Article

Teachers College Record
2022, Vol. 124(11) 179–217

DOI: 10.1177/01614681221140963
journals.sagepub.com/home/tcz
SAGE

Initial Steps in Developing Classroom Observation Rubrics Designed Around Instructional Practices that Support Equity and Access in Classrooms with Potential for "Success"

Jonee Wilson, PhD[1]

Abstract

Background: The field of mathematics education has made progress toward generating a set of instructional practices that could support improvements in the learning opportunities made available to groups of students who historically have been underserved and marginalized. Studies that contribute to this growing body of work are often conducted in learning environments that are framed as "successful." As a researcher who is concerned with issues of equity and who acknowledges the importance of closely attending to the quality of the mathematical activity in which students are being asked to engage, my stance on "successful learning environments" pulls from both Gutiérrez's descriptions of what characterizes classrooms as aiming for equity and the emphasis on the importance of conceptually oriented goals for student learning that is outlined in documents like the *Standards*. Though as a field we are growing in our knowledge of practices that support these successful learning environments, this knowledge has not yet been reflected in many of the observational tools, rubrics, and protocols used to study these environments. In addition, there is a growing need to develop empirically grounded ways of attending to the extent to which the practices that are being outlined in research literature actually contribute to the "success" of these learning environments.

Purpose: The purpose of this article is to explore one way of meeting this growing need by describing the complex work of developing a set of classroom observation rubrics (the *Equity and Access Rubrics for Mathematics Instruction*, EAR-MI) designed

[1]University of Virginia, Charlottesville, VA, USA

Corresponding Author:
Jonee Wilson, Ph.D., University of Virginia, Charlottesville, VA 22903-1738, USA.
Email: vma8ze@virginia.edu

to support efforts in identifying and observing critical features of classrooms characterized as having *potential* for "success." In developing the rubrics, I took as my starting place findings from an analysis that compared a set of classrooms that were characterized as demonstrating aspects of successful learning environments and a set of classrooms that were not characterized as successful. This paper not only describes the process of developing the rubrics, but also outlines some of the qualitative differences that distinguished more and less effective examples of the practices the rubrics are designed to capture.

Research Design: In designing the rubrics, I engaged in multiple cycles of qualitative analyses of video data collected from a large-scale study. Specifically, I iteratively designed, tested, and revised the developing rubrics while consistently collaborating with, consulting with, and receiving feedback from different experts in the field of education.

Conclusions: Although I fully acknowledge and recognize that there are several tensions and limitations of this work, I argue that developing rubrics like the EAR-MI is still worthwhile. One reason that I give for continuing these types of efforts is that it contributes to the work of breaking down forms of practice into components and identifying key aspects of specific practices that are critical for supporting student learning in ways that make potentially productive routines of action visible to and learnable by others, which may ultimately contribute to the development of more successful learning environments. I also argue that rubrics like the EAR-MI have the potential to support researchers in developing stronger evidence of the effectiveness of practices that prior research has identified as critical for marginalized students and in more accurately and concretely identifying and describing learning environments as having potential for "success."

Keywords

Mathematics, Classroom Observation, Rubrics, Equity, Instructional Practices

In recent years, the field of mathematics education has made progress toward generating a set of instructional practices that have the potential to support current and future mathematics teachers to improve learning opportunities among groups of students who historically have been underserved and marginalized (e.g., emergent multilinguals or African American students who have been labeled as underperforming). As a demonstration of this progress, in focusing on Black or African American learners as a specific subgroup that has been underserved, Robert Berry (2020) posted a message, while serving as National Council of Teachers of Mathematics (NCTM) president, listing several resources that have been published in the last decade that address issues around supporting Black students in mathematics (see Bonner, 2010; Clark et al., 2013; Frank et al., 2018; Id-Deen, 2018; Jett et al., 2015; Joseph et al., 2019; White et al., 2016; Wilson et al., 2019). Studies that contribute to this growing body of work are often conducted in contexts that are framed as "successful." In fact, Rochelle

Gutiérrez (2012) called for the study of "successful learning environments for students who have been marginalized by society" to (1) provide "existence proofs . . . that these environments and their associated student outcomes can be created" and (2) inform the field regarding how to "build more such contexts for learning" (p. 17).

Although, as a field, we are learning more and more about practices that support such environments, this knowledge has not yet been reflected in many of the observational tools, rubrics, and protocols used to study these environments. Thus, as work in the field progresses and researchers continue to respond to Gutiérrez's call, there is a growing need to develop empirically grounded ways of attending to the extent to which the practices that are being outlined in research literature actually contribute to the "success" of these learning environments. In other words, in order to support the work of providing "existence proofs" and support the development of more of these learning environments, we need to be able to directly connect the success of these environments to the practices observed in these environments using qualitative and quantitative research methods. In fact, in regard to achievement as *one* measure of success, Joseph Allen and colleagues (2013) stated that "despite the strong theoretical interest in identifying qualities of teacher–student interactions linked to student achievement, scientific evidence is quite sparse regarding our capacity to identify and observe the critical features of teacher–student interactions that actually predict student learning" (p. 79).

The purpose of this article is to explore one way of meeting this need by describing the complex work of developing a set of classroom observation rubrics designed to support efforts in identifying and observing critical features of successful mathematics learning environments. Specifically, the rubrics have the potential to further flesh out some of the instructional practices that have been identified in research literature for their great potential to support improved learning opportunities for students who are often marginalized. The rubrics also have the potential to contribute to the development of more such successful learning environments by highlighting features that appear to distinguish more and less effective examples of these practices. In addition, the rubrics have the potential to support the work of attending to the extent to which these practices actually support and improve learning opportunities for students. In what follows, I outline the definition of success that guided the development of these rubrics. I also review some of the classroom observation rubrics that are used in the field and comment on the effectiveness and efficiency with which they could be used in attending to practices that characterize "successful" classrooms. I then share the process I used in developing rubrics that attend to such practices as well as some of the key distinctions between more and less effective examples of a few of these practices. I end by discussing some of the tensions, limitations, and implications of this work.

Defining "Success": Learning Environments Organized Around Conceptually Oriented Activity *and* Characterized as Aiming for Equity

In order to identify, examine, and attend to the success of mathematics learning environments, it is important to first take a stance on what "success" looks like.

Mathematical "*Standards*" like the NCTM and Common Core Standards (NCTM, 2000; CCSSM, 2010), have supported the field of mathematics education in developing *one* image of "success." Specifically, both sets of *Standards* describe a concrete set of learning goals for students that includes the development of conceptual understanding, productive problem-solving capabilities and dispositions, as well as procedural fluency in a range of mathematical domains. Instruction that supports students in attaining these goals typically provides time for students to engage in disciplinary practices of mathematics. In other words, this instruction often incorporates opportunities to explore challenging or cognitively demanding tasks, supports connections between multiple mathematical representations, elicits students' thinking, promotes participation in mathematical argumentation, and encourages students to provide justifications. This type of instruction has been called reform-, inquiry-, or conceptually-oriented and has been contrasted with more "traditional" forms of instruction (e.g., direct teaching of mathematical procedures) (Parks, 2010).

The image of success outlined within the *Standards*, though widely used and referenced in the field of mathematics education, has been critiqued for being incomplete, especially in light of supporting groups of students who have historically been marginalized in mathematics classrooms and within the field of mathematics in general. For example, the goals of supporting students in developing enduring, conceptual understandings of key mathematical ideas and engaging in disciplinary practices of mathematics outlined in the *Standards* do not yet explicitly include the "critical mathematical literacy" that students need in order to understand and change systems designed to privilege some and oppress others (Martin, 2015).[1] One reason such an omission is problematic is because, as Martin and McGee (2009) argued, mathematics that does not support critical mathematical literacy is irrelevant to specific groups of students, like African American students.

In addition, other than advocating for instruction in which "all" students are supported to substantially participate in mathematically rigorous activity, the image of success supported by the *Standards* does not directly or substantially attend to the extent to which a learning environment is equitable. For example, as far back as the year 2000, the NCTM *Standards* (2000) stated that "excellence in mathematics education requires equity—high expectations and strong support for all students" (p. 12). However, research indicates that in the United States, mathematics instruction remains inequitable and specifically points to students of color and students for whom English is not their first language as groups that are often poorly supported to learn mathematics (Gutstein & Peterson, 2005; Ladson-Billings, 1997; Martin, 2009; Martin et al., 2017; Nasir & Cobb, 2002; Nasir et al., 2009; Tate, 1995, 2008). As Martin (2015) pointed out, these inequities continue to occur notwithstanding the recommendations of the NCTM and "on their institutional watch" (p. 19). "High expectations" and "strong support," as delineated in the *Standards*, are not meaningfully unpacked or concretely defined. Thus, the principles of equity outlined in the *Standards* have been inadequate in supporting equitable learning environments for students and are incomplete in supporting the development of a clear image of instruction that is equitable.

Rochelle Gutiérrez (2012) contributed to developing a clearer image of equitable instruction by providing four dimensions for attending to the extent to which learning environments might be characterized as aiming for equity: access, achievement, identity, and power. *Access* relates to resources (e.g., high-quality teachers, a rigorous curriculum, and an environment that invites participation) that students have available to them in mathematics. *Achievement* relates to both standardized test scores and "participation" in mathematics (which includes contributions in a given mathematics class, mathematics course-taking patterns, and participation in the "math pipeline"). *Identity* concerns issues like the extent to which students have opportunities to see themselves in the curriculum, use mathematics to make sense of the world, find mathematics "meaningful to their lives," and draw from their own cultural and linguistic resources (Gutiérrez, 2012, p. 20). *Power* entails using mathematics to identify and work against inequalities that exist in local and broader contexts. It includes attending to (1) which voices are privileged in the classroom, (2) opportunities to use mathematics as "an analytic tool to critique society," and (3) the extent to which "alternative notions of knowledge" are valued (p. 20). Note, substantial attention to these and other issues of identity and power could support mathematics educators in addressing concerns raised when only the *Standards* are considered in developing an image of success (e.g., concerns around critical mathematical literacy and the relevance of the mathematics in which students are asked to engage). In other words, the delineation of these dimensions (particularly the dimensions of identity and power) better define and unpack the "strong support" necessary for equity and excellence that were outlined in the *Standards*. Access and achievement comprise what Gutiérrez (2012) called the "dominant axis" of equity because it consists of "the components students will need to be able to show mastery of in the discipline as it is currently defined" and "measures how well students can play the game called mathematics" (p. 20). Identity and power make up what Gutiérrez (2012) called the "critical axis" of equity because critical mathematics attends to and builds mathematics around students' cultural identities in ways that highlight the perspectives of marginalized groups and address social and political issues in society (Gutiérrez, 2007).

Gutiérrez (2012) challenged researchers concerned with issues of equity to think about the nature of the contexts in which they do research in light of these four dimensions. Therefore, as a researcher who is concerned with issues of equity and who acknowledges the importance of attending to the quality and the nature of the goals of the mathematical activity in which students are being asked to engage, when framing a learning environment as "successful," I am looking for evidence of both; that is, I am looking for evidence that students are being supported to develop along Gutierrez's four dimensions or the *Axes of Equity* while maintaining a focus on students being supported to participate in conceptually oriented mathematics.

