arbitrates student’s reasoning and sense making (Wagner and Herbel-Eisenmann, 2014). Such authority structures shape the learning environment and constrain students’ sense of agency. In order for students to take ownership of their learning environment they should have opportunities to interpret the data generated by them. In this way, the classroom develops an anthropological authority, an authority that “does not aim to preserve any social structure or the rules to any system but to cause a community to grow and be renewed” (Amit and Fried, 2005, p. 151).

Developing interpretations of how identity shapes/ is shaped by the learning environment

Fourth, a practical measure tool for equity should enable teachers to engage in conversations and develop interpretations that illustrate the way race, gender, and identity more broadly shape and are shaped by the learning environment. Researchers have documented in numerous ways the discourses that racialize and gender students’ learning experiences (e.g., Shah, 2017; Ernest et al., 2019; Gholson and Martin, 2019). As such, the practical measurement tool should reveal the local manifestation, for example, by disaggregating data by race and gender to discern what patterns and/or disparities may exist. We caution, though, that the “web of deficit discourses” (Adiredja and Louie, 2020, p. 42) may make it difficult for one to interpret the data in productive ways. Because of the pervasive insidious deficit, racist, and sexist discourses, Louie et al. (2021) argue that teachers need to interrogate the sociopolitical frames that produce deficit discourses and support to create new frames that alternatively position students in non-dominant, asset-based ways. As noted above, collegial conversations can support teachers to make sense of classroom artifacts; however, the teacher communities need resources (e.g., human, protocols) that support them to not reify deficit narratives and to alternatively position students in non-dominant, asset-based ways (Horn, 2005, 2007). This leads us to the next measurement routine.

Surfacing and building on minoritized students’ strengths

Finally, the tool should enable teachers to surface students’ strengths and/or assets. The identification of students’ strengths and/or assets is important for at least two reasons. First, it is through the teachers’ recognition of strengths and/or assets that teachers create new frames to interpret students’ behaviors in equitable ways. Once a teacher has new frames, they can be used within instructional routines (measurement routine 2) (Jilk, 2016). Secondly - and this is relevant to the case we describe below related to status hierarchies in the classroom - teachers must identify and make public marginalized students’ strengths (Cohen and Lotan, 2014). The teacher can use their* disciplinary authority (Amit and Fried, 2005) to position marginalized students as competent

and change the students' expectations of one another. Tools and resources can be designed to support teachers' noticing of students' thinking in strength based ways (Kalinec-Craig et al., 2021). We argue that a practical measurement tool should similarly support teachers' development of asset-based frames.

The five measurement routines that we describe here are not meant to comprise a comprehensive set of criteria of practical measurement for equity and justice - this is not a checklist. Rather, we propose these as ways to inquire and understand how practical measurement may be used to move organizational processes toward equity and justice. We imagine that some will be more salient than others depending on the nature of the focal organizational process and the specifics of a given context.

An exploratory example

We explored the measurement routines for equity and justice as an analytic tool using a practical measure focused on leveling status in small-group work in middle school mathematics classrooms. This measurement tool was designed in partnership with the CARE Network led by the High Tech High Graduate School of Education. In this section, we first provide a description of the CARE Network and its focus on advancing equity in math classrooms across sites. We then offer a detailed description of the design process we engaged in with the CARE Network and a description of the measurement tool, which can be found in [Appendix A](#).

The CARE network

The CARE Network was formed through the partnerships between The Center for Research on Equity and Innovation at High Tech High Graduate School of Education, CREATE (Center for research on Equity, Assessment, and Teaching Excellence) at the University of California San Diego and the California Math Project. The Network's aim is to increase the number of students who are Black, Latine, Indigenous, or from low-income backgrounds who have a strong academic identity and are on-track in 8th grade to graduate high school and successfully enter college and career. At the time of the writing of this article, the network supports 17 public schools across San Diego county in such efforts to improve students' learning experiences as well as their sense of belonging through the following four critical drivers: 1) Establishing early warning systems for noticing when students are struggling, and working collaboratively to design and track interventions that help students succeed (Allensworth et al., 2014); 2) Developing strong student-teacher relationships where adults know students well, see them as whole people, and nurture their positive identity development (Wang and Eccles, 2013); 3) Implementing culturally responsive instruction that disrupts inequitable status dynamics and builds students' cognitive capacities, while affirming their identities as learners and their sense of belonging (Ladson-Billings, 1992; Ladson-Billings, 1995); and 4) Supporting schools in establishing equitable grading practices that support students' growth through regular routines of feedback and by providing students opportunities for re-learning (Feldman, 2018).

School teams that were interested in pursuing the third driver, implementing culturally responsive instruction, formed

teams composed of 3–5 math teachers to explore, test, and refine culturally responsive practices in their classrooms. Math was chosen as the subject area of focus as math is one of the largest contributors to why students' are off-track (Gutman and Midgley, 2000). As a network, this driver was explored through the lesson study professional development model. Lesson study is a system of research and development in which teachers refine ideas as "best practices" through active instruction and observation (Lewis and Hurd, 2011).

In efforts to understand students' experiences with mathematics, teachers implemented a math agency survey. The survey provided teachers insights into how students were perceiving their experience with math instruction as well as their growing mathematical identities. Through multiple cycles of lesson study and collection of survey data, teachers as well as lesson study coaches noticed that the survey question, "The same few students answer most of the math problems in class" was stagnant (only 8% of students disagreed with this statement - the preferred response). In fact, the students' responses to this question demonstrated that there may be a deeper issue happening in participating teachers' classrooms. This finding led coaches to believe that "status" may be at the heart of the issue of why students were strongly agreeing with this survey question.

The CARE Network's use of these surveys, the network convenings where teachers used the data, and the lesson study structure constituted the existing set of measurement routines that the network enacted. The WestEd team sought to co-design measurement tools that enabled CARE Network leaders to integrate the tools within these established measurement routines. Having described the CARE Network's goals, aspirations, and work, we turn to describe status as a central construct for designing a practical measure.

Small group work, status, and equity

Status in a math classroom is defined as the hierarchical social ranking that happens between students based upon their perceptions of who is competent and who is not (Cohen, 1994). One key site where status issues take place is during small group instruction. This became the site of the organizational routines at the center of this practical measurement work. As students work together in small groups, they already have or are forming perceptions of who has high social and/or academic status. Students' perceptions of who has high status can be based on characteristics such as race, gender, age, popularity, and more (Cohen, 1994; Featherstone et al., 2011). Status hierarchies shape equity and justice in math classrooms, because they impact how students work together and the ways in which students bring forth their ideas, share their strengths, and access learning. In small groups, those who are perceived as more competent tend to take up more space, own most of the work and learning, or may exclude others from the math experience. Because this happens within group work, teachers may not know how to identify status at play in small groups and may not have the tools to counteract status issues in small groups.

