
Title	Measuring the Complexity of Teaching Practice for Equity: Development of a Scenario-format Scale
Authors	Wen-Chia C. Chang, Larry H. Ludlow, Lexie Grudnoff, Fiona Ell, Mavis Haigh, Mary Hill, Marilyn Cochran-Smith
Source	<i>Teaching and Teacher Education</i> , 82, 69-85
Published by	Elsevier

Copyright © 2019 Elsevier

This is an Accepted Manuscript of an article published by Elsevier in *Teaching and Teacher Education* on 08/03/2019, available online:

<https://doi.org/10.1016/j.tate.2019.03.004>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source.

Citation: Chang, W-C., Ludlow, L., Grudnoff, L., Ell, F., Haigh, M., Hill, M., & Cochran-Smith, M. (2019). Measuring the complexity of teaching practice for equity: Development of a scenario-format scale. *Teaching and Teacher Education*, 82, 69-85.

<https://doi.org/10.1016/j.tate.2019.03.004>

This document was archived with permission from the copyright holder.

Measuring the Complexity of Teaching Practice for Equity: Development of a Scenario-format Scale

Wen-Chia C. Chang ^{a, *}

Larry H. Ludlow ^b

Lexie Grudnoff ^c

Fiona Ell ^c

Mavis Haigh ^c

Mary Hill ^c

Marilyn Cochran-Smith ^b

^a National Institute of Education, Nanyang Technological University, 1 Nanyang Walk
Singapore 637616

^b Lynch School of Education, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA
02467, United States

^c Faculty of Education and Social Work, University of Auckland, Symonds Street, Auckland
1150, New Zealand

*Corresponding author

E-mail addresses: wenchia.chang@nie.edu.sg (W-C. Chang), ludlow@bc.edu (L. Ludlow),
l.grudnoff@auckland.ac.nz (L. Grudnoff), f.ell@auckland.ac.nz (F. Ell),
m.haigh@auckland.ac.nz (M. Haigh), mf.hill@auckland.ac.nz (M. Hill)
cochrans@bc.edu (M. Cochran-Smith)

Abstract: This study presents the development and implementation of a novel approach to measure enactment of equity-centered teaching practice. The construct of practice for equity is drawn from international syntheses and the details of the scenario-format scale development are emphasized. The result is a 15-item Teaching Equity Enactment Scenario (TEES) Scale that measures a meaningful continuum of levels of teaching practice for equity. The scale also provides a narrative interpretation of participants' levels of enacting equity-centered practice based on their simple raw scores. The implications for research on teaching and teacher education are discussed.

Keywords: Teacher education, teaching for equity, Rasch measurement, scenarios, vignettes, measurement instrument

Final draft before proofing

Measuring the Complexity of Teaching Practice for Equity: Development of a Scenario-format Scale

Introduction

Over the past two decades, there has been significant growth in teacher education programs designed to address the challenges posed by continuing educational inequality and inequity in the US and across the globe (Cochran-Smith et al., 2016). Preparing and supporting teachers to enact teaching practice that responds to diversity, challenges educational inequalities and inequities, and promotes fairness and justice is a pressing yet daunting and complex task. On the one hand, some studies have examined the resistance, emotional struggle, and non-engagement in teacher candidates as they wrestle with new ideas related to justice and/or as they strive to enact equity-centered teaching practice in school contexts (e.g., Achinstein & Ogawa, 2012; Buehler, Gere, Dallavis, & Haviland, 2009). On the other hand, some studies have documented the successes of programs designed to prepare teachers to teach with a commitment to challenging inequality and inequity by working closely with local communities (e.g., Kretchmar & Zeichner, 2016; Lampert & Burnett, 2015; Zygmunt & Clark, 2016) or through innovative and integrated program designs with justice goals (e.g., Grudnoff et al., 2017a; Quartz, 2003).

Terms such as “teaching/teacher education for social justice” and “teaching/teacher education for equity” have often been used in the literature in ways that are conceptually ambiguous or inadequately theorized (e.g., Arnold, Edwards, Hooley, & Williams, 2012; Cochran-Smith, Barnatt, Lahann, Shakman, & Terrell, 2009; Cochran-Smith, 2010; Grant & Agosto, 2008; McDonald & Zeichner, 2009; Sleeter, 2008, 2009; Zeichner, 1993). It is not within the scope of this article to provide an in-depth conceptual history or theoretical analysis of

these terms here, although we have done so previously in this journal and elsewhere (Cochran-Smith et al., 2009; Cochran-Smith, 2010; Cochran-Smith et al., 2016; Cochran-Smith et al., 2018). However, it is important to note that although these concepts are not the same, they are conceptually related to one another in important ways. Along these lines, it has been argued that social justice is not possible without equity, particularly what has been called “strong equity” (Cochran-Smith et al., 2018), which includes not only the redistribution of educational opportunities but also representation of all stakeholders in the development of what are taken to be common objectives and goals as well as recognition of the societal and educational systems that produce and reproduce inequity (Fraser, 2009).

More research is needed to understand how teachers’ characteristics and prior knowledge interact with the arrangements of teacher education programs, which then shape teacher candidates’ experiences in the programs and how these multiple interactions connect to their teaching practice (Zeichner, 2005). In particular, there is a need for the development of better measures to capture how and the extent to which new teachers come to understand ideas related to teaching for justice and/or how and to what extent they learn to enact equity-centered practice (Hollins & Torres Guzman, 2005; Zeichner, 2005). Measures of this kind could yield evidence to inform program improvement and/or build theories about how teacher candidates and novice teachers learn to teach.

The focus of this article is learning to teach for equity, specifically discussion of the development of a new instrument designed to measure complex teaching practice grounded in a theory of teaching for equity. The article describes the alignment between the items on the instrument and the conceptual and measurement frameworks, psychometric results of the instrument, and its utility and implications for teacher education stakeholders.