It is important to note that Gutierrez's axes likely presume a valuing of conceptually oriented goals for students. Here, I am proposing a synthesis as a way of explicitly valuing and focusing on dimensions that support equity *and* conceptually oriented learning simultaneously. As an example of this type of synthesis, in previous work, I and others shared findings from an analysis in which we identified several forms of practice that appeared to distinguish two sets of classrooms: (1) classrooms in which there was evidence of conceptually oriented instruction and in which African American students performed better than predicted by their previous state assessment scores and (2) classrooms in which there was evidence of conceptually oriented instruction but in which African American students did *not* perform better than predicted on previous state assessment scores (Wilson et al., 2019). In this analysis, what characterized the set of classrooms that was ultimately framed as the more successful set was not only the evidence of conceptually oriented instruction, but also the evidence of instructional practices that support students' developing identities as well as their access to and achievement in mathematics.

When looking for evidence that students are being supported in their development and learning using instructional practices that characterize successful classrooms, it is also important to note that I am looking for potential and progress, not necessarily perfection. This is because I agree with Gutierrez's argument that for any given situation, it may not be appropriate or even possible for all of the components of this framing of success to be equally or fully present and that, at times, some of the dimensions may need to "temporarily shift to the background" (Gutierrez, 2012, p. 21). In addition, the work of identifying classrooms as successful would have to include things like examining student learning of mathematics that goes beyond achievement on assessments, investigating the extent to which students' mathematics learning impacts the developing of their many identities, following students to determine whether they select or have access to math-based majors and careers, and determining the extent to which mathematics incorporated students' resources and references in ways that helped to empower and "build critical citizens." In other words, classifying environments as successful involves aspects of students' lives, learning, and development that go beyond a given moment or instance and beyond a given class or even a given school year. Therefore, this framework is being used in this paper to identify classrooms as having *potential* for "success."

Although the *Standards* and the *Axes of Equity* both support images of ideal learning environments and, combined, contribute to the development of a vision of instruction that supports mathematics classrooms with potential for success, there are still important questions left for researchers and educators to investigate in studying the instruction of successful mathematics learning environments. For example, what do these images of instruction actually look like in practice, and what are some ways of recognizing aspects of the practices that support successful learning environments (i.e., How will we "know it when we see it"? Are there key distinctions that appear to matter in implementation? Is there a difference that makes a difference in terms of quantity or quality of the enactment of these practices?). In other words, these images of instruction do not provide the necessary guidance to identify, examine, or attend to aspects of the *Axes of Equity* and the *Standards* in practice. In addition, because teachers vary in the extent to which they provide instruction and in the extent to which they enact specific practices, researchers and educators also need a more nuanced way of thinking about the quality of the instruction offered in these environments than can be abstracted from the *Axes of Equity* and the *Standards* alone. The current paper describes a set of classroom observation rubrics that are designed to support the field in further developing an image of instructional practices that support successful learning environments and in identifying and examining these practices. Specifically, I describe key distinctions within the practices that are outlined in the rubrics.

Existing Classroom Observation Rubrics

In recent years, classroom observation rubrics have been used to evaluate the quality of lessons and to uncover key differences in teacher–student interactions that influence opportunities for student learning (Boston, 2012). In fact, several classroom observation rubrics have already been developed and are currently being used to assess the quality of mathematics instruction. However, many of these existing tools have limits in terms of the extent to which they can be used to measure instruction that appears to characterize potentially successful learning environments (as defined previously). For instance, a number of the existing measures do not focus specifically on mathematics learning and instruction. For example, the Classroom Learning Assessment Scoring System (CLASS and CLASS-Secondary) assesses the instructional and emotional supports provided in classrooms by examining specific dimensions of classroom interactions. Among the indicators used within this tool are aspects of teacher–student interactions that could support successful learning environments (e.g., indicators of positive climate [relationships, affect, respect, communication] and indicators of concept development [analysis/reasoning, creativity, integration]) (Pianta & Hamre, 2009). However, the CLASS was designed to capture broad interactional patterns and apply them "across diverse content areas at the secondary level" (Allen et al., 2013, p. 79) and thus would not be useful in attending to mathematics instruction that characterizes classrooms aiming for equity.

Other existing measures focus more specifically on mathematics but are not conducive to large-scale data analyses. For example, the Inside the Classroom (ITC) Observation and Analytic protocol (Horizon Research, 2003) is an observational measure that addresses some components of successful mathematics learning environments (e.g., rating the extent to which teacher questioning enhances the development of students' understanding and problem-solving; examining the extent to which there is evidence of a climate of respect for students' ideas, questions, and contributions; evaluating the extent to which "appropriate connections" are made to other disciplines and/or the real world). This measure uses an extensive protocol (Horizon Research, 2000) that requires a lot of time to gather the necessary data and demands much from researchers in analyzing many aspects of the classroom and making judgments about the lessons. Thus, the ITC has the potential to generate rich data; however, using this measure may be cumbersome for larger projects, and achieving and maintaining reliability may prove to be very difficult.

There are a few existing measures that capture *some* aspects of instruction that support successful mathematics learning environments but not others. For example, the Mathematical Quality of Instruction (MQI) rubrics are used to analyze five-minute segments of videotaped lessons. These rubrics "characterize the rigor and richness of the mathematics of the lesson, including the presence or absence of mathematical errors, mathematical explanation and justification, mathematical representation, and related observables" (Hill et al., 2008, p. 431). In many ways, the MQI examines some of the conceptually oriented instructional practices that could support successful learning environments. There are other measures that also focus on inquiry- or reform-oriented instruction (e.g., the Reformed Teaching Observation Protocol [RTOP], the Mathematics Scan [M-Scan]) (Piburn & Sawada, 2000; Walkowiak et al., 2014). Though these tools support researchers in defining, identifying, and measuring aspects of conceptually oriented instructional practices, they fall short in terms of illuminating practices that support an environment that might be characterized as aiming for equity; thus, when used alone, they are inadequate for examining instructional practices that characterize classrooms with potential for "success."

Building on this work around conceptually oriented instruction, I set out to develop rubrics that could be used to complement an existing measure that has been validated and is widely used in mathematics education research, the Instructional Quality Assessment (IQA) (Boston, 2012; Matsumura et al., 2008). The IQA is a measure that is based on the Mathematical Tasks Framework (Stein & Lane, 1996) and is consistent with the *Standards*. Specifically, the IQA is designed to measure the cognitive demand of a task as it appears in curricular materials (using the *Task Potential* rubric), the cognitive demand of a task as implemented (using the *Task Implementation* rubric), and the quality of whole-class discussion (using the *Academic Rigor of the Discussion* rubric as well as the *Participation*, *Teacher Linking*, *Student Linking*, *Teacher Asking*, and *Student Providing* rubrics) (Boston & Wolf, 2006; Matsumura et al., 2008). All of the IQA rubrics range from 0 to 4, where scores of 2 and below indicate a relatively low quality in terms of the task potential, task implementation, or rigor of the discussion, and scores of 3 and 4 indicate higher quality task potential, task implementation, and discussion or lessons in which there is strong evidence of conceptually oriented instruction. Notably, the IQA is not unlike the aforementioned measures in that it was designed without an explicit focus on practices that might characterize classrooms as aiming for equity. The focus on mathematics-specific instructional practices that support conceptual-oriented learning combined with the range of five different scores (0–4) that correspond to five different levels of enactment made the IQA an attractive measure to build on. Specifically, I saw the concreteness with which the IQA breaks down conceptually oriented practices into components and the specific way the IQA identifies key aspects of the practices in determining qualitative differences in terms of the levels of enactment as a great model for developing rubrics. In the following section, I describe the process of developing rubrics designed to complement the IQA and that more deliberately attend to instruction with the potential to support students in developing along the *Axes of Equity* while providing conceptually oriented opportunities for mathematics learning.

Developing Rubrics for Practices that Characterize Successful Classrooms

In developing rubrics that may complement the IQA, I decided to take findings from previous work as my starting place (Wilson, 2017; Wilson et al., 2013, 2019). More specifically, I started with practices that emerged as important when comparing conceptually oriented classrooms in which African American students' performance on state assessments appeared to improve to conceptually oriented classrooms in which this was not the case (Wilson et al., 2019). In the following section, I briefly review the process for identifying the practices that emerged in Wilson and colleagues' 2019 analysis and I outline the rubric development process.

Identifying the Practices

In previous work, I and others analyzed instructional practices that went beyond typical markers of conceptually oriented mathematics activity by conducting a comparative analysis of teaching in (1) eight lessons in which there was evidence of conceptually oriented instruction and in which African American students performed better than predicted by their previous state assessment scores and (2) 14 lessons in which there was evidence of conceptually oriented instruction but in which African American students did not perform better than predicted on previous state assessment scores. We used video-recorded lessons and student achievement data that were collected in years 1–3 (2007–2010) of an eight-year research project investigating instructional improvement in middle-grades mathematics teaching and learning in four school districts, located in three states in the United States (Cobb et al., 2018). Tables 1 and 2 present demographic data about the teachers and students from the four districts because teacher and student interactions were the focus of our analysis. (For more details and descriptions of the districts, see Cobb et al., 2018.)

We used the IQA to determine whether there was evidence of conceptually oriented instruction. We used student achievement data, based on standardized test scores, to identify teachers whose African American students showed, on average, better (and worse) growth on mathematics achievement tests than would be predicted by their performance in the previous two years (Wilson et al., 2019). In setting up our "selection criteria" this way, we attended to Gutiérrez's *dominant axis* of equity in our analysis. In other words, we attended to *access* (i.e., using the IQA as one way of distinguishing classrooms where there was evidence of "high-quality" teachers, rigorous curricula, and environments that invite student participation) *and achievement*

Table 1. Demographic Information for Participating Teachers.

District	Number of Teachers	Mean Years of Experience Teaching Math	% White	% Black	% Fully Certified
A	28	13.2	89.3%	3.6%	100%
B	26*	8.9	69.2%	19.2%	80.8%
C	28	9.2	24.0%	62.1%	93.1%
D	32*	8.7	84.4%	12.5%	87.5%

Note: * Indicates the number does not represent the full sample from the district. We do not have demographic information for 3 teachers (2 in District B and 1 in District D) who participated in the study.
Source: Boston & Wilhelm, 2015.

Table 2. Student Demographic Information for Districts A, B, C, and D.

District	Number of Students	% White	% Black	% Hispanic	% LEP	% Free/Reduced Price Lunch
A	35,000	30%	40%	15%	10%	65%
B	80,000	15%	25%	60%	30%	70%
C	160,000	15%	30%	65%	35%	85%
D	95,000	55%	35%	5%	5%	55%

Note: To protect the anonymity of the districts, the number of students is rounded to the nearest 5,000 and percentages are rounded to the nearest 5%.
Source: Boston & Wilhelm, 2015.