Status dynamics are an equity concern not only because of the identity characteristics (race, gender, etc.) that are imbued

with certain level of status and associated with mathematical competence, but because the presence and enactment of status hierarchies among students in math classrooms reflects an underlying belief that mathematical ability is inherently unequally distributed among a population. We approach equity through the notion of rehumanizing mathematics (Gutiérrez, 2018). As Gutiérrez puts it, “schooling often creates structures, policies, and rituals that can convince people they are no longer mathematical” (Gutiérrez, 2018, p. 2). Rehumanizing students’ as capable mathematical being can occur in many forms. Gutiérrez (2018) provides eight dimensions of classroom practice for consideration - in this project we focus on the dimension of participation/position. Dominant notions of equity have typically attended to students’ opportunities to access rich mathematics and students’ outcomes in terms of achievement (Gutiérrez, 2018). Attention to participation and positioning considers the processes occurring inside classrooms as students learn mathematics. One aspect of these process is the ways in which students interact with one another in a respectful manner that values their peers’ mathematical thinking and learning; an aspect that Boaler (2008) terms relational equity. A second aspect of these processes is considers “the fair distribution of both participation opportunities and participation itself”-or participatory equity (Shah and Lewis, 2019, p. 423).

We now turn to the process through which we developed a practical measure intended to further equity in math classrooms.

Design process

Prior to working with the CARE Network, the WestEd team developed a conceptual map that articulated a wide range of outcomes, routines, and practices for advancing equity and justice in mathematics. This map was generated through a series of interviews with expert math educators and math education researchers, as well as through engagement with literature on equity in mathematics (e.g., Bartell et al., 2017; Gutiérrez, 2009; Gutstein, 2003). A simplified version of this concept map can be seen in Figure 1.

Following the construction of this conceptual map, the WestEd team began a partnership with the CARE Network given their focus on advancing equity in their mathematics classrooms and their use of networked improvement science as an organizing approach to their work. Work with the CARE Network began in the fall of 2022; using the concept map, the WestEd team asked the network’s leadership to identify outcomes and practices that were the focus of their network that would be the subject of practical measurement development. Additionally, the WestEd team attended multiple CARE Network convenings and a lesson study to understand existing routines around measurement, feedback, and testing. This enabled the WestEd team to be sensitized to these routines to consider when designing measures. After a period of learning more about the CARE Network’s work and aspirations, the WestEd team and the CARE Network jointly coalesced on increasing multivocal interactions in small groups. The CARE Network found this part of the concept map appealing and relevant because they viewed it as aligned with their work on disrupting classroom hierarchies using small groups and groupworthy tasks. The CARE Network

used complex instruction (Cohen and Lotan, 2014; Lotan, 2003) as an organizing frame for their improvement work; thus, attending to small group work and status within small groups was a core feature of equitable mathematics classrooms for them.

The process of developing the small-group status practical measure (SGS-PM) is depicted in Figure 2. To begin developing the prototypes, (the left-most box in Figure 2) the authors drew on literature on status in mathematics classrooms to generate ideas for how to gain insight into status hierarchies in small groupwork. After a series of readings, the authors honed in on a central aspect of status that seemed to be of interest to the CARE Network: That status beliefs are third-order beliefs. That is, status beliefs are “necessarily beliefs about what most people (i.e., the generalized other) believe” (Ridgeway, 2018, p. 317). In the case of mathematics classrooms, this means the student does not have to actually believe they are less competent than another student to invoke the status belief; the student has to believe only that others think there is a hierarchy with one person being more competent than another person. As such, individuals/students will act in accordance with the assumed set of beliefs about the hierarchy since they will be judged by their peers if they do not modify their behavior to match the groups’ expectations (“I have a role to play and will be held accountable to play that role”). Drawing on this component of status, we held a design session with CARE Network leaders to jointly design a range of prototypes to gain insight into status hierarchies in small groupwork. As a result of this design session, the CARE Network and WestEd chose to focus on how students might express status and how they might experience status in the behaviors and interactions in small groups. This led to a focus on actions and, in particular, students’ perceptions of their group mates’ actions toward them and their ideas.

Through interviews with two teachers in the CARE network (Figure 2, Testing 1), we emerged with a short survey that asked students, “Did your team think you had good ideas today?” In addition to the central item, students are also given a second multiple-choice question depending on their response to the first:

- If students respond “yes,” they are asked, “What did your team do to make you feel like you had good ideas?”
- If students respond “no,” they are asked, “What did your team do to make you feel like you did not have good ideas?”
- If students responded, “I’m not sure because I didn’t share my ideas,” they are asked, “Why didn’t you share your ideas?”

A visual representation of the final survey and its branching can be found in Appendix A.

Following testing with teachers, the authors then conducted cognitive interviews (Drennan, 2003) with ten 6th and 8th grade students (Figure 2, Testing 2), who were enrolled in a school that was part of the CARE Network. We asked teachers to connect us with a diverse group of students - in terms of identity characteristics and math achievement experiences - but otherwise left the selection of students to the teachers. Minor revisions were made based on these cognitive interviews and, importantly, served as evidence to the authors that students were responding to items in ways that were consistent with how the authors had intended.

The authors then tested the survey by having teachers administer the survey to students, discuss the results with the