This study is part of the international Rethinking Initial Teacher Education (RITE) research program led by teacher education and measurement scholars in New Zealand and in the US. The work of Project RITE is informed by the theoretical framework that links complexity theory (CT) with critical realism (CR). As we have discussed elsewhere (Cochran-Smith, Ell, Ludlow, Grudnoff, & Aitken, 2014; Cochran-Smith et al., 2014), the principles of the CT-CR framework inform our views of teaching, learning, and learning to teach as non-linear, complex, and contingent causal processes, as well as our belief that teacher education research needs to capture this complexity. The overall goal of Project RITE is to develop an explanatory theory of teachers' learning during the critical period of initial teacher education that helps us understand the complex factors that influence whether, how and to what extent teacher candidates enact practice for equity and understand that part of their work as teachers is joining with others to challenge inequities (Cochran-Smith et al., 2014). Building on the previous work of Project RITE (e.g., Cochran-Smith et al., 2014; Cochran-Smith et al., 2016; Grudnoff et al., 2017b), the group has developed an instrument to capture the extent to which teachers enact equity-centered practice, understood as an intended outcome of teacher education programs.

In framing educational inequality, Project RITE rejects the equity-as-equality approach, which assumes that providing equal (that is, the "same") access to effective teachers, curriculum, and resources for marginalized students will produce equity (Cochran-Smith et al., 2016). This notion of equity, more or less consistent with distributive views of justice (Fraser, 2003; Fraser & Honneth, 2003), does not acknowledge the classroom, school, and societal systems and structures that perpetuate inequity (injustice). Nor does it adequately recognize the absence of the voices and epistemologies of historically disenfranchised groups in participating in the development of curriculum, educational policies, and practices (Cochran-Smith, 2010;

McDonald & Zeichner, 2009; Young, 2011). Rather, Project RITE takes a critical sociopolitical view of equity/inequity, “acknowledging that inequality is rooted in and sustained by much larger, long-standing, and systemic societal inequities in the first place” (Cochran-Smith et al., 2016, p. 69). This notion of equity, which combines redistribution, recognition, and representational views of justice (Fraser, 2009), suggests that equal distribution of resources and opportunities without attending to and challenging structural inequities and institutionalized oppression is insufficient in achieving equity (Cochran-Smith, 2010; Fraser & Honneth, 2003; McDonald & Zeichner, 2009; Young, 2011).

Our approach to conceptualizing equity as a complex and systemic societal issue informs the fundamental premises of the goals of teaching and teacher education. First, we assume that the purpose of teacher education is to prepare teachers who promote marginalized students’ learning while simultaneously recognizing and challenging systemic social and school inequities. This means that while we recognize the responsibility and moral purpose of equity-centered teacher education in preparing teachers who have the capacity to shape the knowledge, skills, and life chances of students, we also acknowledge that teachers alone cannot address societal inequities. Second, we assume that teaching and learning and the process of learning to teach are complex and non-linear processes (Cochran-Smith et al., 2016). Teaching is inherently value-laden and never detached from the socio-political context. This means that teaching practice includes beliefs, attitudes, habits, interpretive frames, and actions that combine in multifaceted ways to influence student learning.

Of course, our approach to conceptualizing teaching practice for equity is not the only way to do so. Our approach is informed by and roughly aligned with relevant studies and theories of teaching practice for marginalized students internationally. These include, for

example, the idea of culturally relevant pedagogy (Ladson-Billings, 1995, 2009), culturally responsive teaching (Gay, 2000; Villegas & Lucas, 2002), multicultural education (Sleeter & Grant, 1987), and teaching for social justice (Cochran-Smith, 2010; Grant & Agosto, 2008; McDonald & Zeichner, 2009; Sleeter, 2008, 2009; Zeichner, 1993) in the US. Other work along similar lines includes Gore's (2001) "productive pedagogy" that prepares teachers with critical stances about equity in Australia, and *Te Kotahitanga*, a professional development program intended to help teachers and schools support Maori students' learning in New Zealand (Bishop, Berryman, Cavanagh, & Teddy, 2009).

Like our approach to conceptualizing teaching practice for equity, all the educational theories noted above take a critical socio-historical view of equity (Cochran-Smith et al., 2016; Tan & Barton, 2012), acknowledging the systemic, institutionalized, and inequitable practices that produce and reproduce educational inequalities (Au, 2016; Ayres, Kumashiro, Meiners, Quinn, & Stovall, 2016). While these theories of teaching practice have different roots and emphases, they explicitly seek to improve the opportunities and outcomes of historically marginalized learners. In other words, they assume that it is part of the job of teachers and teacher educators to recognize and work with others to both eliminate the social structures and educational practices that perpetuate inequities and improve student learning outcomes. Further, these theories reject the notion that there is a set of discrete teaching practices that is uniformly effective regardless of students' backgrounds and the contexts of schools and classrooms. Rather, teachers' classroom actions involve their view of knowledge and knowledge construction and the beliefs and values through which they view their work, decide on their instructional approaches, support student learning, and form relationships with others (Grudnoff et al., 2017b).

Teaching Practice for Equity: Six Principles

To conceptualize teaching practice for equity, Project RITE sought to examine internationally published empirical evidence about teaching practice that contributes to student learning broadly defined, has a positive impact on diverse learners, and reflects a complex view of teaching. Using these criteria, we ultimately identified five international research programs and syntheses (Cochran-Smith et al., 2016; Grudnoff et al., 2017b). They include three New Zealand Best Evidence Syntheses (Aitken & Sinnema, 2008; Alton-Lee, 2003; Anthony & Walshaw, 2007), the Teaching and Learning Research Project [TLRP] in the UK (James & Pollard, 2006), the Measures of Effective Teaching in the US (MET Project, 2013), the Te Kotahitanga Effective Teaching Profile (Bishop et al., 2009), and the Center for Research on Education, Diversity, and Excellence's [CREDE] five standards for effective pedagogy (Dalton, 2007). We then conducted a content analysis to identify common practices that support all students' learning, especially the learning of those who have been historically marginalized, across all five syntheses (Grudnoff et al., 2017b). Table 1 presents the linkages between principles of practice drawn from the five international syntheses/research programs and the six interconnected principles (Grudnoff et al., 2017b).