(i.e., using students' performance on standardized test scores and students' participation in mathematics lessons) within the contexts we analyzed.

Our qualitative analysis of video recordings from these two sets of classrooms focused on identifying forms of practice that were not already captured by the IQA and that appeared to support students along the *Axes of Equity* as they participated in conceptually oriented mathematical activity. We identified and coded seven forms of practice that appeared to distinguish the two sets of classrooms: Making Expectations Explicit, Coaching Students, Attending to Students' Local Context, Attending to Language, Attributing Mathematical Authority to Students, Positioning Students as Competent, and Attending to Classroom Community. In addition to coding for these practices, we applied subcodes to capture distinctions we noticed for four of the identified practices. For one of the practices (Positioning Students as Competent), the distinction focused on whether the practice was directed toward individual students or a group of students. For the other three practices (Attending to Students' Local Context, Making Expectations Explicit, and Coaching Students), the distinction centered on whether the practice was targeted specifically at supporting students' *mathematics* participation or whether it was targeted at supporting their participation in class more generally. For example, when coding for the practice of Attending to Students' Local Context, we noticed that sometimes the teacher was attending to students' local context in a way that directly connected aspects of students' lives to the mathematics task at hand. We also noticed that this connection often served the specific purpose of supporting the students' mathematics understanding and participation (e.g., using students' experiences with running down different hills in the schoolyard to support their understanding of the mathematics concept of slope). Other times, we noticed that the teacher was attending to students' local context to draw students in, engage them, and support student participation in class more generally (e.g., using a previous conversation about pizza being the favorite food for many of the students as a way to motivate them to solve a scenario about sharing pizza evenly among a group of fictional friends).

Given the selection criteria used in selecting the classrooms that we analyzed and the lens through which the qualitative analysis used to identify these practices was examined, these seven practices are directly connected to the *Axes of Equity*. For example, by enacting the practice of Attending to Students' Local Context, teachers from our analysis attended to concerns of *identity*. More specifically, using Gutiérrez's framework, this practice could be used to serve students by providing opportunities for students to see themselves in the curriculum, make meaningful connections to their lives, and draw from their own cultural resources.

In order to consistently apply these codes and subcodes, we needed to determine what counted as an "instance." In our analysis, an instance was an interaction involving a teacher with one or more students. The start and end of an interaction was signaled by a change in the topic of conversation and/or the teacher's physical movement through the class. These codes and decisions around instances (which are outlined more fully in the methods of Wilson et al., 2019) formed the base from which I developed the Equity and Access Rubrics for Mathematics Instruction (EAR-MI).

Developing the Equity and Access Rubrics for Mathematics Instruction

In designing the EAR-MI rubrics, I engaged in multiple cycles of iterative design (Cobb et al., 2003) (see Table 3). In the first cycle of design, I capitalized on the codebook created during Wilson and colleagues' 2019 analysis. Specifically, focusing on one practice at a time, I went through each of the 22 already coded lessons to examine every clip that was coded for each practice. In my reexamination of the lessons, I analyzed each clip and gauged the extent to which the clips could be categorized as "rich" examples of enactment for each practice. I based the distinction of rich and less rich examples on the extent to which each enactment appeared to (or at least showed strong potential to) support students in substantially participating in class, gaining access to and identifying with the mathematics activity at hand, and/or making sense of the discussions that took place. It is important to note that for the practices where there were subcodes, I paid particular attention to the extent to which the observed distinction that originally justified the creation of a subcode influenced the "richness" of the example. For instance, for the practice of Attending to Students' Local Context, the

Table 3. Cycles of Rubric Development.

Cycle	Number of Lessons/ Videos Analyzed	Development or Revisions Made	Who Was Consulted	Focus of Consultations
First cycle	22 lessons	• Identified clips • Categorized each clip for each practice (as rich or less rich examples) • Developed initial descriptions for these two categories for each practice based on trends observed • Developed more fleshed out definitions for the practices	• Graduate research assistant from the Middle-School Mathematics and the Institutional Setting of Teaching (MIST) project • Project manager/co-principal investigator (PI) for MIST • PI of MIST project	• Interrogated the language used in the categories of descriptions • Interrogated the language of practice definitions • Tested the clarity in the distinctions between the two categories
Second cycle	80 lessons	• Accounted for variety in the examples of implementation • Further fleshed out definitions • Better described distinctions between the rich and less rich examples • Developed additional distinctions within the two existing categories (yielding four categories) • Highlighted examples for explaining the practices and the distinctions	• The same three members of the MIST team: • Graduate research assistant from MIST • Project manager/co-PI for MIST • PI of MIST project	• Edited the definitions of the practices • Tested the clarity of the distinctions

(continued)

Table 3. (continued)

Cycle	Number of Lessons/ Videos Analyzed	Development or Revisions Made	Who Was Consulted	Focus of Consultations
Third cycle	90–100 lessons/videos 80 lessons previously analyzed plus videos of mathematics instruction provided by IQA developer (an estimate of 10–20 additional lessons/clips analyzed)	• Assigned each category a level between 1 and 4 • Created a description to capture what a level "zero" may look like for each practice • Checked that 0–4 scores reflected patterns of the IQA • Selected examples of the practices to include within the rubrics • Made edits and revisions based on feedback	• Experts in measuring quality of mathematics instruction (e.g., an original developer of the IQA) • Researchers on improving learning opportunities for marginalized and minoritized students (e.g., director of a Teaching English Language Learners program, a director of a center for urban education, and an editor of a journal for urban education) • The same three members of the MIST team	• Used the rubrics to examine lessons • Gauged the extent to which the rubrics represent practices outlined in literature • Assessed alignment between practices and the corresponding indicators for each rubric
Fourth cycle	95–110 lessons/videos An estimate of 5–10 additional lessons/video clips analyzed	• Re-examined gradations and language from rubrics (based on coding/testing) • Adjusted and refined the rubrics based on feedback from previous cycles	• Second author from original analysis • Researchers with expertise in the fields of mathematics education and issues of equity, diversity, and inclusion (e.g., professors who research, edit journals, and teach courses on race, identity, and learning in mathematics classrooms)	• Reviewed gradations of each rubric • Tested the clarity of rubric gradations • Used rubrics to assign codes for a variety of videos • Provided insights and feedback on the rubrics
Fifth cycle	95–110 lessons/videos Reused the same videos from previous work	• Adjusted and refined the rubrics according to the insights and the feedback from previous cycles	A new group of researchers with similar expertise (e.g., professors who research and teach courses on race, identity, and learning in mathematics classrooms)	• Provided suggestions about the practices in general • Careful and in-depth review of the language used within the rubrics • Provided suggestions/line-by-line edits of the text in the rubrics

example of a teacher connecting students' experiences with running down the hills of the schoolyard to the concept of slope would be considered a richer example of the practice compared to the example of a teacher using students' common love for pizza as a "hook" to engage students in the lesson. In this case, the richer example not only demonstrates knowledge of the students while taking advantage of an experience considered common among the students, but also more purposefully connects the mathematics that is being discussed in the lesson to the students' lived experiences. Thus, within the practice of Attending to Students' Local Context, the subcode indicated a key distinction in determining the richness of the implementation for the practice.

I officially categorized each clip for each practice as either a "rich" or "less rich" example of the practice. I then developed initial descriptions for these two categories for each practice based on trends observed and developed more fleshed-out definitions for the practices. I made note of the examples that for one reason or another seemed tricky to categorize. I then collaborated regularly with and received feedback and suggestions from three colleagues. In particular, we watched selected clips from the 22 lessons (starting with the "tricky" clips) to guide us in interrogating the language used in the descriptions of the categories and in the definitions of the practices. We also tested the clarity in the distinctions between the two categories.

In the second cycle of design, I used the updated definitions of the practices along with the descriptions of the "rich" and "less rich" categories to reexamine specific clips from the 22 lessons. Also, based on my well-developed knowledge of the full data from the eight-year project, I selectively supplemented the set of 22 lessons with additional lessons that I knew would include examples of enactment that would push on the definitions of the practices and the descriptions of the categories. In addition, I chose to add lessons that were not originally included in Wilson and colleagues' 2019 analysis because they did not meet one of our selection criteria; for example, our analysis deliberately focused on lessons from years 1–3 of the project in which there was evidence of conceptually oriented instruction and in which African American students performed better than predicted by their previous state assessment scores. I added lessons from years 1–3 in which there was evidence of conceptually oriented instruction and in which students who were identified as learning English as a second language performed better than predicted by their previous state assessment scores. (Note, this added 16 lessons to the set.) Adding these 16 lessons had the potential to expose me to other ways of enacting the practices and also the potential for me to find additional practices that support participation and achievement based on observations of teachers working with emergent multilinguals. For example, the speed at which teachers talked, the deliberateness with which they chose specific words to use when addressing the class, and the extent to which they paired their verbal instructions with gestures, visuals, and images were themes that came up in the original analysis but were emphasized and even more common in these classrooms. Some of what was observed in these 16 lessons contributed to the refinement of indicators within the rubrics (e.g., what it means and looks like to provide an "image" when explicitly stating expectations). Other aspects of instruction that were observed were noted. I planned to return and further investigate these aspects of instruction to ensure they were not already accounted for in another rubric (e.g., in an IQA or another EAR-MI rubric) and to see if a case could be made for developing additional rubrics to account for them. As another example, the Wilson and colleagues (2019) analysis deliberately focused on lessons in which there was evidence of conceptually oriented instruction by only including lessons that scored high on the IQA, so I expanded to include more "traditional" instruction by adding lessons that scored lower on the IQA. These additional lessons helped in further fleshing out some of the indicators for the lower scores for the rubrics (e.g., I had a better image of what explicitly stating social expectations looks like when students were expected to work independently). I added a total of 58 more lessons to the set of lessons that I viewed while developing the rubrics, bringing the set to a total of 80 lessons. The additional videos supported me in accounting for the possibility of more variety in the examples of implementation for the practices. In addition, they helped me in further fleshing out the definitions of the practices and in better describing the distinctions between the rich and less rich examples.

During this cycle, I also focused on additional distinctions that emerged within the two existing categories, concentrating on aspects of enactment that made examples within the less rich category seem weak or strong and doing the same for the rich category. I continued iteratively consulting with the same three colleagues, editing the definitions of the practices and testing the clarity of the developing distinctions. By the end of this round, I had four categories for the levels of enactment for each practice that centered around this new distinction of strong versus weak within the previously established rich versus less rich categories. I also began highlighting examples that appeared to be the most useful in explaining the practices and these distinctions.

In the third cycle of design, because the rubrics were deliberately developed to complement the IQA, I tested the four categories to see if the qualitative differences in the enactment of the practices captured within the categories mirrored different levels or gradations incorporated into the IQA. In other words, I officially assigned each category a level of 1–4 and created a description to capture what a "zero" for each practice may look like in implementation. I then checked that these 0–4 scores reflected the general patterns of the 0–4 scores within the IQA—namely, that scores of 3 or 4 indicate "more effective" examples of enactment for a given practice; scores of 1 or 2 indicate relatively "less effective" examples; and a zero indicates that there was little to no attempt to implement the practice during the coded lesson.