Conceptualizing Math Justice & Equity

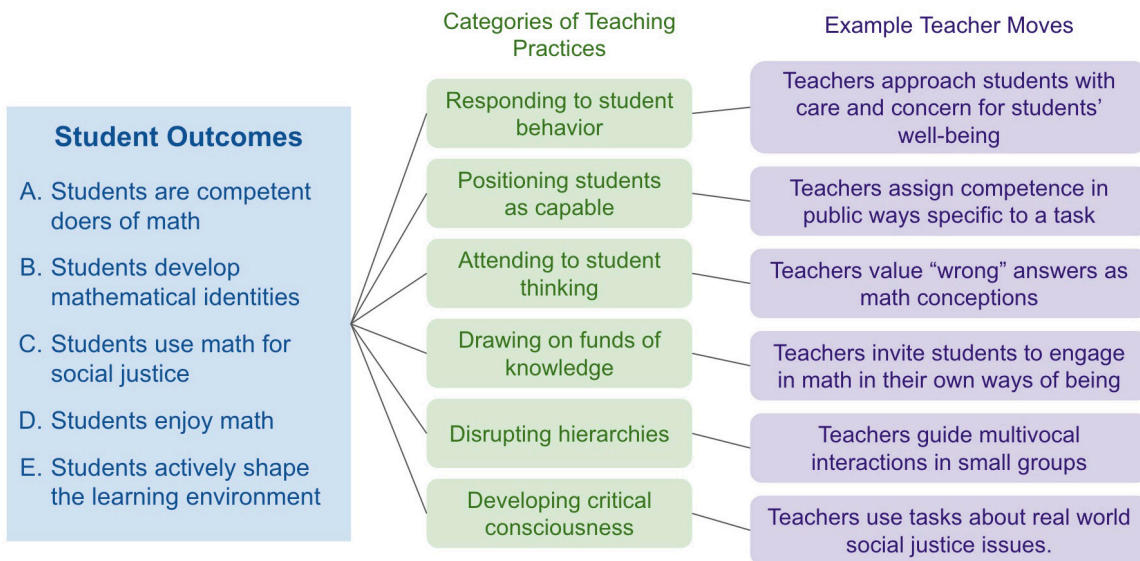
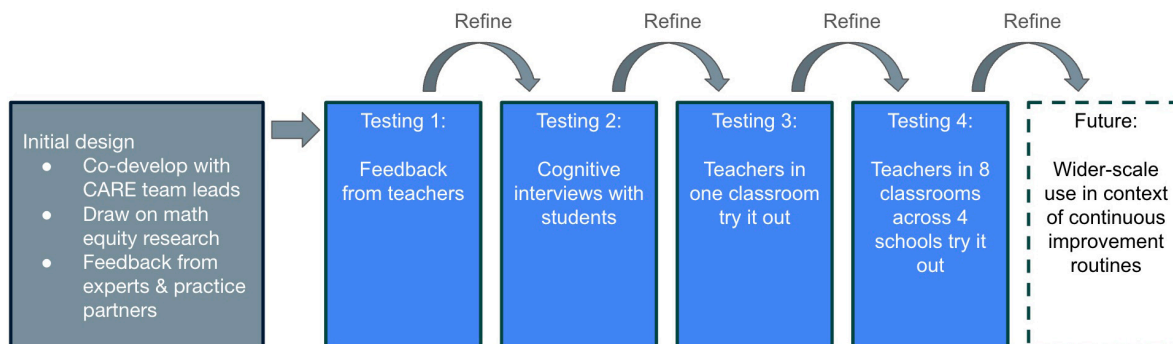


FIGURE 1 Mathematics for equity concept map.

Small Group Status Practical Measurement Development Process



Adapted from Jackson, K., Henrick, E., Cobb, P., Kochmanski, N., & Nieman, H. (2016). *Practical measures to improve the quality of small-group and whole-class discussion [White Paper]*. Retrieved January 2022 from University of Washington: <http://www.education.uw.edu/pmr/files/2016/09/White-Paper.pdf>

FIGURE 2 The small-group status practical measurement development process.

authors, and then facilitate a conversation with students about the results (Figure 2, Testing 3 & 4). The purpose of this round of testing was to understand how teachers and students interpret results, and how teachers would use the data for and with their students. Because of our focus on measurement routines, it was critical for us to understand what transpired in these sensemaking spaces. Although this round of testing did not result in substantive changes to the survey, it did generate insight into what meaning students and teachers made of the survey.

Having described the measure, its design, and the context of its testing, we now turn to interrogate its design and use

using the five patterns of action for centering equity and justice using measurement.

Teachers' reflections on the use of the small-group status practical measure

Four teachers from the CARE network across four middle schools tested the small group status practical measure (SGS-PM)

in a total of eight classrooms, from fall 2023 to spring 2024. The teachers were recommended by the CARE network leads because they were regularly using small group work as a core feature in their classrooms. Three of the four Southern California schools served predominantly Latine students (75% to 92%), while the fourth school served a student population that was approximately half Latine (52%) and one-third white students (32%).

We conducted five classroom observations and five, semi-structured interviews with the teachers. During the classroom observations, we witnessed teachers administering the survey to students in one classroom and discussing survey results with students in four classrooms. Observers took notes, but did not engage students in conversation. During interviews with teachers, we asked them to reflect on their perceived purpose of the tool, logistics of survey administration, and students' reactions to the data. We asked the teachers their impression of the survey results and any other reflections that came up for them. Additionally, we also asked teachers whether the data sparked any ideas for what actions they would take with their students. Interviews were held after teachers administered SGS-PM to their students at least once. Four of the five interviews were individual conversations, and the fifth interview was a group interview with two of the teachers who frequently collaborate with each other (Gabriela¹ and Mazie).

After conducting interviews between Fall 2023 and Spring 2024, interviews were later transcribed and coded for instances of the five features of practical measures for equity and justice. We also noted other emergent themes related to equity in math classrooms. Our discussion here focuses mainly on what teachers shared during these interviews.

Attending to and reflect on day-to-day practices and processes for equity & justice; and identifying instructional moves teachers can test to advance equity & justice

While making sense of the data from the SGS-PM, teachers reflected on their students' experiences of small group work in the classroom, and the practices and processes that enable students to voice their thinking and treat each other's thinking respectfully in small groups. These reflections sometimes led to discussions about next steps in their practice, so we will share the teachers' words that touch on the first two features of practical measures for equity together.

In some cases, teachers gained insight into students' experiences that were surprising to them. In Louise's class after the first administration of SGS-PM, 12 students reported that their team thought they had good ideas, two reported that their team did not think they had good ideas, and seven said that they did not share their ideas (Figure 3A). Of the two students who responded "No," one reported that "My team didn't ask me about my ideas," while another student indicated that "My team shut down my ideas" (Figure 3C). Louise was troubled by the negative responses. She remarked, "When it came to the 'No's,' it broke my

heart because even if it's just one student, 'my team shut down my ideas,' oh my gosh, it broke my heart to be honest." This reflection, in turn, led her to want to resurface discursive classroom norms to her class, in hopes that they engage one another differently. She continued, "So I think it's a good moment to remind kids, 'Hey, we're in this together, this is a safe environment. Let's be kind to one another. Let's listen to other ideas.'" We interpreted Louise's comments as reflecting on what she thought was acceptable and positive data in her class—in this case, finding it unacceptable to have one student feel as if their ideas were shut down—and generating an action that she could take.