That analysis yielded six principles of practice related to the construct of teaching practice for equity (Grudnoff et al., 2017b):

- 1) Selecting worthwhile content and designing and implementing learning opportunities aligned with valued outcomes: Characteristics include selecting content and setting learning goals and outcomes, designing and selecting learning opportunities, and implementing planned learning experiences aligned with valued outcomes;
- 2) Connecting to students as learners and their lives and experiences: Characteristics include identifying and recognizing students' home and community cultures, and making

connections between home and school cultures by planning and implementing linguistically and culturally sensitive curricula, instruction, and assessments;

3) Creating learning-focused, respectful and supportive learning environments:

Characteristics include fostering a caring, respectful, and inclusive learning environment, and orchestrating classroom procedures and physical space to facilitate collaborative learning;

4) Using evidence to scaffold learning and improve teaching: Characteristics include designing assessments and using evidence to improve instruction and scaffold learning;

5) Taking an inquiry stance for further professional engagement and learning:

Characteristics include taking an inquiry stance for individual and collective learning, and collaborating with involved stakeholders to advocate for supportive and sustainable learning environments and professional culture;

6) Recognizing and challenging classroom, school, and societal practices that reproduce inequity: Characteristics include recognizing and challenging one's own deficit thinking and theorizing on historically marginalized students, taking an agentic position in one's practice, and accepting professional commitment and responsibility.

Evidence from the selected international syntheses indicates that the six principles are *interconnected* (Grudnoff et al., 2017b). In other words, it is the enactment of multiple principles as a whole rather than the enactment of one single principle that enhances student learning. For example, “selecting worthwhile content and designing and implementing learning opportunities aligned to valued outcomes” (i.e., principle one) must rely on the understanding and recognition of students’ lives, cultural experiences, and prior knowledge (i.e., principle two). “Connecting to students as learners and their lives and experiences” (i.e., principle two) then

facilitates the “creation of learning-focused, respectful, and supportive learning environments” (i.e., principle three). The enactment of these three principles also requires the consistent use of assessments that scaffold learning and improve teaching (i.e., principle four), constant reflection, and a strong commitment to professional responsibility and continuous learning (i.e., principle five). Finally, the enactment of these five principles requires an asset-based framework to approach diversity brought by students, a critical view of knowledge and knowledge construction, and an agentic position to advocate for structural change (i.e., principle six).

Existing Relevant Instruments

Keeping the construct of practice for equity in mind, we reviewed 16 existing instruments designed to measure constructs related to teaching with a commitment to equity or teaching diverse learners or related concepts. Our review shows that the majority of existing instruments measure teachers’ attitudes (e.g., Baluch, Greig, Ponterotto, & Rivera, 1998), beliefs (e.g., Ludlow, Enterline, & Cochran-Smith, 2008), awareness (e.g., Henry, 1986), self-efficacy (e.g., Siwatu, 2007), and knowledge (e.g., D’Andrea, Daniels, & Noonan, 2003), rather than practice. While these constructs may be important predictors of teachers’ practice in the classroom, scholars have been calling for research in teacher education to measure teachers’ practice (e.g., Sleeter, 2001, 2014; Zeichner, 2005). One new instrument along these lines is the Framework for Equitable and Excellent Teaching (FEET), which is an observational protocol that captures teachers’ performances based on four dimensions of critical-oriented equitable and excellent teaching (Carmen Salazar, 2018). Generally, however, although some validated instruments that measure teachers’ practice in the classroom do exist, such as Danielson’s Framework for Teaching Evaluation Instrument (Danielson, 2014) and the Reformed Teaching

Observation Protocol [RTOP] (Piburn & Sawada, 2000), these instruments are not grounded in theories of teaching for equity or for social justice.

Additionally, a close examination of the theoretical foundations of existing instruments suggests that the political and critical aspects of teaching culturally and linguistically diverse learners are left out of most discussions (e.g., Baluch et al., 1998; D'Andrea et al., 2003; Siwatu, 2007). In other words, the advocacy aspect of teaching practice that recognizes and challenges structural inequalities perpetuated by social, school, and classroom practice on behalf of students is often diluted. In many of the instruments we reviewed, authors discussed some concept or concepts related to teaching for equity or teaching for social justice, such as culturally responsive teaching (Gay, 2000). These discussions generally described teaching practices that incorporate students' home culture as a conduit to facilitate the teaching-learning process, promote culturally compatible learning environments, use multiple approaches to assess student learning, and provide students with the knowledge and skills to function in mainstream society. However, we could find very few discussions in the instrument development literature, about teacher education programs that help students recognize and challenge social inequities, build students' capacity to investigate knowledge, and share decision-making power with communities as well as students in the classroom. In other studies, there is often inadequate theorization of terms such as "equitable teaching" or "multicultural competencies." When the theoretical foundation of an assessment does not reflect the political and critical aspects of teaching or when the construct is not clearly conceptualized, the items are also unlikely to capture such aspects.

Moreover, the majority of existing instruments apply Classical Test Theory to develop discrete, simple, and narrowly focused items that are intended to be highly correlated with one another to maximize traditional reliability estimates (Lord & Novick, 1968). This means that

each item only measures a fragment of a complex construct. In other words, if a construct such as teaching practice for equity comprises teachers' knowledge, interpretive frames, methods, and advocacy, then items should capture the complex nature of this construct. Although most instruments provide useful summary information based on mutually correlated discrete items, none of the scales we located offer a holistic and qualitative interpretation of what a person's summary score means in relation to the construct of interest. Altogether, this suggests the need for an innovative measurement approach that supports the development of a different type of scale to measure the extent to which teachers enact teaching practice for equity.