During this cycle, when consulting with my colleagues, we also selected examples of the practices from the set of 80 lessons to include within the rubrics. In addition to my three colleagues, during this cycle, I consulted regularly with experts in measuring quality of mathematics instruction. Just as I did with my colleagues, we watched videos from the original project. We also added videos of mathematics instruction provided by a co-developer of the original IQA rubrics to test and further develop the EAR-MI rubrics. We used the rubrics while examining these videos to gauge the extent to which the set of rubrics collectively represented practices that support historically marginalized students and to assess the alignment between each practice and the corresponding indicators on the given rubric. I also received suggestions and feedback from researchers with expertise in teaching and examining learning opportunities of students who have been historically marginalized in classroom interactions. I revised and edited the rubrics based on their feedback.

The fourth and fifth cycles involved similar iterations of design and revisions. For the fourth cycle, I consulted with researchers with expertise in the fields of mathematics education and issues of equity, diversity, and inclusion. I also invited the second author of the original analysis to give her insights and feedback on the rubrics. We reviewed the gradations of each rubric, tested out the clarity of the gradations by using the rubrics to assign codes to a variety of videos, and re-examined where the gradations were set and the language used to describe them. I then adjusted and refined the rubrics according to the insights uncovered and the feedback I received. The fifth cycle mirrored the fourth cycle except I consulted with a new group of researchers and they did a more careful and in-depth review of the language used within the rubrics. Specifically, they gave suggestions about the practices themselves and provided line-by-line edits on the written descriptions of each practice, as well as the text included in the distinctions within and across the levels of the rubrics for the various practices.

The Practices

A large part of the process of developing the rubrics involved taking the larger themes and codes that were observed in the original analysis of 22 lessons and fleshing out the practices into more detailed and observable concrete teacher moves. Also, as described previously, this process involved a lot of revisions and alterations to ensure that people, other than the two coders that were involved in the original analysis, understand the practices, can identify the practices (i.e., know it when they see it), and can accurately use the rubrics to categorize the different levels of enactment for a given practice. Thus, the descriptions of the practices were revised since the original analysis. In Table 4, I name and briefly describe the practices. (See Wilson et al. [2019] for detailed examples and how the practices connect with the dimensions of equity.)

As outlined previously, after careful examinations of various examples, consultations with experts, and many rounds of revisions, the original seven forms of practice were developed into 11 rubrics. Part of the rubric development process was creating gradations to help distinguish between more and less effective examples of the practices based on their potential to support students. In what follows, I share some of the qualitative differences that distinguished more and less effective examples of enactments that were used in creating these rubrics. For the sake of space, I focus on three specific rubrics as examples: the rubric focused on supporting connections and engagement between student context and the mathematics learning environment and the two rubrics focused on attending to language. To provide concrete images of the qualitative distinctions within each practice, I include vignettes created based on interactions observed within and across classroom video data taken from the MIST project. I also

Table 4. Name and Descriptions of Each Practice.

Name of Practice	Brief Description	Notes
Explicitly Stating Expectations	Teachers make the implicit and otherwise invisible or abstract class expectations explicit and more concrete (Kazemi & Stipek, 2001; Yackel & Cobb, 1996).	For this practice, two separate rubrics were created: one for explicitly stating social expectations for participation and one for explicitly stating mathematical expectations.
Coaching Students	Teachers support students in negotiating productive ways of participating and meeting expectations without decreasing the rigor of the task at hand by deliberately intervening, scaffolding, or providing additional supports (Staples, 2007).	As was the case with expectations, two separate rubrics were created to account for the distinction between social coaching and mathematical coaching.
Supporting Connections and Engagement Between Student Context and the Mathematics Learning Environment	Teachers connect students' lives to discussions and interactions that take place in mathematics class and support students in viewing issues, problem-solving contexts, and other scenarios discussed in a mathematics class as significant or real to either the students themselves or a broader audience (Banks & McGee, 2001; Gay, 2002; Jackson et al., 2013; Ladson-Billings, 1995).	
Attending to Language[2]	Teachers support all students in understanding the language used in the classroom (Battey et al., 2016; Moschkovich, 1999, 2002).	Two separate rubrics were created—one for cultural dialects and another for mathematical.
Attributing Responsibility to Students in Response to Their Requests for Assistance	Teachers support student agency to work through and solve mathematics problems by ascribing or "pushing back" responsibility to students (particularly when the students' questions get at the essence of the mathematics they are being asked to wrestle with). In so doing, teachers can (1) communicate that "the struggle is real," (2) normalize the feeling and convey that what students are wrestling with is common (i.e., it is not an individual or private struggle that is exclusive to them and them alone), and (3) reveal that what they are working through is intentional or mathematically difficult by design (Boaler & Staples, 2008; Cobb, 1995; Lampert, 1990; Staples, 2007).	
Positioning Students as Competent	Teachers frame students' actions and statements as intellectually valuable by explicitly and publicly identifying and acknowledging their actions and statements (Bartell, 2011; Battey et al., 2016; Hattie & Timperley, 2007; White et al., 2018).	This is not "appointing" or "giving" students competence—all students already have the capability and know-how to do important and brilliant things notwithstanding whether a teacher recognizes them or their brilliance.
Supporting a Nurturing Environment	Teachers generate dialogue, establish personal relationships, and develop a sense of community in the classroom (Timmons-Brown & Warner, 2016) as one way of developing a space in which students are more likely to feel comfortable taking the risks necessary in using complex, non-algorithmic thinking to create strategies for solving problems; making conjectures or forming generalizations; using other students' ideas while providing evidence for or justifying their own ideas; and drawing connections between, building on, or disagreeing with each other's ideas.	Two rubrics were created—one attending to the extent to which teachers build relationships and reinforce "classroom values" and one attending to how teachers respond to what may *appear* to be off task.

discuss how these distinctions may contribute to differences in learning opportunities made available for students.

Key Distinctions Between "More" and "Less" Effective Examples of Supporting Connection and Engagement Between Student Context and the Mathematics Learning Environment[3]

This practice focuses on teachers' attention to context and the extent to which they support connections between students' lives outside of the mathematics classroom and the discussions that occur within the classroom. Making these types of connections is suggested as one of the key aspects of culturally relevant pedagogy (Ladson-Billings, 1995) and multicultural education (Banks & McGee, 2001; Gay 2002).

In examining the video data for instances of teachers supporting connections and engagement between student context and the mathematics learning environment, I looked for examples of the teacher making explicit statements that signaled an awareness of or an attempt to connect to students' lives (e.g., school context, home or community context, or a reference to popular culture). For example, a teacher who is teaching at a school that is running a canned food drive to assist tornado victims in a neighboring state selects and edits a task that was originally about fundraising so that the task relates more directly to the philanthropic efforts being made in the school community. The teacher then facilitates a whole class discussion about the school food drive and connects the discussion to the scenario in the task. By altering the mathematics task to account for events going on in their school community, the teacher was making an effort to connect students' in-class and out-of-class experiences. Also, by facilitating a discussion about the task and how it relates to the food drive, the teacher made this connection explicit.

In coding for this practice, an aspect that appeared to make a difference—in terms of the supports provided for students—was the extent to which the references or connections to student life were made in service of furthering mathematical understanding (Jackson et al., 2013). Therefore, in developing the rubric designed to attend to this practice, the distinction between the more and less effective examples of implementation of the practice were mainly based on whether the teacher facilitated conversations about students' context that were relevant to the mathematics or the task that students were working on at the time. It is important to note that making connections to students' lives notwithstanding, mathematics and mathematical understanding is also important. In fact, many of these types of connections (that were made irrespective to mathematics or the task) were accounted for and marked in coding for the "supporting a nurturing environment" rubric that focuses explicitly on building relationships.

More Effective Examples: Attending to Students' Local Context in Service of Mathematics. In the more effective examples, teachers attended specifically to aspects of the problem-solving scenario that were directly related to the mathematics and the mathematical ideas that students needed to be able to understand in order to solve the problem. The richer examples of this practice usually took on the form of teachers facilitating discussions about the main task in ways that connected components of the students' lives to key aspects of the problem. For example, one teacher connected the fictional scenario of two brothers in a race (where the older brother wanted the younger brother to win in a close race) to a recent school field trip where the students biked around a lake. The teacher used the distance around the lake and the speed of a famous athlete to support the students' understanding of the need for a head start (Wilson et al., 2019).

In many of these enactments of attending to students' context, the teachers often took tasks and activities that could be "justified in terms of student learning opportunities as [an] initial point of reference," and then they identified "adjustments to the activities or additional supports that might enable particular groups of students to participate substantially" (Hodge & Cobb, 2019, p. 868). In other examples, the teachers capitalized on the interests, concerns, and cultural practices of students as a starting place when selecting and designing tasks and activities in ways that promote a positive view of students' cultures and practices and that affirm and confirm the communities in which they participate as potential resources for learning and instruction (Hodge & Cobb, 2019). For example, one teacher connected with her students' love for fashion and shopping by frequently creating tasks related to shopping dilemmas; the teacher-researcher described in Wilson and Hunt (2022) created problem-solving scenarios based on her students' love for snacks, aliens, and video games. Occasionally, there were "micro moments" where teachers contextualized mathematical concepts by using cultural references or things with which many students have experiences. For example, there was a teacher who noticed a student having difficulty understanding the concept of perimeter, so he used the student's familiarity with ants and suggested that the student imagine an ant starting at one point and walking around the shape that was drawn on the student's paper. These teachers used their students' interests and experiences as a resource while also supporting the students in being able to see themselves and their interests/experiences in the mathematical work that they all did together in class.

Less Effective/Non-examples: Attending to Students' Local Context Not in Service of Mathematics. The distinction between more and less effective enactments of this practice was primarily based on the extent to which there were substantial attempts to connect to students' home or community lives. Often, in the less effective examples, only the teacher participated in conversations that related to or brought in student contexts (e.g., there were lessons where the teacher talked at length about the context of the problem at hand within a task but did not stop to allow time for students to share their own personal experiences or connections to the task or problem-solving scenario). There were also frequent enactments where there was no relevance to the mathematical task at hand. Aguirre et al. (2013) use "meaningful connections" and "emergent connections" in describing a similar distinction between substantial and superficial attempts to connect students' cultural funds of knowledge and their multiple mathematical knowledge bases.

The less effective enactments often took the form of the teacher asking yes or no questions about something that they thought the students might relate to or be excited about, notwithstanding whether it was relevant to the mathematics being discussed. For example, while introducing a task that incorporates a scenario involving Spider Man, the teacher asks students whether they have seen the latest installment in the Marvel movies franchise without making any direct or significant connections to what the students would be learning or the problem they would be solving. Other common examples involve teachers making quick references or simply switching names or changing the contexts of word problems to appear to be more relevant for students. (See Aguirre et al. [2013] for descriptions of other examples of less effective attempts to connect students' context to the mathematics being discussed in class.)