Another teacher, Olivia, thought the data for her classes might be "too good to be true," sharing her reflections that she was expecting more equal proportions of yeses and nos. However, she also shared a similar sentiment as Louise's about even one student saying "no" or "I didn't share my ideas" is reason for the classroom community to reflect on their actions and improve. (In Olivia's first class, 20 students responded "yes," one responded "no," and four responded "I didn't share"; in her second class, 14 students responded "yes" and 8 responded "I didn't share.") She explained, "there's still one person who feels this way. So we're a team or community. How can we make that person feel included or feel more comfortable with our community. . .?" However, Olivia did not generate an idea for actions she could take to make students feel more comfortable. We interpreted Olivia's comments as an indication that the tool did support her in reflecting on day-to-day moves she makes to center equity, though it did not necessarily support her in generating an action she could take to do so.

For one of her classes, Gabriela expressed surprise by a response from a particular student, and speculated about what might have led to that response. "I was really shocked that he said that he didn't share, because he's usually pretty vocal. I'm thinking that the girls in that group might have taken over." This led to ideas about group structures and processes that she might employ. "I think that one of the things that in terms of my next steps, when I saw that, was like, huh, I think I need to make sure that I at least add that one role of making sure that there's a monitor of everybody's voice being heard."

In a second interview 4 months later, Gabriela emphasized the importance of the structure of teamwork, allocating speaking turns and times to start a conversation to ensure that each team member voices their ideas, before allowing for free-flowing conversation. "I always make sure that I start the group work in that very structured, organized way of who is speaking so that everybody's voice, like, it's almost like a warmup, right?" She also discussed how she prompts students to be in collaborative conversation about their ideas, to "muddle through" ideas together to head in the right direction on rigorous tasks. She referred to a poster she has in her class showing many different ways to be mathematical - not only one way, and emphasized how making mistakes and being confused are "essential to learning."

Gabriela's comments highlighted a range of reflections on her classroom and her students, and included potential actions she could take based on the data. Her reflection that she was "shocked" that a vocal student did not share illustrates how the tool may have provided her with information that was different than her perceptions of students. Her subsequent comment that she was going to assign a monitor to ensure equity of voice illustrates how

¹ All names included in this manuscript are pseudonyms.

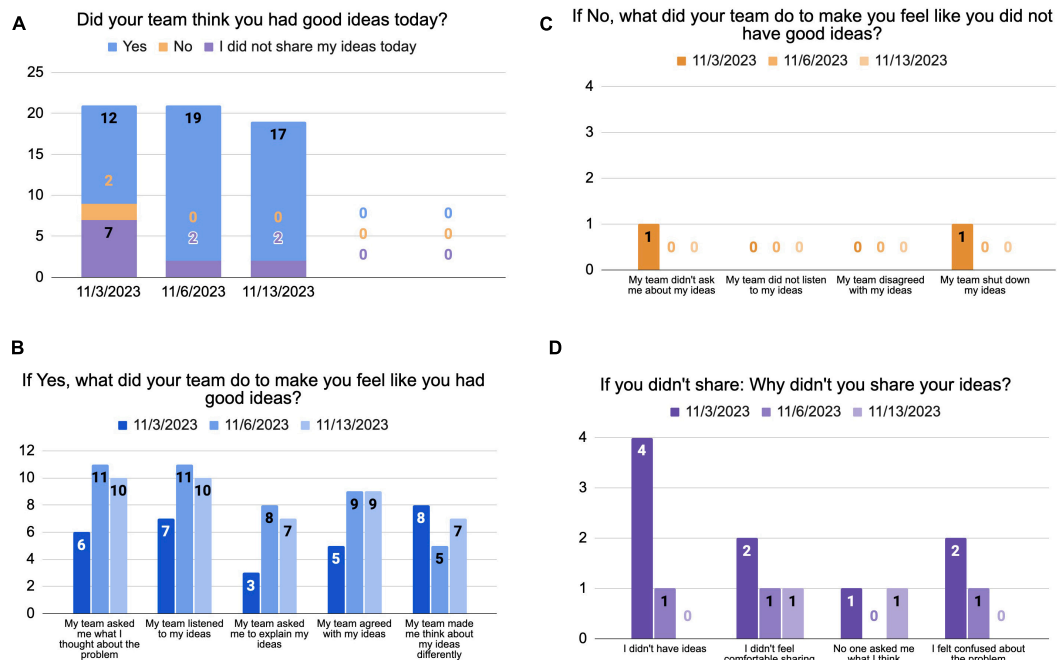


FIGURE 3 (A) Small-group status practical measure data for Louise’s class. Responses to the question, “Did your team think you had good ideas today?” (B) Small-group status practical measure data for Louise’s class. Responses to the question encountered only by students who responded “Yes” to the first question. (C) Small-group status practical measure data for Louise’s class. Responses to the question encountered only by students who responded “No” to the first question “What did your team do to make you feel like you did not have good ideas?” (D) Small-group status practical measure data for Louise’s class. Responses to the question encountered only by students who responded “I didn’t share my ideas” to the first question: “Why didn’t you share your ideas?”

the data—and the conversation about the data—generated ideas for new and different actions she might take in her classroom.

While the SGS-PM did appear to enable teachers to attend to day-to-day practices that surface inequities in participation, we note that a shortcoming of the tool was that it did not surface patterns of inequities in status as measured by the tool. For instance, the measure did not appear to surface whether or not Black students were especially negative in their responses compared to their Latine or white peers. This shortcoming is explored in detail in later sections.

Positioning students to take ownership over the learning environment

We highlight this emergent routine around enabling students to take ownership over the classroom and its norms using the case of Olivia. Olivia described the data from the SGS-PM as being for students, in contrast to other student survey data that are mainly for her as a teacher:

“It’s more from them rather than just me and my observations. I don’t see everything, it’s nice to hear them [...] Because [all the other] surveys have been more for me, and I guess it’s nice to frame it to them. ‘Here, this is for you. It’s from you, for you.’ [...] I really like the idea of framing it more for them and how they can create community for themselves, support each

other and be real about where they are as members of their community with each other.”

We interpreted Olivia’s comments that other “surveys have been more for me” as illustrating how the SGS-PM was more uniquely suited to soliciting insight from students and enabling them to “create community for themselves.” In our analysis of this quote, we conjectured that the measurement tool enabled teachers to reposition students as actors who can actively shape how the classroom unfolds to better support shared discursive norms around hearing and supporting one another.