Previous studies have shown the success of developing instruments by applying Rasch measurement theory (Rasch, 1960/1980) alone or in combination with Guttman facet theory (Borg & Shye, 1995; Guttman & Greenbaum, 1998) to measure complex educational and psychological constructs. For example, some applied Rasch principles (Rasch, 1960/1980) to measure teaching with a social justice or equity lens (Ludlow et al., 2008; Carmen Salazar, 2018). Additionally, a recent study combined Rasch principles and Guttman's facet theory and sentence mapping technique (Borg & Shye, 1995) to develop a scenario-format scale to measure engagement in later-life activities (Ludlow et al., 2014; Ludlow, Matz-Costa, & Klein, 2018). The resulting scenario items provide a rich and meaningful description of how a person's score may be interpreted based on his/her level of engagement along a ladder-like continuum.

The Teaching Equity Enactment Scenario Scale

This article describes the development and implementation of the self-report Teaching Equity Enactment Scenario (TEES) Scale by applying Rasch measurement theory (Rasch, 1960/1980) and Guttman facet theory (Borg & Shye, 1995). The instrument is intended to be

formative, i.e., for assessing the extent to which teachers enact equity-centered teaching in the initial critical years of learning to teach.

Method

This study involved three phases. Phase one included the lengthy iterative process of clarifying the complex construct of equity-centered teaching practice and a pre-pilot of 15 scenario-format items with key informants from both New Zealand and the US. Phase two involved administering the pilot survey and a think-aloud exercise for item revision purposes. Phase three was the final calibration of the revised scale. Below we describe how the Rasch principles and premises of Guttman facet theory and sentence mapping techniques guided the procedures of constructing scenario-format items that define the TEES Scale.

Measurement Approach

Rasch measurement principles guided the specification of the construct, the content of the scenario-format items used to capture experiences related to the construct, and the psychometric modeling approach (Rasch, 1960/1980). We first theorized teaching practice for equity as a complex construct characterized by six interconnected principles of practice for equity, which function simultaneously to influence teachers' enactment of equity-centered practice. We also hypothesized that the construct of enacting equity-centered teaching practice could be usefully conceptualized as a ladder-like continuum ranging from lower to higher levels of enacting equity-centered practice (see Figure 1).

At the lower range of the continuum, practice reflects a technical view of teaching. Here the teacher is more or less unaware of the impact of his/her own practice and cultural positioning and instead tends to attribute educational success and underachievement to the merit or deficits of learners and their families. Teaching at the lower range of the continuum reflects high

expectations for only some students, and the learning opportunities offered are rarely relevant to students' lives, cultural experiences, or valued learning outcomes. Further, teaching at the lower range of the continuum does not integrate assessment into the instructional activities to help students learn. There is also little collaboration and shared responsibility among students and between students and the teacher. Moving along the continuum of equity-centered practice, at the higher range, practice reflects a more complex view of teaching wherein there is an explicit rejection of deficit thinking regarding historically marginalized students. There is also evidence of more critical self-reflection, shared decision-making with students, and advocacy for structural change on behalf of students. Practice that is equity-centered includes cognitively challenging learning goals for all students, clearly communicated, and deliberate use of a variety of sources, instructional materials, and pedagogical strategies. Practice at the higher end of the continuum functions to build collaborative learning communities in the classroom and integrate various assessment approaches into instructional practice to scaffold learning and improve teaching.

Practice at the lower range of the practice-for-equity continuum is based on a technical approach to teaching wherein knowledge and classroom practice are assumed to be neutral or value-free. In contrast, practice at the higher level of equity-centered practice involves flexibility, risk-taking, and teachers' efforts to be change agents who challenge the "status quo" and the hegemony of knowledge and pedagogical practice (Arnold et al., 2012). Developing equity-centered practice consistently and fully is difficult because to do so, teachers must deal with multiple and inevitable tensions, including the tension between the idea that there is a standard knowledge base that all students must know and master, on one hand, and acknowledgment of the fact that the standard knowledge base privileges some students while

marginalizing those who have been historically minoritized, on the other hand (Author, 2010). Achieving higher levels of equity-centered practice is often made even more difficult for teachers who must also deal with the constraints of conservative and stringent policies imposed on schools (e.g., high-stake test-based accountability policy) (Arnold et al., 2012; Cochran-Smith, 2010; Sleeter, 2009).

To measure this hypothesized hierarchical continuum, the instrument was constructed according to the following Rasch measurement principles: (a) the items address only the construct of enactment of teaching practice for equity (i.e., unidimensionality); (b) the items cover a wide range of levels from relatively easy to more difficult to enact (i.e., variation); (c) the variation in the items captures a meaningful progression from the lower to higher levels of enactment of practice for equity (hierarchical order); (d) the items should spread uniformly to form a ladder-like continuum over their hierarchical progression (i.e., continuum); (e) items have equal discrimination in terms of differentiating between higher and lower scoring persons; (f) the items possess local independence in the sense that an answer to one is not dependent upon the answer to another; and (g) the empirical data should confirm the hypothesized unidimensional, continuous and hierarchical item ordering as evidence of construct validity (Ludlow et al., 2014; Rasch 1960/1980). When the items have been well constructed according to these principles, and the data fit the measurement model, persons with a higher capacity of enacting equity-centered practice are expected to score higher on any item than lower capacity persons, and items capturing the lower levels of equity-centered practice are expected to be easier than higher level items for any participant to endorse.

Facet theory is an approach that systematically explores a content domain of interest

(Guttman & Greenbaum, 1998). Following the previous study by Ludlow and colleagues (2014), we employed a facet design procedure to develop the scenario-format items. A *facet* is defined as “a set of attributes that together represent underlying conceptual and semantic components within a content universe” (Guttman & Greenbaum, 1998, p. 17). Within facets, there are elements called *structs* that define the different levels that describe the variation within a facet (Guttman & Greenbaum, 1998). In our study, the six interconnected principles of practice for equity were regarded as six facets that together represent the content domain of equity-centered teaching practice. The variation within each facet was conceptualized as three levels of practice (i.e., structs) – low, moderate, and high.