Summary. Supporting connections between student context and the mathematics learning environment is an important practice because it supports efforts toward empowering students, honoring their identities, supporting their access to and participation in mathematics, and promoting their achievement (i.e., all four of Gutiérrez's dimensions). Teachers can support students in making sense of the world, in critiquing society, and in using mathematics to work against inequalities and injustices through this practice. In addition, teachers can honor and empower students as they appreciate their own histories, cultures, and communities while also providing opportunities for them to learn about the histories, cultures, and communities of others. This practice is also important in helping students to keep the complexities and nuances of "real life" connected to the mathematics they are learning rather than forcing students to detach or abstract the mathematics from the scenarios described in the problems they solve. Implementing this practice demonstrates that the teacher does not assume that their students perceive of a task as "real" simply because it is situated in a real-world context nor that their students have a familiarity with or an understanding of the given context (Boaler, 1993). Instead, by having discussions about the problems and mathematical concepts and then making direct connections to students' lives, teachers support students in making experientially real associations with problems they are charged with solving and the concepts they are charged with learning.

Supporting connections between student context and the mathematics learning environment is one way that teachers intentionally embed the mathematics content in socially meaningful contexts that matter to students (Herzig, 2005). Through this practice, teachers can support students in finding mathematics meaningful to their lives and in drawing from their own cultural resources. Thus, supporting connections between student context and the learning environment is one way of honoring and supporting student identity development. In addition, the experientially real associations could support students by providing mathematical images of the scenario (Jackson et al., 2013; Thompson, 1996) that have the potential to assist students as they use complex, non-algorithmic thinking to create strategies for solving the problem, as they make conjectures and form generalizations, and as they provide mathematical evidence or explanations to support their conclusions. Therefore, supporting connections between student context and the mathematics learning environment is key to supporting students as they participate in and learn from conceptually oriented mathematics activities in learning environments aiming for equity.

Key Distinctions Between "More" and "Less" Effective Examples of Attending to Language

In Wilson and colleagues (2019), attention to language included all instances in which the teacher made an explicit attempt to ensure that students understood the meaning of a word, phrase, or idea being shared. For this practice, two separate rubrics were created—one to account for the extent to which teachers attended to and included cultural dialects and one to account for the extent to which teachers attended to the mathematical language used in class.

Attending to Language and Including Cultural Dialects. By cultural dialects, I mean all languages other than the dominant or "White Mainstream English" (Baker-Bell, 2020). Examples include "African-American language" (Ladson-Billings, 1995, p. 482), "African American Vernacular English" (AAVE) or Black Language (BL) (Baker-Bell, 2020), cultural slang, and forms of informal English, as well as languages other than English (e.g., Spanish and Arabic). In looking for instances of teachers attending to and including cultural dialects, I searched the videos for teacher moves that would support students of diverse linguistic backgrounds, particularly as they shared their thinking with the teacher and with one another. For example, Moschkovich (2007) demonstrated ways in which teachers can support students who are learning English as a second language to use their home language as a resource for developing mathematical understanding. I looked for evidence of teachers encouraging students to think of words that sound similar to a specific word in a different language or using the roots of words as a resource. I also looked for whether students were affirmed in their ability to translanguage, or move with facility between languages (Ladson-Billings, 1995). In addition, the practice of revoicing has been connected with supporting students in bridging between informal language and mathematical language (e.g., Battey et al., 2016). Thus, I also looked for teachers using and revoicing contributions made in cultural dialects as well as linking these contributions to the contributions of others. It is important to note that in coding for this practice, I did not include instances where teachers interpreted or decoded student responses in patronizing or condescending ways that negated or censored students' language. Correcting or discouraging cultural dialects is arguably the exact opposite of the intention of this practice.

Attending to and including cultural dialects is about valuing student contributions, and teachers often demonstrated that they valued contributions by marking, positioning, "normalizing," and/or highlighting them as significant. In coding for this practice, I noticed differences in *how* and *when* teachers valued different languages and dialects. Specifically, I noticed a difference in the extent to which the teacher appeared to be open to "public" use of cultural dialects. In other words, I noticed differences in the extent to which the teacher demonstrated that these cultural dialects were "valued on an equal footing with [dominant American] English in the school context" (Bucholtz et al., 2017, p. 52). I also noticed a difference in the extent to which the teacher attended to whether other listening students understood what was shared in these different languages and dialects. Therefore, in developing the rubric designed to attend to this practice, the distinction between more effective and less effective implementations of the practice were mainly based on the extent to which there was evidence that the teacher publicly encouraged the use of cultural dialects in the classroom.

More effective examples of attending to and including cultural dialects. In the more effective examples of this practice, there was evidence that the teacher was open to the use of cultural dialects and consistently treated these different languages and dialects as important contributions to the class discourse. In addition, in the richer enactments, teachers often checked that the other listening students understood what was being shared. In these examples, the teacher encouraged other dialects and languages being shared in public as well as private contexts and was mindful of students who may have been trying to follow the discussion but may not fully comprehend. For example, in a class where students are working on the task about two fictional brothers participating in a race, a student used what Baker-Bell (2020) calls Black Language (BL) while sharing an important epiphany with a partner during group work time. The teacher, who was walking by, stopped and acknowledged what the student shared was an important idea and asked her to share what she discovered during the concluding whole class discussion:

Student: "Well when we were working, it hit me that if the race was that long, the older brother woulda been done gone and left him."

The teacher publicly acknowledged the epiphany as an important contribution and amplified and emphasized it by asking other students to revoice it:

Teacher: "Was everyone listening? She said something very important. Who can say that again in their own words?"

This contribution may not conform to dominant notions of grammatical correctness and thus may have been stigmatized in other classrooms as an indication of a linguistic deficit (Bucholtz et al., 2017). However, in this lesson, the teacher recognized that the student's use of "been done gone" was very purposeful in communicating important points about both distance and time in terms of the two fictional brothers and their participation in the race.

There were also examples of teachers who encouraged students to present their solution strategies to the class in Spanish and then had a groupmate translate what the presenting student said in English for the non-Spanish-speaking students who were trying to follow the conversation. In these examples, the teachers privileged the voices of their Spanish-speaking students while also providing opportunities for students to draw from their own cultural and linguistic resources. These teachers attended to their students' bilingual identities and sustained, honored, and valued their practice as "language brokers" (Bucholtz et al., 2017). In addition to examples of teachers acknowledging cultural dialects as equally important contributions to the class discussion, there were examples of teachers themselves switching in and out of cultural dialects (e.g., a teacher using Spanish while coaching a student to share his ideas turned to another student mid-sentence to ask, "¿Como se dice *table*?").

Less effective/non-examples of attending to and including cultural dialects. We observed an inconsistent or lack of openness to the use of cultural dialects and languages other than English in some classes. For instance, in a lesson where a student was using language and an algorithm from her home country to explain how she solved a problem, the teacher interrupted and said, "I know that's how you used to do it but you're in America now, you have to say it and do it the American way." In some classes, though there may have been evidence that the teacher was open to the use of cultural dialects, this use of diverse language was strictly restricted to one-on-one conversations with students or while students worked together in small groups. In other words, the teacher appeared to only allow other dialects and languages when used in relatively private contexts. For example, in a lesson where a teacher checks in on a pair of students working together during a think-pair-share and overhears one of the students using Spanish while explaining how she used slope to plot points and graph a line, the teacher nods and tells the students to stay partnered up and to continue working with one another. However, during the whole class discussion when the same student makes an observation in comparing the slopes of two graphed lines, she says, "they're not [the] same cada segundo" and the teacher simply moves on to the next raised hand without attending to or even really acknowledging this contribution. In this example, there is evidence that the teacher may view the use of Spanish or even "translanguaging"[4] (Bucholtz et al., 2017) as appropriate only when students are working privately. As another example, while a teacher is walking around the classroom during independent student work time, she is seen responding to students who are using Black Language (BL) and even engages them in conversation using BL herself. However, during whole class discussions, she often stops her students mid-sentence to edit their contributions (e.g., "You mean he walks faster than his brother. Not, 'He walk fast and his brother don't'."). In these examples, the teachers' responses to the more public use of non-dominant dialects and languages convey a deficit orientation. By ignoring, publicly correcting, or discouraging certain student contributions, teachers treat these ways of participating as unacceptable, inappropriate, or unimportant.

Summary. Attending to cultural dialects is critical to the work of empowering students, honoring their identities, supporting their access to and participation in mathematics, and promoting their achievement (Gutiérrez, 2012). Implementing this practice is one way teachers provide diverse opportunities for participating, support the valuing of alternative notions of knowledge, and promote various voices and perspectives opposed to privileging certain voices over others. Attending to cultural dialects demonstrates a level of understanding and respect for the linguistic diversity that is present within a given class (Franke et al., 2007). By encouraging students to freely communicate without focusing on using the "dominant" language of English or even having to use "White Mainstream English" (Baker-Bell, 2020) to explain or express what they are thinking, this practice also has potential to support students in embracing their developing identities rather than feeling pressured into abandoning them. In other words, attending to cultural dialects is one way that teachers can support students in "maintain[ing] their cultural integrity while succeeding academically" (Ladson-Billings, 1995, p. 476). It is important to remember that the most sophisticated enactments of this practice appeared to be instances where the teacher attended to the understanding of students who were expected to be listening as well. In making sure that other students understand what is being shared, the teacher can potentially support richer and deeper understandings of the mathematics being discussed, essentially by capitalizing on or maximizing the language that is being used as a resource. Also, in implementing this practice, teachers can demonstrate to students not only that it is important that there is space for them and that they feel comfortable sharing however they need to, but also that it is important that others understand them because the class values each student and what they have to share as a meaningful contribution to the learning community.

Attending to Mathematical Language

Mathematical language is language of the discipline that students learn and use while engaging and participating in mathematics (e.g., math words, phrases, symbols, syntax, etc.). It is not appropriate to assume that all students have the same lexicon. Also, because one way students express their understanding of content is through oral and written language, disparities in the extent to which different students understand the mathematical language being used in class could exacerbate other existing inequalities (e.g., the differences in access, comprehension, participation, and power that already exist). Thus, in addition to attending to cultural dialects and languages, attending to the mathematical language appeared to be an important practice in developing learning environments with potential for success.

In looking for instances of teachers attending to mathematical language, I searched the videos for indications that the teachers valued their students' understandings of mathematics vocabulary, symbols, and syntax. One way I observed teachers enacting this practice was by recognizing and highlighting aspects of words that could help students in understanding the meanings of specific vocabulary. Another way I observed teachers enacting the practice was by having students define and redefine mathematical terms or having them unpack the mathematical vocabulary that they use.

In coding for the practice of attending to mathematical language, I noticed differences in the extent to which teachers involved students in defining terms and unpacking the definitions of the mathematical language used in class. Specifically, I noticed a difference in the extent to which the teacher attended to whether other students, besides the original contributing student, understood the mathematical language used. Therefore, in developing the rubric designed to attend to this practice, the distinction between more and less effective examples of the practice were mainly based on the extent to which the teacher involved multiple students and how the teacher involved the students.