Our design of the SGS-PM included teachers sharing the data back with students as an essential feature of the measurement process. Therefore, for the four teachers, “testing” SGS-PM included not only administering the survey and reviewing the data themselves, but sharing the data back with their students after each survey administration. The graphical displays were additive (a new column for each time point), showing change over time for each class after each survey administration (ex. Figures 3A–D). Crucially, we offered no guidance to teachers around how they ought to share data back with students. For instance, we could have generated a protocol that focused squarely on having students explain their responses to the data. We did not offer guidance to teachers to understand how the tool would be taken up in practice as part of our design process.

In all four classroom observations, teachers shared the data back with students, led a discussion about what students noticed in the data, and then engaged students in generating ideas for

action steps, usually in small groups. Teachers recorded ideas from small groups about actions they would want to take next. Ideas included, “Helping each other and not stressing them out. Not telling them to hurry”; “Don’t make fun of the people for incorrect answers. Help them and explain.”; and “Don’t leave anyone out and make sure everyone spoke.” In one class, students got into a discussion about how to respond when someone in their group shares an answer that they think is incorrect. One student offered that you should show the student how you solved it by saying, “This is how I did it.” Another student chimed in, “show them step by step.” We interpreted teachers’ moves to share data back in ways that enabled students to generate ideas for how to improve small group discussions as evidence that the tool enabled an emerging routine to support students in taking ownership over the learning environment. By having students generate ideas for how to improve data, teachers repositioned students as active participants who can (re)shape their interactions to better live up to shared and agreed-upon norms.

Mazie discussed how students led her to want to increase small group experiences for their students.

I think one of the things that most of them were talking about was just more team experiences. We focus a lot on partner work. And so offering more opportunities for teamwork while also being able to follow the curriculum. So I heard them, I was listening.

While there were numerous instances where the data led teachers to think about the pedagogical decisions and actions, this was one instance in which students directly influenced how a teacher was thinking about structuring her class. Thus, we interpreted Mazie’s comment as evidence that the measurement tool and data had played a role in surfacing what students wanted—more team work as opposed to partner work—and shifted what she then did to be responsive to student voice.

Taken together, we interpret teachers’ quotes and use of the data from the SGS-PM to conjecture that the tool can support measurement routines around enabling students to take ownership over the classroom.

Developing interpretations of how identity shapes/ is shaped the learning environment

Evidence of this aspirational routine was the least present across all of the interviews. In two of the five interviews, race, gender, and other identity markers were not discussed at all. In two other interviews, gender and English learner identities were referenced in passing; however, there was only one interview that dove in more deeply into how student identities impacted group dynamics, and consequently, the survey responses.

Although the lack of conversation around identity portrays that the measurement tool may not support those routines, we highlight an instance in which the measure and resulting data enabled a conversation about Black students in a predominantly

Latine school. One of the interviews we conducted were with a pair of teachers, Mazie and Gabriela, who taught the same subject and grade level and worked closely together at the school. In this interview, Mazie and Gabriela talked to one another about the data and, in turn, issues of race. After their first survey administration, Mazie noticed that out of the students who indicated “no,” their team did not think they had good ideas on that day, one was a Black student in her class of predominantly Latine students who had been experiencing racial marginalization and offensive slurs. Mazie explained:

[U]nfortunately at our school, we don’t have too many Black students. . . . I do feel like that kid particularly feels left out sometimes. And I tell the kids, I love it when they speak Spanish. . . . but if not everyone at your table understands Spanish, that’s when we don’t speak Spanish. So I know that that student has run into that issue before. And then also the other day we had another conversation about, no, we don’t say the N-word. . . . That’s not okay, that’s racist. The history behind it. . . . that student was very appreciative, because apparently this has been going on. So when I look at those particular students that answered, no, I know that it’s not necessarily because of the team dynamic, it’s because of some outside problems that they’re experiencing.

Here, Mazie is making sense of the data by reflecting on what she knows about the group dynamics, and how issues of race, ethnicity and language have impacted one small group. We witness the tension between welcoming multiple languages in the classroom and being careful not to “other” English-only speakers, especially when they are carrying minoritized identities, such as being one of only a few Black students in the school. We also interpreted Mazie’s comment as evidence that the data and response from the Black student in the class sparked her to contextualize that students’ experience in the school and in the classroom within this particular lesson, prompting her to manage the racial dynamics that emerge in a multiracial classroom.

In an interview around her classroom data, Gabriela later commented about her own efforts to support Black students in her class:

I do have two Black students who I know that I’m really aware of making sure how they’re feeling. . . . They’re both really focused. . . . because they’re trying to navigate the space and making sure that they’re trying to block out some of the other stuff. And so I’m just trying to think about some of the issues that they’re facing and how to bring up issues. . . . for example, a lot of the progress that Latinos have made have been because our Black brothers and sisters have elevated. They were the ones who were showing us the way.

For Gabriela, building a strong classroom culture and supporting Black students in her class includes attending to their voices and drawing on the history of cross-racial collaboration and support. In many ways, this reflection feels separate from mathematics, however the subtext here is that both teachers are

aware of underlying tensions – in this case, between Black and Latino students – that may be impacting students' ability to learn together during small group work. Additionally, we interpreted that the remarks about Black students that Mazie surfaced earlier in the conversation enabled Gabriela to discuss issues of race and how Black students were positioned in the school, and that the data played a role in surfacing this conversation among colleagues about race and racial dynamics.

We reiterate that this conversation between Mazie and Gabriela was the only conversation around race that surfaced in our interviews with teachers. We also note that, although the tool was a catalyst for talking about race—given the negative response that a Black student submitted—Mazie and Gabriela very likely had previous conversations about race, racism, and racial dynamics at the school; this measure did not appear to spark this discussion for the first time. We surface this to highlight that although the data were catalytic in some way, the context of the school, as well as the work and conversations these two teachers were engaged in, were tightly connected to the insights that emerged from the data discussion.

Surfacing and building on minoritized students' strengths

We found little evidence in our interviews that the measurement tool or the data generated opportunities for teachers to talk about surfacing and building on minoritized students' strengths. Generally, teachers pointed out obvious strengths in the data (e.g., when the majority of students said they felt comfortable sharing ideas). For instance, Gabriela remarked, "My focus usually is obviously celebrating the fact that a lot of people, you know, feel comfortable and that they feel that it's positive to do group work." In one interview, Louise reflected on how even quiet students should be able to express their ideas through various modes of participation:

I have a couple of shy kids. . . .but that's part of their personality. They're just quiet. But just because they're quiet doesn't mean that they don't have ideas. So how can they express their ideas in a different way?

Similarly, Gabriela discussed with her students that it is okay to show up in different ways:

We're not all meant to have the same strengths. And so I talk to them a lot about that as well, about the fact that we're meant to have different strengths because we're meant to be a community of people who bring something else to the table.