We then employed a sentence mapping technique to construct the scenarios based on the facets and their levels (Borg & Shye, 1995). A mapping sentence consists of formal and informal elements: formal elements are drawn from facet descriptions, and informal elements comprise phrases linking the descriptions of facets together and provide the facets with a meaningful, substantive context (Guttman & Greenbaum, 1998). Combinations of multiple sentences then form each scenario, and each scenario becomes an item. The scenarios (i.e., the items comprising the scale) operationally define the construct/variable of teaching practice for equity. This combined measurement approach is consistent with the CT-CR framework of the Project RITE and allows us to capture the holistic, interactive and complex construct of enacting practice for equity as a “lived experience” whole rather than measuring each facet separately using narrowly focused Likert-like items.

Development of Scenario-format Items

In this section, we briefly describe four steps involved in the development of the scenario-format items. The complete procedures are thoroughly detailed in Chang (2017).

The first step required an in-depth understanding of the construct. With the six equity-centered teaching practice facets in mind, we categorized specific pedagogical practices associated with positive student learning outcomes drawn from each of the five international syntheses into one (and only one) facet. For practices categorized under each facet, we conducted an iterative process of content analysis, which led to the identification of two to three sub-topics (i.e., characteristics) within each facet. For example, Facet 1, *selecting worthwhile content and designing and implementing learning opportunities aligned to valued outcomes*, encompasses three characteristics: (a) selects content and sets learning goals and outcomes; (b) designs and selects learning opportunities; and (c) implements planned learning experiences aligned with valued outcomes. Based on these characteristics, we wrote a rich narrative to describe the practice associated with each facet. We then wrote specific narrative descriptions to capture each of the hypothesized three levels of enactment of practice (i.e., *low* – level 1, *moderate* – level 2, and *high* – level 3) for each facet.

The second step was to determine the structure of the scenarios. Given the six facets and three levels within each facet, 729 combinations of facets (i.e., 729 scenarios) were possible. We employed an “extreme group contrasts approach” for the initial stage of instrument development, meaning that we constructed extreme teaching practice scenarios from just the highest, then the moderate, and then the lowest levels of facet narratives (Ludlow et al., 2014). In addition, because it would be overwhelming for participants to read a scenario containing six facets, each scenario was constructed to encompass just three of the six facets.

Aligning with the theoretical framework of equity-centered teaching practice, we decided that Facet 6, *recognizing and challenging inequities*, would be present in all scenarios. We systematically selected the remainder of the five facets so that there was overlap between

scenarios. For instance, Scenario F126H included Facets 1, 2 and 6; Scenario F236H included Facets 2, 3, and 6; and Scenario F346H included Facets 3, 4, and 6 capturing high-level enactments. Fifteen scenarios with five scenarios capturing each of the three distinct levels of enactment of practice for equity were created (see Table 2). This approach created a chain of linkage and allowed the scenarios to cover the content of all six facets without overwhelming the respondents. We hypothesized that the calibrated scenarios (i.e., estimated difficulty levels) would loosely form three clusters along a hierarchical continuum of the levels of enactment with the five higher-level (“harder”) scenarios on the top, five in the middle, and five lower-level (“easier”) scenarios at the bottom.

The third step was to construct the scenarios. Given that there are five different configurations of scenarios (i.e., F126, F236, F346, F456, and F156), we used a different sentence mapping structure to construct the scenarios for each configuration but kept the same structure across the three levels of enactment within each combination. A total of five sentence mapping structures were developed. Each sentence mapping structure includes a formal part to be filled with descriptions of practice drawn from the specific facets employed in that scenario, and an informal part that remains constant to connect the formal parts and provide a brief context for the scenario. In addition, a person’s name was used in each scenario to enhance its authenticity (Table 3 presents an example of a sentence mapping structure and a scenario constructed from the sentence structure).

The fourth step was to decide on the response options and survey instructions. Participants were instructed to reflect on their practice, compare it against the practice of the individual teacher named in a scenario, and rate their own practice based on a five-point level of

enactment scale. Below we use the scenario capturing the high level of enactment of Facets 1, 5, and 6 (F156) as an example.

“Megan fully embraces her responsibility to identify and challenge classroom and school practices that promote inequities for students. Megan sets cognitively challenging goals and communicates to her students clearly and consistently. She purposefully draws upon a variety of sources to cultivate their conceptual understanding and encourages students to challenge information in textbooks. She deliberately uses various pedagogical strategies to capture students’ interests. Megan also works with others in a professional community to pose questions, reflect on her own assumptions, and proactively respond to student needs.”

Choosing *about the same* (a score of 3) means that respondents conceive of their own practice as similar to Megan’s practice; choosing *slightly lower* (a score of 2) or *much lower* (a 1) means that participants consider their practice to be at a lower enactment level than Megan’s practice; and choosing *slightly higher* (a score of 4) or *much higher* (a 5) means that respondents consider their practice to be at a higher enactment level than Megan’s practice. Higher scores indicate higher levels of enactment of equity-centered practice. It was intended that scenarios representing higher levels of enactment would be harder for respondents to reply with *slightly higher* or *much higher* ratings than scenarios representing lower levels of enactment.

To address face validity, in addition to the careful review by experienced teacher educators of Project RITE throughout the scenario development process, we conducted a pre-pilot and two feedback sessions with key informants who had expertise in either teacher education or measurement. The informants responded to the 15 scenarios on Qualtrics and provided feedback through a focus group session or in writing at the time that they responded to the scenarios. We identified three main problems, which were subsequently addressed. First, we

added a clearer introduction and a moderate-level scenario practice item to the beginning of the survey to address the potential “start-up effect” given the unconventional item format and comparative response task. Second, we revised the scenarios to be more engaging and story-like and to minimize confusion around potential multi-barreled interpretations within scenarios. The third issue concerned the effect of social desirability - especially for the lower-level scenarios. We subsequently revised negative wordings to reduce the influence of social desirability.

Data Collection and Participants

Pilot and final sample data were collected using a single questionnaire on Qualtrics. The questionnaire consisted of the 15-item scenario-format scale and five items asking for participants’ demographic information (i.e., gender and race) and teaching background (i.e., years of teaching experience and levels and subject matter of teaching).