More effective examples of attending to mathematical language. In the richer enactments, there were multiple students involved in unpacking the mathematics language used. For example, in a lesson where students were asked to share their observations of data sets displayed in different statistical charts:

Student 1: Well looking at the dot plot, you have a bunch of numbers in the middle but that number all the way out there looks like an outlier.
Teacher: Oh so you think that may be an outlier. Interesting. Does everyone know what he means by that?
Student 2: It's a number that is in a set but is very different from the other numbers in the set.
Teacher: Okay, say more about . . . in case others do not understand. What does she mean by "very different from the other numbers"?
Student 3: Well it's basically a number that is really big or really small compared to the other numbers in the data set.

In this example, the teacher deliberately asked students to share what they knew about the mathematical term that was used (in this case, it was the word "outlier") and made sure to hear from more than one student before moving on. The teacher had students rephrase or reframe the definition as one way of attending to the understanding of other students.

In the more effective examples, we also saw the class working together toward developing a shared understanding of the words being used. For example, in a class where students were prompted to determine if the following table represented a linear relationship,

x	0	1	2	3	5
y	240	220	200	180	140

the teacher checked to make sure that students understood what *linear relationship* means:

Teacher: When I say a linear relationship, what do I mean?

Student 1: You mean does it go straight down [motions diagonally downward with a pencil in her right hand as if sketching the line in the air].

Teacher: Mmm hmm. So how are you going to know it goes straight down by looking at those numbers? 'Cause this is just a table of numbers. If you were looking at a graph, it would be going straight down [motions diagonally like S1 had done]. How are we going to be able to tell if those numbers are going straight down in the graph based on this table?

Student 1: How big or how small the numbers are.

Teacher: Okay, can you say more about that?

Student 1: Umm, you could kind of visualize where the numbers would go and then kind of see how much they're going up and down by.

Teacher: Okay so from 240 to 220, are the numbers going up or going down?

Choral: Down

Teacher: By how much?

Choral: 20

Teacher: And from 220 to 200?

Choral: 20

Teacher: And from 200 to 180?

Choral: 20

Teacher: And from 180 to 140?

Choral: 40

Teacher: Okay so is this linear?

Choral: Yes

Teacher: Say more about "Yes." In order for it to be linear, what has to happen?

Student 2: 'Cause the 4 is missing but, if it was there, it would have been 160. 'Cause the 4 is in there. But it's missing between the 3 and the 5.

Student 3: Yeah, there's a pattern but it skips 4.

Teacher: So there's a pattern on the time, where it goes 0, 1, 2, 3 and then it skips 4 and goes to 5 . . . so is it going down by 20 every time?

Choral: Yeah

Teacher: Except for the 180 and 140 and that's because they skipped the 4. So yes, this is a linear pattern—or well, the table represents a linear relationship.

In this example, the students and the teacher collaboratively made sense of the term and of the first student's way of explaining it. In addition, while working together to make sense of linear relationships, the teacher took up the students' ways of talking about the term. More specifically, she used the language ("straight down") and gestures (i.e., drawing an imaginary diagonal line in the air) that the first student used without editing her words. Another teacher may have made a correcting move (e.g., saying something like, "a line going straight down would be a vertical line"). In this interaction, the student's gestures were given just as much weight as her words and thus were taken up as a significant contribution to the class discussion. The teacher continued to use the first student's language in asking the class about the *y*-values from the table (". . . are the numbers going up or going down?" and "By how much?"). Then after confirming with the class that the table represented an example of a linear relationship (**Teacher:** . . . is this linear? **Choral:** Yes), the teacher prompted the students to add to their working definition (**Teacher:** Say more . . . In order for it to be linear, what has to happen?). Two other students contribute the idea of there needing to be a "pattern" even if numbers are "skipped." And the teacher once again used their words without editing them. In other words, in the richer instances of this practice, we see students actively participating and demonstrating understanding of key mathematical ideas while working with their teacher in marking and revoicing student contributions, pressing students to support their responses and explicitly state their reasonings/thinking, and connecting ideas to build coherence in the discussion as a way of establishing a shared understanding.

Less effective/non-examples of attending to mathematical language. In the lower quality enactments of the practice, only one person was involved in sharing the meaning of a mathematical term, and this one person was often the teacher. For example, in the case where the student used the word "outlier" in describing numbers included in a data set, a different teacher may have responded by saying, "I know some of you are wondering 'What is an outlier?' Well in this case an outlier is a number that lies outside the normal range of the numbers in the data set." In these examples, the teacher may recognize that there is a mathematical term shared that may not be understood by everyone and may acknowledge the need to define and/or unpack what was shared, but students generally are not involved in the defining and unpacking of the term. In another example, the teacher could take *one* student's demonstration of any level of understanding as an indication that there is no need for further discussion. For instance, in a lesson where students were asked to draw two triangles (based on a given description) that are not congruent to one another:

Student 1: [reads] "For each description below, you will draw two triangles that are not congruent. ....."
Student 2: Yeah, they will be the same shape but not the same size.
Teacher: Let's talk about congruent. What does congruent mean? Does anybody know?
Student 2: Two shapes that are the same exact size.
Teacher: Perfect! Exactly the same. So what our goal is—we don't want anything congruent. We don't want anything that's the same.

Here we see the teacher take Student 2's contributions as a signal that there was no need to unpack what congruent means. Notice that the combination of responses by Student 2 provides evidence that they understand what congruent means, but the articulated definition provided when explicitly prompted to share "what does congruent mean" was incomplete. Attending to the understanding of other students is important because students who are trying to follow the conversation may have missed or misinterpreted the first part of Student 2's contribution and could walk away thinking that any shapes that are the same size could be congruent.

Another less effective example of attending to mathematical language is reading a definition from the textbook and moving on with little to no discussion of what was just read (e.g., a teacher, reading from a textbook, stops and says, "So now they're [the textbook authors] defining unit rate for us. 'A unit rate is a rate in which one of the numbers being compared is one unit.' Here it is up here if you get lost" [points to the textbook definition written out on the front board]).

Summary. Attending to mathematical language use in a mathematics classroom is important because in implementing this practice well, the teacher acknowledges that there are different ways of explaining and understanding the mathematical terms that are used. The teacher also opens up a space for students to hear a variety of descriptions of the same term, which increases the likelihood that more students in the class will develop genuine understanding of the mathematics language compared to if only one definition was offered and immediately taken up as the "shared" definition. Thus, attending to mathematical language is another way to honor students' ways of knowing and explaining mathematical ideas, to empower students, to support their access to and participation in mathematics, and to promote their achievement in mathematics.

In general, attending to the language used in mathematics classrooms (whether attending to cultural dialects or attending to math language) appeared to be useful in supporting students in understanding each other's shared ideas during whole class and small group discussions. Knowing and appreciating what other students are saying is important if students are to make connections between, build on, and/or disagree with each other's ideas. It is also important if students are to make use of other students' ideas while providing evidence for or justifying their own ideas. Being able to comprehend other students' contributions could also potentially support students in expanding their problem-solving capabilities and increasing the solution strategies available to them. In addition, attending to language is essential for environments aiming for equity because it supports a more equitable distribution of intellectual authority and ability (Hand et al., 2015; Louie, 2017). Understanding what is shared during mathematics class is imperative for students to access and participate in discussions and the activities that take place in the classroom and beyond. Thus, attending to language is an essential aspect of supporting students as they participate in conceptually oriented mathematics activity in classrooms aiming for equity.

It's Complicated: Some Tensions and Limitations of This Work

Developing classroom observation rubrics is complicated work, especially when the instructional practices that the rubrics are designed to highlight are practices that support students who have typically been marginalized or underserved. One complication that I ran into while doing this work centered on the fact that designing these types of rubrics ultimately involves describing diverse teacher–student interactions—which are often multifaceted, layered, and nuanced—and framing them in specific and concrete ways that would make them easily and more immediately observable to others. This challenge is potentially further complicated by the fact that practices that aim for equity likely have unobservable dimensions—for example, the interpretations and the meanings that students and teachers make from particular interactions or how these interactions affect the ways they experience the learning environment over time. In particular, how the students themselves are experiencing the racialized space of the mathematics classroom is not currently reflected in the rubrics. As a next step in the development of the rubrics, I and others are currently working to get students' perspectives on these practices as well as other aspects of their experiences in mathematics classrooms. In addition, some components of each of the forms of practice that were observed in the original analysis were "left on the cutting room floor" and, at least for now, are not currently incorporated into the rubrics—for example, attending to the speed at which teachers give instructions (especially when given in English) and the extent to which teachers are deliberate in using gestures and hand motions alongside their verbal directions for the *attending to language* rubrics. There are several reasons certain components were not incorporated. For instance, it may have been difficult to outline tangible markers for specific aspects of the practices in ways that could support users of the rubrics in being able to reliably identify and code for them. In future research, it would be important to examine these aspects of practice separately by exclusively coding for them across various video data and investigating the extent to which there are themes within the coded clips of video that could support the work of naming them as specific practices and distinguishing between more and less effective enactments of these practices.

Although I fully acknowledge and recognize these and other challenges involved in concretely specifying the nuanced and complex practices of teaching, developing rubrics like the EAR-MI is still worthwhile work. Part of the work of developing such rubrics involves providing actual images and representations of the practices they are designed to capture. Also, this type of rubric development involves unpacking practices in ways that highlight the expertise needed in order to develop and implement them well. In addition, this work involves emphasizing aspects of different enactments of the practices that distinguish more powerful and more substantial enactments from others, particularly with regard to students' development and learning opportunities. All of these different components of rubric development contribute to what Grossman and colleagues (2009) called the "decomposing" of the practices of teaching; in other words, dissecting forms of practice into constituent parts in order to make potentially productive routines of action visible to others. This work is also necessary in order to support teachers and other practitioners in learning, developing, and implementing these productive instructional practices, which could ultimately contribute to the development of more learning environments with potential for success.

Another complication centers on the fact that it is not necessarily appropriate to implement all of these practices at all times. Teaching is complex and contingent work that involves many decisions based on things like how well the teacher knows their students and what the learning goals are for a particular lesson (Lampert & Graziani, 2009). In decomposing the practices and designing these rubrics, I am not advocating for "one-size-fits-all" instruction or a recipe for successful teaching. Rather, the practices outlined here are intended to serve as foundational suggestions for potentially productive starting places for teachers looking to improve their instruction. Ultimately, teachers need to get to know their students and use their own discernment in trying out these and other practices and in assessing and adjusting them according to the various dynamics at play within their own learning environments.

On a related note, because it may not be necessary to enact all of the practices at all times, characterizing a teacher's practice based on one or two observations may not be appropriate. As part of the process for validating these rubrics, I and others are currently working to answer questions about how many observations are needed and why as well as whether the number of necessary observations vary for different practices/rubrics. In addition, using rubrics like these to characterize teachers and/or their practice as "weak" is inappropriate. Using the rubrics to support teachers in thinking about how they could improve their instruction in a particular moment or in understanding ways they could elevate a particular interaction is more productive than labeling teachers or their practice as holistically weak or strong.