In each of these comments from teachers, the strengths they surfaced focused on their modes of participation, either what they knew about how the students prefer to participate (e.g., "shy kids") or how they reported on each other's participation in the survey data. However, teachers' comments did not focus on *minoritized* students' strengths, disentangling their modes of participation and engagement from their identities. Given the lack of discussion that emerged around issues of identity more generally,

this was unsurprising to us. We interpret this lack of focus on minoritized students' strengths as evidence that the measurement tool might fall short of enabling conversations that broached what strengths minoritized students bring to classrooms, and how they can be leveraged.

We have sought to illustrate what it might look like to interrogate practical measurement tools from the lens of the kinds of routines they engender using data we collected from the process of designing this measure. While the SGS-PM appeared to be uniquely well-suited to support some measurement routines for equity and justice (i.e., enabling students to take ownership over the classroom; enabling reflection and ideas for action), it did not seem to be conducive to others (i.e., centering issues around identity; surfacing minoritized students' strengths). At the same time, evidence from our design process revealed that the SGS-PM also called attention to an aspect of the day-to-day classroom experience where status hierarchy plays out. It also enabled teachers to engage in reflection and generate ideas for pedagogical action to attend to *whose voice* is valued among students in small groups.

Our aim was not to spotlight a measurement tool as an exemplar, but instead, to illustrate how our aspirational measurement routines for equity and justice might be used to benchmark and test the extent to and modes through which practical measures support those measurement routines. We now return to our discussion of measurement routines for equity and justice and chart paths for possible future research on measurement for improvement in the context of equity and justice work.

Discussion

In this paper, we sought to generate a set of measurement routines for advancing equity and justice in the day-to-day, moment-to-moment work of classrooms and schools. We generated measurement routines, rather than design features or principles, to elevate the kinds of actions that we argue practical measures ought to support if they aim to advance equity and justice. We reject any notion that measurement tools are silver bullets that, on their own, can advance equity and justice. Instead, our focus on the five measurement routines we articulated motivates an attention to the ways in which measurement tools are taken up in practice and in the day-to-day organizational routines, and how their enactments come to give measurement tools meaning.

Although our paper sought to focus on measurement routines rather than design features, we believe our work has implications for how measurement tools are and can be designed to advance equity and justice. In line with our connection to [Pentland and Feldman \(2008\)](#), those who design measures for equity and justice ought to carefully consider and work toward the organizational routines they hope to support, rather than designing measures assuming an emergence of routines. In our paper, we highlighted how an equity-focused measurement tool engendered some equity- and justice-focused measurement routines but fell short of enabling others. Future design work and iteration on this measure ought to focus on how it can more reliably surface conversations about race, gender, and other forms of identity in small group work. For instance, we can imagine connecting the

data to identify characteristics, that could be used to disaggregate students' responses by race or gender. Conversation protocols might include questions asking teachers to attend to these issues. If this tool were developed to ask students about how they felt their teachers treated their ideas, this would raise another set of possibilities for the reflections and actions they might spur (and may come with challenges for teachers who may feel judged and critiqued.) These changes, and their accompanying conjectures about the kinds of routines they support, ought to be tested.

In addition, our work highlights the need for practical measurement designers to attend to the wide range of components of practical measurement. This includes the data that are generated, its analysis and visualization, the sensemaking and connection to action, and all of the discursive practices that comprise each aspect of this process. Designing measurement is not solely about the measurement tool, but also about designing these processes. Our measurement routines, and our illustration of them using an equity-focused measure, highlight how the data that results becomes a resource in the sensemaking and use by educators. They are not a resource in themselves, but they are resourced (Pentland and Feldman, 2008) by teachers who took them up as evidence (to varying degrees) of what students were experiencing in small group work in their classrooms, and as a tool to spur a collective reflection among the class.

The sites of data sensemaking feel particularly critical to attend to. While we interviewed teachers to understand their use and perceptions of SGS-PM, in practice, those interviews also functioned as sites of professional sensemaking, reflection, and consideration of next steps. What are the professional learning routines that are embedded in schools, and how can they be sites of "Bending routines toward racial justice. . ." (Diamond and Gomez, 2023, p.5)? The five measurement routines we propose here point to some activities and discussions that may lead to the kinds of reconsiderations toward more just organizational routines. One type of discussion is to highlight a disconnect between a routine as it is theorized and how it is actually enacted, creating a dissonance that compels educators to change how routines are designed or enacted. Another is to "reflect-in-action," so that educators can engage in their work routines and simultaneously evaluate it. The authors assert that especially doing this in a collective setting with transparency about practice can promote new actions and new routines. In both cases, the practical measurement routines that we have been discussing could offer concrete opportunities for these kinds of "productive ways to pause, disrupt, and deconstruct how routine activities reproduce racial injustice and create more racially just practice" (Diamond and Gomez, 2023, p.5).

One of the avenues for future exploration is how the measurement routines we describe here connect with systemic improvement routines - inquiry cycles, team huddles, and the collective critical reflection and revision of theories of improvement (Grunow et al., 2024). We see opportunities for the five measurement routines to occur within team huddles, in the "study" part of a plan-do-study-act cycle, and other spaces where data are brought in to shape an improvement effort. Yet we did not conduct this work in the SGS-PM example because the work did not reach that phase. We see this as an opportunity for future work and learning.

Furthermore, we did not delve into the measurement development process in this paper, other than to describe the process we engaged in, but this is another avenue of future exploration and understanding. What are the features of an inclusive and equity-centered practical measurement development process? Whose voices are included when and in what ways? We were mindful of co-developing SGS-PM with practice colleagues, including teachers and a diverse group of students in its development process. But there is still much to be understood about how to elevate these voices.

Conclusion

The push for "data-driven decision making" in the standards and accountability movement in education has not led to systemic transformations toward more equitable and just learning experiences for students from historically marginalized groups. Practical measurement may provide another avenue of data use to shift organizational routines toward greater equity and justice. Through the measurement routines that they may foster, we see affordances and opportunities. Yet this is also an emergent field of work. We urge the field of continuous improvers to join in this shared learning for the goals of system transformation. How do the five measurement routines we have described here play out for different practical measures for equity and justice in varying contexts? Are there other measurement routines to add? What does it take for these measures to promote redesigning organizational routines? Not all organizational routines happen at the classroom level. What does this look like for practical measurement at the larger school or district level? What if the users are school and district leaders rather than teachers? Do the core ideas of the five routines still apply?