The target participants were teachers and teacher candidates with classroom experience teaching academic subjects in K-12 settings. Convenience and snowball sampling were used for both the pilot and final sample. The pilot study was conducted over the course of two weeks, between October 31 and November 14, 2016. One hundred and twenty-one participants responded to the pilot questionnaire. We accepted 73 respondents with full response sets or only one missing item for the pilot analysis. The six participants who skipped one scenario completed the rest of the survey, including five items on demographics and teaching background, and thus were kept in the sample. No imputation method was necessary to replace the missing responses since the psychometric analyses employed a pairwise estimation algorithm (Linacre, 2016). Of the 73 participants, 23 were pre-service teachers in New Zealand and 50 participants were pre- or in-service teachers in the US. For the final calibration study, participants were pre-service and novice teachers (i.e., less than five years in the teaching profession) prepared in eleven different

university-based teacher education programs in the US. Data were collected between December 5 and December 20, 2016. Ninety-seven participants responded and 57 participants with complete response records or with only one missing item response (without replacement) were selected. Due to the public holidays in New Zealand and the timing of the semester in the US, we could not obtain a larger sample over the short data collection time frame for both studies.

The majority of participants (approximately 80%) for both the pilot and calibration analysis were White or European females, which represent the homogeneous teaching force and the “overwhelming presence of Whiteness” in teacher education indicated in the literature (Sleeter, 2001, 2017). In both analyses, the participants included pre- and in-service teachers who were prepared by programs in different higher education institutions. The participants also reported various years of teaching experience ranging from less than one year to more than ten years and had taught different subject areas including English language arts, social studies, mathematics, and science at the elementary through secondary levels.

Data Analyses

The Rasch rating scale model (Andrich, 1978; Wright & Masters, 1982) was used to test the extent to which the hypothesized hierarchical scenario scale structure was supported by the empirical data. In the statistical model (see Equation 1 below), the probability of person n responding in category x to scenario i (π_{nix}) is governed by the person’s enactment level (β_n) and the enactment difficulty of scenario i (δ_i). Each scenario has five response categories ($x = 5$), and τ_j is the threshold difficulty parameter of moving from one response category to the next higher one. In the Rasch rating scale model, the difficulty of moving from one step to the next (e.g., from choosing *about the same* to *slightly higher*) is presumed to be the same for all 15 items. Person and scenario estimates are reported in measurement units called *logits* (Ludlow &

Haley, 1995). High-scoring individuals have positive logits, while low-scoring individuals have negative logits. Scenarios that are harder for participants to respond with higher ratings have positive logits, and easier scenarios have negative logits.

$$\pi_{nix} = \frac{e^{\sum_{j=0}^X n_i [\beta_n - (\delta_i + \tau_j)]}}{\sum_{k=0}^m e^{\sum_{j=0}^k [\beta_n - (\delta_i + \tau_j)]}} \quad (\text{Equation 1})$$

The WINSTEPS software package was used for the analyses (Linacre, 2015, v3.91.2).

In the following section, we report psychometric results including the “variable map,” goodness-of-fit statistics, category characteristic curves, and separation and reliability estimates. A variable map is a graphic representation of the order and spread of calibrated items and persons based on the empirical data. This graphic illustrates the extent to which the scenarios define the hierarchical continuum of equity-centered practice from lower to higher levels of enactment as hypothesized. Fit statistics indicate the discrepancy between the expected and the actual responses of persons to the items (Wright & Masters, 1982). Unexpected responses often indicate a need for either scenario revisions or further investigations into respondents’ response patterns. We report the weighted (Infit-MS in WINSTEPS software) and unweighted (Outfit-MS in WINSTEPS software) mean squares of standardized residuals between the observed and the expected responses. The Infit-MS indicates inconsistent response patterns, while the Outfit-MS signals outliers (Wright & Masters, 1982). We adopted a liberal threshold of 1.2 ~1.3 for both unweighted and weighted fit statistics and the corresponding t statistics (i.e., $t > |\pm 2|$ or $t > |\pm 3|$ depending on the sample size) to identify misfitting scenarios and persons. Category characteristic curves graphically present the extent to which the five response options were used by the participants and whether the probability of moving from one category to the next followed the expected ordered pattern for all scenarios. Lastly, the item separation index indicates whether items are spread out enough to meaningfully define a difficulty progression along the

construct, while the person separation index indicates whether the sample of people is sufficiently spread out along the continuum.

Results

Pilot Study

The pilot study results (reported in detail in Chang, 2017) confirmed that the locations of the scenarios followed the hypothesized order, and a “proof-of-concept” for the measurement procedures was thereby established. In other words, the five scenarios (i.e., Scenarios F126L, F236L, F346L, F456L, and F156L) capturing the low-level of equity-centered practice were the easiest for participants to give “higher” ratings. The five moderate-level items (i.e., Scenarios F126M, F236M, F346M, F456M, and F156M) were all harder than the five low-level scenarios. The five high-level scenarios (i.e., Scenarios F126H, F236H, F346H, F456H, and F156H) were the hardest for participants to give “higher” ratings than the practices captured in the scenarios.

We also examined whether the locations of the scenarios were identical, or nearly so, between New Zealand and the US participants. Figure 2 presents the separate variable maps for the two subgroups. In the maps, the scenarios (on the right of the vertical line) are ordered from the easiest to enact at the bottom to the hardest to enact at the top. The participants (on the left of the vertical line) are ordered from the lowest scoring to the highest scoring. The “M” symbols shown on the left and right of the vertical line indicate the mean enactment levels of participants and scenario difficulty, respectively. As shown in Figure 2, the order of the scenarios for the subgroups is similar; that is, the scenarios capturing the lower-level of equity-centered teaching are the easiest to give higher ratings followed by the moderate-level scenarios, and the higher-level scenarios are the hardest to give higher ratings. The Pearson correlation between the respective pairs of item estimates was 0.965.