The institutional contexts and the policies in place in these contexts present other complications and limitations for this work. For example, though the video data used to develop the rubrics was collected from a larger project that deliberately selected contexts that in many ways could be described as "typical" (e.g., large public school districts facing the challenges of limited resources, high teacher turnover, large numbers of students identified as low-performing in mathematics, etc.), these contexts played a role in the kinds of practices that I was and was not able to observe. For instance, many of the teacher participants from the larger project taught in districts that had adopted and invested in the use of specific curricula. Thus, these teachers did not have much agency in terms of providing students with opportunities to make decisions about the curriculum used. In addition, the lack of observations of students using mathematics to interrogate power relations or to critique society may be attributable to (1) the curriculum frameworks, textbooks, and other resources provided to the teachers by the districts and policies around teachers using these materials and/or (2) larger systemic concerns (e.g., No Child Left Behind accountability measures that were being implemented at the time that data were being collected). It could also be that these empowering practices are currently not being used widely. In districts where there may be no pressure to keep up with pacing guides or to reach specific benchmarks while using required textbooks, teachers may create projects that provide space for students to collect data and find ways to "mathematize" student concerns or problematic policies (e.g., using mathematics to problematize office referral rates or to highlight relationships between race, gender, and the likelihood of students being expelled or suspended from school as a disciplinary action in response to specific behaviors). Therefore, though it is disappointing, it is not surprising that rubrics developed from observations within these contexts do not adequately attend to Gutiérrez's (2012) dimension of power.

The practices outlined in the original analysis and the distinctions and gradations included in the subsequent rubrics all derived directly from actual observations taken from real classrooms; thus, what was ultimately incorporated within the rubrics was directly dependent on what was actually observed. In other words, if we did not see it, we did not include it, so there are likely other practices or additional distinctions that are important in developing learning environments with potential for success that are not outlined here. Also, in looking within other contexts, these practices and distinctions may look different or take on different forms. Therefore, the practices and distinctions included in this paper are not intended to be all-inclusive in terms of practices that are implemented in learning environments with potential for success. Additional research is necessary to investigate other contexts—for example, examining the practices that emerge as important in learning environments that are identified as "successful" and that *do not* have an adopted curriculum.

In addition, these rubrics were developed based on findings from an analysis with a specific focus on African American students and that used achievement as one indicator of "success." Future research that focuses on other groups that have been marginalized or underserved may add new insights in terms of the practices captured in the rubrics as well as additional practices that are enacted in learning environments with potential for success. Also, additional research is necessary to investigate other indicators of success—for example, investigations that use the other dimensions of Gutierrez's *Axes of Equity* in selecting indicators of success (e.g., Smith & Wilson, 2019).

Discussion and Implications

Rubrics like the EAR-MI have the potential to support researchers in more accurately and concretely identifying and describing learning environments as "successful" by defining and outlining key distinctions of teaching practices implemented in classrooms characterized by conceptually oriented instruction and identified as aiming for equity. In addition, rubrics like the EAR-MI can support researchers in developing stronger evidence of the effectiveness of practices that prior research has identified as critical for marginalized students. For example, the EAR-MI has been designed to be used with large-scale data sets and could be used along with student achievement data to contribute quantitative and mixed methods analyses to growing efforts toward demonstrating the effectiveness of specific practices. This growing evidence could be even further strengthened if the EAR-MI is used with large-scale data and is considered along with measures that capture *other* aspects of successful learning environments; for example, identity and power have both been named as critical dimensions of classrooms aiming for equity. A lot of work has been done to develop surveys and interview protocols that can be used in revealing issues around identity and power, specifically as they relate to race and mathematics (e.g., Youth Survey of Race and Mathematics [English-Clarke, 2011], Multidimensional Inventory of Black Identity-Teen [Scottham et al., 2008]). When paired with these types of interviews and surveys, these rubrics have great potential to move the field toward developing stronger evidence of the effectiveness of practices that have been identified as anti-oppressive, culturally sustaining, culturally responsive, and culturally relevant.

Rubrics like the EAR-MI not only have the potential to influence the perspectives and the work of researchers, but also have the potential to influence the ways preservice teachers are prepared and the ways in-service teachers are trained. Specifically, using observation rubrics is one way to support teachers by providing shared lenses with which to view instruction and a common language to use in discussing their own practice. As Boston et al. (2015) state:

> Through the lens of a specific tool, teachers may be able to see aspects of instruction that previously blended into the myriad classroom activities occurring throughout a lesson. Once aspects of instruction are made visible, tools can provide a concrete structure for the development of new practices by specifying criteria and identifying standards for the implementation of the intended practice. Finally, tools can foster formative assessment and self-evaluation by focusing teachers' reactions on emerging or existing practices to identify strengths and/or motivate change. (p. 154)

The EAR-MI has four gradations for each practice built into the structure of the rubrics, so it is designed to generate data and results in a way that could be useful in supporting improvements in instructional practice. Thus, these rubrics could be a powerful tool for facilitating reflections and for providing feedback to teachers as they work toward developing specific practices. I and others are currently working to tailor the EAR-MI rubrics to better meet the needs of teachers and to support instructional coaches in using the rubrics in their work to support teachers (Litke et al., 2022).

Although I view this work as making a contribution, more work needs to be done around identifying successful learning environments, outlining practices that are implemented in these environments, and making direct connections between the practices observed and the success of the environments. One major implication that emerged from this work is the need for more work that focuses squarely on investigating issues of power and identity. The practices that the EAR-MI rubrics were developed to attend to were largely in service of providing students with access to dominant forms of engaging in conceptually oriented classroom mathematical activity. This is, in part, due to the nature of the data we had access to and could examine as part of the original analysis. Future research that incorporates more careful attention to specific students and their experiences as well as the classroom teachers' own perspectives would be necessary in order to contribute to the field's understanding of the critical axis of equity (Gutiérrez, 2012). For instance, a productive next step could be engaging in work that explicitly asks students themselves about instructional practices and teacher moves that *they* view as supportive in terms of their developing identities and in terms of empowering them to use mathematics to critique society.

Another implication of this work is the need to validate classroom observation rubrics and specify important criteria regarding the use of such instruments. As I stated earlier, rubrics like the EAR-MI have great potential to influence the ways researchers analyze and attend to the quality of classroom instruction as well as the ways that preservice and in-service teachers are trained. Unfortunately, observation rubrics also have great potential to be misused. For example, federal legislation has put states and districts under enormous pressure to improve teaching quality through evaluation (U.S. Department of Education, Office of Planning Evaluation and Policy Development, 2010). This pressure has resulted in the use of various observation protocols to make large-scale, high-stakes personnel decisions, creating an increased need for research on the validity of these protocols (Bell et al., 2012). In addition to identifying appropriate uses for the rubrics, validation studies also contribute to efforts toward determining the impact of the practices that the rubrics are intended to evaluate, thus adding to the empirical findings that directly connect the practices theorized to be important for students to concrete outcomes.

If there is any truth to the notion that we value what we measure or we measure what we value, then it is telling what we as a field of mathematics education researchers have and have not included in the tools we use to observe, examine, and assess the quality of classroom instruction. To date, most of the existing observation protocols and rubrics focus on only one dimension of successful learning environments—the extent to which they support conceptually oriented mathematics activity. If we value the more recent theories and findings that have emerged in terms of the instructional practices that are implemented in classrooms aiming for equity, then the rubrics and observational protocols that we use should be replaced, revised, or expanded to more adequately include these practices.

Acknowledgments

I would like to thank the participating teachers for allowing our project to learn from their practice. Additionally, I would like to thank Denice Kelley, Paul Cobb, Erin Henrick, H. Richard Milner, Robert Jimenez, Melissa Boston, Mahtab Nazemi, Marielle Myers, Sandra Crespo, Dorothy White, Toya Frank, Charles Munter, Maxine McKinney De Royston, and Kara Jackson for their constructive feedback on versions of the rubrics discussed in this paper and for their ongoing support.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the National Science Foundation under Grant Numbers DRL-1908481 and and DRL-1119122.

Notes

1. Some mathematics educators may argue that developing the practices of the discipline of mathematics is a necessary part of demonstrating critical mathematical literacy and thus, by supporting students in developing the disciplinary practices, educators are essentially supporting students in being able to demonstrate critical mathematical literacy. For example, the ability to understand and consume quantitative arguments and the ability to mathematize everyday phenomena are essential components to demonstrating critical mathematical literacy that are also practices of the discipline. However, developing critical literacy involves the development of agency and the development of cultural identities that could not be supported by *Standards*-based or conceptually oriented instruction alone.
2. Here, *language* refers to anything from written and spoken words (e.g., phrases, syntax, symbols, etc.) to body language (e.g., gestures)
3. Note, the term *context* refers to students' experiences and lives at school and home, as well as in their local community context and broader society. It is also important to note that by *mathematics learning environment* I am referring to all aspects of a mathematics class including the tasks, interactions, and discussions that occur in class, whether they directly relate to mathematics or not.
4. A linguistic practice; in this case, combining elements of English and Spanish, which is commonly referred to as "Spanglish."

References

Aguirre, J. M., Turner, E. E., Bartell, T. G., Kalinec-Craig, C., Foote, M. Q., Roth McDuffie, A., & Drake, C. (2013). Making connections in practice: How prospective elementary teachers connect to children's mathematical thinking and community funds of knowledge in mathematics instruction. *Journal of Teacher Education*, *64*(2), 178–192.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. *School Psychology Review*, *42*(1), 76–98.

Baker-Bell, A. (2020). Dismantling anti-black linguistic racism in English language arts classrooms: Toward an anti-racist black language pedagogy. *Theory Into Practice*, *59*(1), 8–21.

Banks, J. A., & McGee, C. A. (2001). Families and teachers working together for school improvement. *Multicultural Education: Issues and Perspectives*, *4*, 402–420.

Bartell, T. (2011). Caring, race, culture, and power: A research synthesis toward supporting mathematics teachers in caring with awareness. *Journal of Urban Mathematics Education*, *4*(1), 50–74.

Battey, D., Neal, R. A., Leyva, L., & Adams-Wiggins, K. (2016). The interconnectedness of relational and content dimensions of quality instruction: Supportive teacher–student relationships in urban elementary mathematics. *The Journal of Mathematical Behavior*, *42*, 1–19.

Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, *17*(2–3), 62–87.

Berry, R. Q. (2020, February). "How do we help teachers teach math to Black kids?" My response [NCTM Blog Post]. https://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Robert-Q_-Berry-III/%E2%80%9CHow-do-we-help-teachers-teach-math-to-Black-kids_%E2%80%9D-My-Response/

Boaler, J. (1993). The role of contexts in the mathematics classroom: Do they make mathematics more “real”? *For the Learning of Mathematics*, *13*(2), 12–17.

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record*, *110*(3), 608–645.

Bonner, E. P. (2010). *Unearthing culturally responsive mathematics teaching: The legacy of Gloria Jean Merriex*. University Press of America.

Boston, M. (2012). Assessing the quality of mathematics instruction. *Elementary School Journal*, *113*(1), 76–104.