With increasing interest among educators in leveraging improvement science methods to advance equity and justice in schools, we have an opportunity to move beyond the dominant educational narrative about differential outcomes by race, to home in on and work toward changing the organizational routines that marginalize non-dominant student groups in the day-to-day processes of schooling. Instantiations of measurement routines for justice and equity have meaning in enacted practice. In this paper, we have outlined the ostensive features of what these measurement routines are; but they are imbued with meaning only as actual, live routines - as performative routines.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ST: Writing – review and editing, Writing – original draft. CS: Writing – review and editing, Writing – original draft. BJ: Writing – original draft. JC: Writing – review and editing, Writing – original draft. CT: Writing – original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The work described in this manuscript was funded by the Bill & Melinda Gates Foundation. The ideas expressed in this manuscript are the authors' own, not those of the Gates Foundation.

Acknowledgments

We are grateful to the other members of the CARE Network community, especially Daisy Sharrock, Yekaterina Milvidskaia, Ruth Smith, and the teachers and students who generously and thoughtfully participated in this development effort. Thank you to Reginald Thomas who made the data visualizations that supported this work.

References

- Adiredja, A. P., and Louie, N. (2020). Untangling the web of deficit discourses in mathematics education. *Learn. Maths.* 40, 42–46.
- Allensworth, E., Julia, A., Gwynne, J. A., Moore, P., and de la Torre, M. (2014). *Looking forward to high school and college: Middle grade indicators of readiness in Chicago public schools*. Chicago, IL: Consortium on Chicago School Research.
- Amit, M., and Fried, M. N. (2005). Authority and authority relations in mathematics education: A view from an 8th grade classroom. *Educ. Stud. Maths.* 58, 145–168.
- Bartell, T., Wager, A., Edwards, A., Battey, D., Foote, M., and Spencer, J. (2017). Toward a framework for research linking equitable teaching with the standards for mathematical practice. *J. Res. Maths. Educ.* 48, 7–21.
- Benak, S. L. (2022). *Noticing-for-equity: Using video clips to understand preservice secondary school mathematics teachers' equitable noticing*. [Ph.D. thesis]. San Diego, CA: San Diego State University.
- Bensimon, E. M. (2018). Reclaiming racial justice in equity. *Change Magazine High. Learn.* 50, 95–98.
- Bertrand, M., and Marsh, J. (2021). How data-driven reform can drive deficit thinking. *Phi Delta Kappan* 102:3936.
- Boaler, J. (2008). Promoting 'relational equity' and high mathematics achievement through an innovative mixed-ability approach. *Br. Educ. Res. J.* 34, 167–194.
- Bryk, A. S., Gomez, L. M., and Grunow, A. (2011). "Getting ideas into action: Building networked improvement communities in education," in *Frontiers in sociology of education*, ed. M. Hallinan (Dordrecht: Springer Netherlands), 127–162.
- Busby, L., Goff, C., Haton, D., Herbel-Eisenmann, B., Leah, J., Loeffert, C., et al. (2017). "Supporting powerful discourse through collaboration and action research," in *Access and equity: Promoting high-quality mathematics*, eds A. Fernandez, S. Crespo, and M. Civil (Reston, VA: NCTM), 59–76.
- Cobb, P., Jackson, K., Henrick, E., and Smith, T. M. (2020). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Cambridge, MA: Harvard Education Press.
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Rev. Educ. Res.* 64, 1–35.
- Cohen, E. G., and Lotan, R. A. (2014). *Designing groupwork: Strategies for the heterogeneous classroom*, 3rd Edn. New York, NY: Teachers College Press.
- Cunningham, J. (2019). Missing the mark: Standardized testing as epistemological erasure in U.S. schooling. *Power Educ.* 11, 111–120.
- Cunningham, J., Takahashi, S., and Hirschboeck, K. (2024). *Using the EQUIP observation tool to advance equity in math classrooms*. San Francisco, CA: WestEd.
- Datnow, A., Greene, J. C., and Gannon-Slater, N. (2017). Data use for equity: Implications for teaching, leadership, and policy. *J. Educ. Admin.* 55, 354–360.
- Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociol. Educ.* 80, 285–313.
- Diamond, J. B., and Gomez, L. M. (2023). Disrupting white supremacy and anti-Black racism in educational organizations. *Educ. Res.* doi: 10.3102/0013189X231161054
- Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *J. Adv. Nurs.* 42, 57–63.
- Engeström, Y. (1999). Activity theory and individual and social transformation. *Perspect. Activ. Theory* 19, 19–30.
- Ernest, J. B., Reinholz, D. L., and Shah, N. (2019). Hidden competence: Women's mathematical participation in public and private classroom spaces. *Educ. Stud. Maths.* 102, 153–172.
- Featherstone, H., Crespo, S., Jilk, L., Oslund, J., Parks, A., and Wood, M. (2011). *Smarter together: Collaboration and equity in the elementary math classroom*. Reston, VA: National Council of Teachers of Mathematics.
- Feldman, J. (2018). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms*. Thousand Oaks, CA: Corwin Press.
- Feldman, M. S. (2000). Organizational routines as a source of continuous change. *Organ. Sci.* 11, 611–629.
- Feldman, M. S. (2004). Resources in emerging structures and processes of change. *Organ. Sci.* 15, 295–309.
- Garner, B., Thorne, J. K., and Horn, I. S. (2017). Teachers interpreting data for instructional decisions: Where does equity come in? *J. Educ. Admin.* 55, 407–426.
- Gholson, M. L., and Martin, D. B. (2019). Blackgirl face: Racialized and gendered performativity in mathematical contexts. *ZDM* 51, 391–404.
- Grunow, A. J., Park, S., and Bennett, B. (2024). *Journey to improvement: A team guide to system change in education, health care, and social welfare*. Lanham, MD: Roman & Littlefield Publishers.
- Gutiérrez, R. (2009). Framing equity: Helping students "play the game" and "change the game". *Teach. Excell. Eq. Maths.* 1:1.
- Gutiérrez, R. (2018). *The need to rehumanize mathematics: Rehumanizing mathematics for black, indigenous, and latinx students*. Reston, VA: National Council of Teachers of Mathematics, Annual Perspectives in Mathematics Education.
- Gutman, L. M., and Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African American students during the middle school transition. *J. Youth Adolesc.* 29, 223–249.
- Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban. *Latino Sch. J. Res. Maths. Educ.* 34, 37–73.
- Hinnant-Crawford, B. N. (2020). *Improvement science in education: A primer*. Gorham, ME: Myers Education Press.
- Horn, I. S. (2005). Learning on the job: A situated account of teacher learning in high school mathematics departments. *Cogn. Instruct.* 23, 207–236.
- Horn, I. S. (2007). Fast kids, slow kids, lazy kids: Framing the mismatch problem in mathematics teachers' conversations. *J. Learn. Sci.* 16, 37–79.
- Jilk, L. M. (2016). Supporting teacher noticing of students' mathematical strengths. *Maths. Teach. Educ.* 4, 188–199.
- Jilk, L. M., and Erickson, S. (2017). Shifting students' beliefs about competence by integrating mathematics strengths into tasks and participation norms. *Access Eq.* 78, 6–8.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