A Mantel-Haenzel differential item functioning (DIF) analysis was performed on each pair of item estimates for the US and NZ subgroups. Although the tests were non-significant in every instance ($p > 0.05$), three easier scenarios (i.e., F156L, F126L, F456M) displayed a relatively large DIF contrast of 0.82, 0.65, and 0.59, respectively. Specifically, Scenarios F156L and F456M were relatively more difficult for the US participants while F126L was more difficult for the NZ participants after controlling for participants' levels of enacting practice for equity. Nevertheless, the non-significance in these results suggests that the meaning of enacting from lower to higher levels of teaching practice for equity in the classroom was statistically invariant across the subgroups in these two different country contexts.

We then examined the combined sample goodness-of-fit statistics for scenarios and participants and investigated the unexpected response patterns. Two relatively easy and poorly fitting scenarios (i.e., F156L and F236L) were identified with unexpected responses from several higher-scoring participants who gave lower than expected responses. A think-aloud exercise with key informants who were teachers or teacher educators was then conducted. The think-aloud exercise involved having the informants respond to a select number of good and poorly fitting scenarios and a discussion to identify specific issues in the misfitting scenarios and ways to address them.

Altogether, we revised six scenarios for the final implementation. The think-aloud exercise identified specific words and phrases that caused confusion in Scenarios F156L and F236L; they were clarified. In addition, for the scale to be useful for differentiating individuals' teaching practices, the scenarios should be uniformly distributed along the continuum of difficulty estimates and extend to higher-level, harder to enact practice. Therefore, scenario F346H was revised to be easier by eliminating some positive adjectives that convey high-level

teaching practice, Scenarios F236M and F456H were modified to be harder by using words that convey even more positive teaching images, and Scenario F456L was modified to be easier by making the lower-level description clearer.

Final Implementation and Calibration

Variable map. The variable map from the calibration sample is presented in Figure 3. Figure 3 confirms that the order of the scenarios is consistent with both our theoretical definition of teaching practice for equity and the pilot sample results, and more importantly, reflects the intended scenario revision effects.

Regarding goodness-of-fit, we identified five potentially misfitting scenarios, among which four were lower-level scenarios with the two misfitting scenarios in the pilot study remaining problematic. We also identified five misfitting participants ($MS > 1.6$ and t statistic > 6.0) who gave unexpected responses (i.e., lower than expected responses to the low-level items and higher than expected responses to the high-level ones) to both misfitting and non-misfitting items. Further investigation suggested that the five participants might not have followed the comparative survey task instructions. Because the purpose of this study was to construct a structurally sound and generalizable scale, we removed these five participants and re-ran the analyses with $n = 52$ as suggested in the Rasch literature (Linacre, 2016; Smith, Linacre, & Smith, 2003).

Figure 4 presents the variable map for the 15 scenarios after removing the five poorly fitting respondents. Importantly, the order of the scenarios is unchanged. Also note that the logit estimates are converted back to their corresponding raw scores and the horizontal lines correspond to the average raw scores of 1 (*much lower*), 2 (*slightly lower*), 3 (*about the same*), 4 (*slightly higher*), or 5 (*much higher*). This variable map now provides a convenient and

meaningful way to interpret participants' levels of enacting practice for equity along the scale based on their summative raw scores (see Table 4 for score interpretations).

For example, participants scoring in the range of 40 - 45 are expected to choose *slightly lower* (score = 2) for moderate-level scenarios (e.g., Scenarios F236M and F346M) and *about the same* (score = 3) for the scenarios below their location. Moving up the scale, participants scoring in the 45 – 60 range are expected to choose *about the same* (score = 3) for the three scenarios capturing the moderate/high-level enactment (e.g., Scenarios F236H, F346H, and F126M) and *slightly higher* (score = 4) for the scenarios below their location. Participants with raw scores of 60 to 75 are expected to select *slightly higher* (score = 4) for the three high-level scenarios (i.e., Scenarios F156H, F456H, and F126H) and *much higher* (score = 5) for the scenarios below their location.

Reliability and separation. The person separation (Wright & Masters, 1982) was 2.08 (an increase from 1.76 before the deletion of extremely misfitting participants) along with a person reliability of 0.81 (an increase from 0.76). The item separation was 8.74 (an increase from 5.70) with an item reliability of 0.99 (an increase from 0.97). A separation index of above two for persons and above three for items is desirable (Linacre, 2016). As a result of removing poorly fitting cases, participants and scenarios both became more reliably differentiated and spread out along the continuum of enactment of practice for equity.

Category characteristic curves. Category characteristic curves represent the probability of responding to each of the five categories. The Andrich thresholds ($\tau_1 = -4.65$, $\tau_2 = -1.83$, $\tau_3 = 2.20$, and $\tau_4 = 4.28$) exhibit an increasing difficulty pattern. Additionally, mean enactment levels of the participants (-1.99, -1.27, 0.36, 3.05, and 5.38) increased with each successive response category. These results indicate that the response categories functioned as intended.

Goodness-of-fit. After removing the five misfitting individuals, we still identified two scenarios with unexpected response patterns (i.e., Scenarios F346L and F126H). Analysis of the unexpected responses revealed they were a matter of degree (i.e., responding with a 5 rather than a 4) in contrast to a reversed response of 5 instead of a 1.

Lastly, we conducted a principal component analysis on the residuals to check the unidimensionality assumption. The results reveal a residual first component consisting of a cluster of correlated higher-level scenarios along with a cluster of correlated lower-level scenarios. These clusters correspond to the upper and lower level regions of scenarios in Figure 3. This finding suggests that once the variance in the responses has been accounted for under the Rasch model, there is evidence of a slight dependence effect, i.e., if a person scores high on a lower level scenario then the person is likely to score high on all the lower level scenarios while the same holds true for a person scoring high on the higher level scenarios. While there is no strong evidence of multidimensionality based on these correlated clusters of residuals, further investigation is possible through future applications of the TEES.