Boston, M., Bostic, J., Lesseig, K., & Sherman, M. (2015). A comparison of mathematics classroom observation protocols. *Mathematics Teacher Educator*, *3*(2), 154–175.

Boston, M., & Wilhelm, A. G. (2015). Middle school mathematics instruction in instructionally focused urban districts. *Urban Education*, 1–33. doi: 10.1177/0042085915574528

Boston, M., & Wolf, M. (2006). *Assessing academic rigor in mathematics instruction: The development of the instructional quality assessment toolkit: CSE technical report 672*. University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Bucholtz, M., Casillas, D. I., & Lee, J. S. (2017). Language and culture as sustenance. In D. Paris & H. S. Alim (Eds.), *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world* (pp. 43–59). Teachers College Press.

Clark, L. M., Badertscher, E. M., & Napp, C. (2013). African American mathematics teachers as agents in their African American students’ mathematics identity formation. *Teachers College Record*, *115*(2), 1–36.

Cobb, P. (1995). Mathematical learning and small-group interaction: Four case studies. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 25–129). Lawrence Erlbaum Associates.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

Cobb, P., Jackson, K., Henrick, E., & Smith, T. M. (2018). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Harvard Education Press.

Common Core State Standards Initiative. (2010). Common Core State Standards for mathematics. http://www.corestandards.org/assets/CCSSI_Math Standards.pdf

English-Clarke, T. L. (2011). *Things my family told me about math: African American youths’ perception and use of racial and mathematical socialization messages* [Doctoral dissertation]. University of Pennsylvania.

Frank, T. J., Khalil, D., Scates, B., & Odoms, S. (2018). Listening to and learning with Black teachers of mathematics. In I. Goffney & R. Gutiérrez (Eds.), *NCTM annual perspectives in mathematics education: Rehumanizing mathematics for students who are Black, Indigenous, and/or Latin@/x* (pp. 147–158). National Council of Teachers of Mathematics.

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256). Information Age.

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education*, *53*(2), 106–116.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055–2100.

Gutiérrez, R. (2007). (Re)defining equity: The importance of a critical perspective. In N. S. Nasir & P. Cobb (Eds.), *Improving access to mathematics: Diversity and equity in the classroom* (pp. 37–50). Teachers College Press.

Gutiérrez, R. (2012). Context matters: How should we conceptualize equity in mathematics education? In B. Herbel-Eisenmann, J. Choppin, D. Wagner, & D. Pimm (Eds.), *Equity in discourse for mathematics education* (pp. 17–33). Springer.

Gutstein, E., & Peterson, B. (Eds.). (2005). *Rethinking mathematics: Teaching social justice by the numbers*. Rethinking Schools.

Hand, V., Kirtley, K., & Matassa, M. (2015). Narrowing participation gaps. *Mathematics Teacher*, *109*(4), 262–268.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Herzig, A. (2005). Goals for achieving diversity in mathematics classrooms. *Mathematics Teacher*, *99*(4), 253–259.

Hill, H., Blunk, M., Charalambous, C., Lewis, J., Phelps, G., Sleep, L., & Ball, D. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, *26*(4), 430–511.

Hodge, L., & Cobb, P. (2019). Two views of culture and their implications for mathematics teaching and learning. *Urban Education*, *54*(6), 860–884.

Horizon Research. (2000). *Inside the classroom observation and analytic protocol*. http://www.horizon-research.com/horizonresearchwp/wp-content/uploads/2013/04/cop1.pdf

Horizon Research. (May 2003). *Looking inside the classroom: A study of K–12 mathematics and science in the United States*. http://www.horizon-research.com/insidetheclassroom/reports/looking/complete.pdf

Id-Deen, L. (2018). "It will help me do better in mathematics": Building trusting relationships with black students. In A. Fernandes, S. Crespo, & M. Civil (Eds.), *Access and equity: Promoting high-quality mathematics in grades 6–8* (pp. 131–142). National Council of Teachers of Mathematics.

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education*, *44*(4), 646–682.

Jett, C. C., Stinson, D. W., & Williams, B. A. (2015). Communities for and with Black male students. *Mathematics Teacher*, *109*(4), 284–289.

Joseph, N., Hailu, M., & Matthews, J. (2019). Normalizing black girls' humanity in mathematics classrooms. *Harvard Educational Review*, *89*(1), 132–155.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *Elementary School Journal*, *102*(1), 59–80.

Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory into Practice*, *34*(3), 159–166.

Ladson-Billings, G. (1997). It doesn't add up: African American students' mathematics achievement. *Journal for Research in Mathematics Education*, *28*(6), 697–708.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27*(1), 29–63. doi:10.3102/00028312027001029

Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *The Elementary School Journal*, *109*(5), 491–509.

Litke, E., Wilson, J., & Hill, H. (2022). *Supporting teachers to develop equitable mathematics instruction through rubric-based coaching: Year one annual report*. University of Delaware, North Carolina State University, & Harvard University.

Louie, N. L. (2017). The culture of exclusion in mathematics education and its persistence in equity-oriented teaching. *Journal for Research in Mathematics Education*, *48*(5), 488–519.

Martin, D. B. (2009). Researching race in mathematics education. *Teachers College Record*, *111*(2), 295–338.

Martin, D. B. (2015). The collective black and “principles to actions.” *Journal of Urban Mathematics Education*, *8*(1), 17–23.

Martin, D. B., Anderson, C. R., & Shah, N. (2017). Race and mathematics education. In *Compendium for research in mathematics education* (pp. 607–636). National Council of Teachers of Mathematics.

Martin, D. B., & McGee, E. (2009). Mathematics literacy and liberation: Reframing mathematics education for African-American children. In B. Greer, S. Mukhopadhyay, A. B. Powell, & S. Nelson-Barber (Eds.), *Culturally responsive mathematics education* (pp. 207–238). Routledge.

Matsumura, L. C., Garnier, H. E., Slater, S. C., & Boston, M. D. (2008). Toward measuring instructional interactions “at-scale.” *Educational Assessment*, *13*(4), 267–300.

Moschkovich, J. (1999). Supporting the participation of English language learners in mathematical discussions. *For the Learning of Mathematics*, *19*(1), 11–19.

Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning*, *4*(2–3), 189–212. doi:10.1207/S15327833MTL04023_5

Moschkovich, J. (2007). Bilingual mathematics learners: How views of language, bilingual learners, and mathematical communication affect instruction. In N. S. Nasir & P. Cobb (Eds.), *Improving access to mathematics: Diversity and equity in the classroom* (pp. 89–117). Teachers College Press.

Nasir, N. S., Atukpawu, G., O’Connor, K., Davis, M., Wischnia, S., & Tsang, J. (2009). Wrestling with the legacy of stereotypes: Being African American in math class. In D. B. Martin (Ed.), *Mathematics teaching, learning, and liberation in the lives of Black children* (pp. 231–248). Routledge.

Nasir, N. S., & Cobb, P. (2002). Diversity, equity, and mathematical learning. *Mathematical Thinking and Learning, 4*(2–3), 91–102.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. NCTM.

Parks, A. N. (2010). Explicit versus implicit questioning: Inviting all children to think mathematically. *Teachers College Record*, *112*(7), 1871–1896.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109–119.

Piburn, M., & Sawada, D. (2000). *Reformed teaching observation protocol (RTOP) reference manual*. ACEPT Technical Report No. IN00-3. https://www.public.asu.edu/~anton1/AssessArticles/Assessments/Chemistry%20Assessments/RTOP%20Reference%20Manual.pdf

Scottham, K. M., Sellers, R. M., & Nguyên, H. X. (2008). A measure of racial identity in African American adolescents: The development of the Multidimensional Inventory of Black Identity—Teen. *Cultural Diversity and Ethnic Minority Psychology*, *14*(4), 297.

Smith, E., & Wilson, J. (2019). *What does attending to power in teaching look like? Poster presented at the 41st annual conference of the North American Chapter of the International Group for the Psychology of Mathematics Education*, *St. Louis, MO*.

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction*, *25*(2–3), 161–217.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, *2*(1), 50–80.

Tate, W. F. (1995). Returning to the root: A culturally relevant approach to mathematics pedagogy. *Theory into Practice*, *34*(3), 166–173.

Tate, W. F. (2008). The political economy of teacher quality in school mathematics: African American males, opportunity structures, politics, and method. *American Behavioral Scientist*, *51*(7), 953–971.

Thompson, P. W. (1996). Imagery and the development of mathematical reasoning. In L. P. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 267–283). Erlbaum.

Timmons-Brown, S., & Warner, C. (2016). Using a conference workshop setting to engage mathematics teachers in culturally relevant pedagogy. *Journal of Urban Mathematics Education*, *9*(1), 19–47.

U.S. Department of Education, Office of Planning Evaluation and Policy Development. (2010, March). ESEA: A blueprint for reform: The reauthorization of the Elementary and Secondary Education Act. https://files.eric.ed.gov/fulltext/ED508795.pdf

Walkowiak, T., Berry, R. Q., Meyer, J. P., Rimm-Kaufman, S. E., & McCracken, E. R. (2014). Introducing an observational measure of mathematics instructional quality: Evidence of validity and score reliability. *Educational Studies*, *85*, 109–128.

White, D. Y., DuCloux, K. K., Carreras-Jusino, Á. M., González, D. A., & Keels, K. (2016). Preparing preservice teachers for diverse mathematics classrooms through a cultural awareness unit. *Mathematics Teacher Educator*, *4*(2), 164–187.

White, D. Y., Gomez, C. N., Rushing, F., Hussain, N., Patel, K., & Pratt, J. (2018). Assembling the puzzle of mathematical strengths. *Mathematics Teaching in the Middle School*, *23*(5), 268–275.

Wilson, J. (2017). Identifying and assessing instructional practices that support historically marginalized students' engagement and participation in rigorous mathematics activities. *Paper presented at the 2017 NCTM conference*, *San Antonio, TX*.

Wilson, J., & Hunt, J. H. (2022). Marginalized within the margins: Supporting mathematics meaning making among students with learning disabilities. *The Journal of Mathematical Behavior*, *67*, 2002, 100982. https://doi.org/10.1016/j.jmathb.2022.100982

Wilson, J., Nazemi, M., & Jackson, K. (2013). Identifying forms of practice that support African American students' learning of mathematics: An investigation of classroom interactions. *Paper presented at the Annual Meeting of the American Educational Research Association*, *San Francisco, CA*

Wilson, J., Nazemi, M., Jackson, K., & Wilhelm, A. G. (2019). Investigating teaching in conceptually oriented mathematics classrooms characterized by African American student success. *Journal for Research in Mathematics Education*, *50*(4), 362–400.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458–477.

Author Biography

Jonee Wilson, PhD is an Assistant Professor of Mathematics Education in the School of Education and Human Development at the University of Virginia. Her research focuses on examining and outlining instructional practices that empower and honor historically marginalized students, specifically in the context of conceptually oriented mathematics classrooms. Dr. Wilson also focuses on professional development opportunities and other resources that support educators in learning, developing, and implementing such practices.