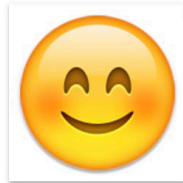
All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Kalinec-Craig, C. A., Bannister, N., Bowen, D., Jacques, L. A., and Crespo, S. (2021). "It was smart when:" Supporting prospective teachers' noticing of students' mathematical strengths. *J. Maths. Teach. Educ.* 24, 375–398.
- Ladson-Billings, G. (1992). Reading between the lines and beyond the pages: A culturally relevant approach to literacy teaching. *Theory Pract.* 31, 312–320.
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *Am. Educ. Res. J.* 32, 465–491.
- Latour, B. (2007). *Reassembling the social: An introduction to actor-network-theory*. Oxford: Oxford University Press.
- Lewis, A. E., and Diamond, J. B. (2015). *Despite the best intentions*. Oxford: Oxford University Press.
- Lewis, C., and Hurd, J. (2011). *Lesson study step by step: How teacher learning communities improve instruction*. Portsmouth: Heinemann.
- Lotan, R. A. (2003). Group-worthy tasks. *Educ. Leadersh.* 60, 72–75.
- Louie, N., Adiredja, A. P., and Jessup, N. (2021). Teacher noticing from a sociopolitical perspective: The FAIR framework for anti-deficit noticing. *ZDM Maths. Educ.* 53, 95–107.
- McDuffie, A., Foote, M. Q., Bolson, C., Turner, E. E., Aguirre, J. M., Bartell, T. G., et al. (2014). Using video analysis to support prospective K-8 teachers' noticing of students' multiple mathematical knowledge bases. *J. Maths. Teach. Educ.* 17, 245–270.
- Nussbaum, M. C. (2000). *Women and human development: The capabilities approach*, Vol. 3. New York, NY: Cambridge University Press.
- Orlikowski, W. J., and Scott, S. V. (2008). Sociomateriality: Challenging the separation of technology, work and organization. *Acad. Manag. Ann.* 2, 433–474. doi: 10.1080/19416520802211644
- Pentland, B. T., and Feldman, M. S. (2008). Designing routines: On the folly of designing artifacts, while hoping for patterns of action. *Inform. Organ.* 18, 235–250.
- Ridgeway, C. L. (2018). "Status construction theory," in *Contemporary social psychological theories*, 2nd Edn, ed. P. J. Burke (Redwood City, CA: Stanford University Press), 315–339.
- Shah, N. (2017). Race, ideology, and academic ability: A relational analysis of racial narratives in mathematics. *Teach. Coll. Rec.* 119, 1–42.
- Shah, N., and Lewis, C. M. (2019). Amplifying and attenuating inequity in collaborative learning: Toward an analytical framework. *Cogn. Instruct.* 37, 423–452.
- Sherin, M., and van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *J. Technol. Teach. Educ.* 13, 475–491.
- Takahashi, S., Jackson, K., Norman, J., Ing, M., and Krumm, A. E. (2022). "Measurement for improvement," in *The foundational handbook on improvement research in education*, eds D. J. Peurach, J. L. Russell, L. Cohen-Vogel, and W. R. Penuel (Lanham, MD: Rowman & Littlefield), 423–432. doi: 10.1097/CEH.0000000000000331
- Valdez, A., Takahashi, S., Krausen, K., Bowman, A., and Gurrola, E. (2020). *Getting better at getting more equitable: Opportunities and barriers for using continuous improvement to advance educational equity*. WestEd.
- van Es, E. A., and Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *J. Technol. Teach. Educ.* 10, 571–596.
- van Es, E. A., and Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM Maths. Educ.* 53, 17–27.
- van Es, E. A., Hand, V., Agarwal, P., and Sandoval, C. (2022). Multidimensional noticing for equity: Theorizing mathematics teachers' systems of noticing to disrupt inequities. *J. Res. Maths. Educ.* 53, 114–132.
- Wagner, D., and Herbel-Eisenmann, B. (2014). Mathematics teachers' representations of authority. *J. Maths. Teach. Educ.* 17, 201–225.
- Wang, M. T., and Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learn. Instruct.* 28, 12–23.
- Williams, A. D. P., Higgs, J. M., and Athanases, S. Z. (2020). Noticing for equity to sustain multilingual literacies. *J. Adolesc. Adult Lit.* 63, 457–461.

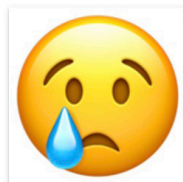
Appendix

Did your team think you had good ideas today? *

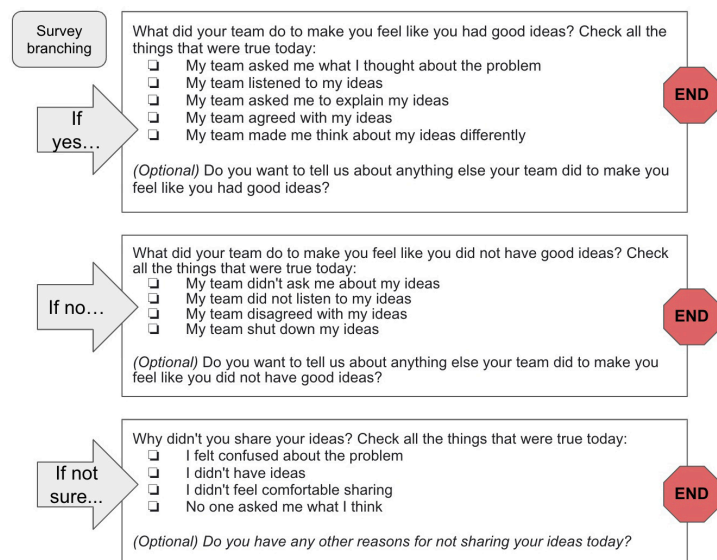
Yes



No



I'm not sure because I didn't share my ideas



APPENDIX A
Measurement tool for status experiences in small group work.