Discussion and Conclusions

The purpose of this study was to develop a Rasch-based scenario-format scale that measures the enactment levels of teachers' equity-centered practice: The Teaching Equity Enactment Scenario (TEES) Scale. We theorized teaching practice for equity as a complex variable characterized by six interconnected facets of practice for equity that function simultaneously to influence teachers' levels of enactment of equity-centered practice. The lower levels of enacting teaching practice for equity reflect a technical view of teaching: teaching is executing a set of strategies to transmit knowledge to students regardless of their backgrounds; and teachers do not assume the responsibility to advocate on behalf of their students and

challenge inequitable practices. The higher levels of equity-centered practice on the other hand reflect a more complex view of teaching: teaching requires critical self-reflection and connecting to students' life and learning experiences; and teachers are agents of change.

The results provide a proof of concept for this novel approach, and the outcome is 15 scenario-format items that are sufficiently distributed along a hierarchical continuum to measure increasing levels of enactment of practice for equity. Because the scenario-format items were constructed to capture the variation of enacting equity-centered teaching practice, we can obtain a meaningful description of an individual's practice by locating the person's raw score along the scale and identifying the scenario(s) next to the score. The results suggest that this approach to measuring a complex construct, first proposed by Ludlow et al. (2014) to measure engagement in later-life activities, is a viable alternative to traditional narrowly defined stem-and-Likert response formats.

Moreover, the measurement invariance of the scales between the New Zealand and US participants in the pilot study suggests that the structure of enacting equity-centered teaching practice is statistically and conceptually similar for both groups. In other words, enacting equity-centered teaching practice and the difficulty progression in enacting the complex construct are similar across two countries. This result supports the findings of a previous study that compared the participants' responses using the Learning to Teach for Social Justice - Belief Scale [LTSJ-B] across three countries (Cochran-Smith, Ludlow, Ell, O'Leary, & Enterline, 2012). This implies that participants from both contexts find high-level practice for equity, which reflects a more complex view of teaching, the most challenging.

The promising results of the TEES Scale have three main implications for research on teaching and teacher education. First, as mentioned earlier, the TEES Scale presents a unique

advantage in that the items allow a rich and comprehensive description of an individual's equity-centered teaching practice based on the person's summary raw score. Although existing instruments measuring related constructs do provide useful summary information based on mutually correlated discrete items, none of the scales provide a holistic and qualitative interpretation of what a person's score means in relation to the complex construct. Moreover, a close examination of the theoretical foundations of existing instruments suggests that the political and critical aspects of teaching culturally and linguistically diverse learners with a commitment to equity or social justice are left out of most discussions. When the theoretical foundation does not reflect the political and critical aspects of teaching, the items of particular instruments are also unlikely to capture the characteristics. Recognizing the limitations of existing instruments, the development of the TEES Scale puts equity front and center in the conceptualization of equity-centered teaching practice and the deliberate decisions made during item development.

Second, the self-report TEES Scale provides a new way to capture teachers' equity-centered practice, which can be one outcome indicator for teacher education programs intended to prepare teachers who teach for equity and challenge inequity. This instrument contributes to research that seeks to improve programs or investigates how, why, to what extent, and under what conditions teacher candidates learn to enact this kind of practice, and the connection between practice and student learning outcomes. Research programs such as this can allow the teacher education community to better understand how to prepare and support teachers to enact teaching practice that responds to diversity, challenges educational inequities, and promotes social justice in a democratic society of a changing and uncertain world.

Third, the TEES Scale, which is intended to be formative, could serve as a platform for initiating and guiding teacher candidates' reflections on their practice. The notion of reflection was first defined by Dewey as the "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends" (1993 cited in Jay & Johnson, 2002, p. 74). In teacher education specifically, Zeichner and Lipton (1987, 1996) argue that *reflective teaching* could facilitate student teachers' self-directed growth and professional development. Moreover, reflective teaching does not merely focus on solving a practical problem in the classroom but also entails the examination of one's beliefs, values, worldviews, and goals that guide one's work (Zeichner & Lipton, 1987, 1996). Recently, studies have proposed different dimensions or typologies of reflection for fostering reflective activities (e.g., Jay & Johnson, 2002) and various tools such as portfolio and performance assessments that facilitate reflective thinking in teacher education programs (e.g., Jay & Johnson, 2002).

In light of the literature on reflective teaching, the TEES Scale could be a useful professional development tool to facilitate critical reflection on one's process of making sense of and learning to enact equity-centered practice. Specifically, teacher candidates could first use the TEES Scale to assess their teaching during and at the end of their student teaching. The equity-centered practice captured in each scenario then offers a point of entry into discussions with mentor teachers on specific situations, puzzling problems, concerns, and challenges with students in the classroom. For example, if a teacher candidate selects "*about the same*" to a moderate-level scenario, the following questions can include: What happens in the classroom? What is working and not working for whom? What informs the person's classroom and instructional decisions? Once the areas or matters for reflection are identified, discussions about

alternative perspectives and approaches can be explored, discussed, and considered. Examples of reflective questions can include: What are the alternative perspectives of what occurs in the classroom? What are the obstacles or struggles? Because the scenarios capture the complex nature of equity-centered teaching practice that does not merely reflect discrete teaching strategies, discussions stimulated by the scenarios could involve deeper reflection on how one's assumptions and world views inform one's teaching practice.

Limitations

Developing a scenario-format scale that captures teachers' enactment of practice for equity was itself a complex process that involved decision-making on multiple trade-offs (e.g., between the complexity of equity practice and the cognitive demand of items for the participants). While the results are promising, we recognize several limitations. One primary limitation relates to the small sample size ($n = 73$ for the pilot and $n = 52$ for the final study) and representativeness of the sampled participants. While a large sample size is not required for conducting a Rasch analysis, the sample sizes of both phases were lower than the desirable number (between 75 and 150 - five to ten times of the number of scenarios) recommended by the general guideline for item analysis. Due to our sampling approaches, the participants of this study may not be representative of K-12 teachers and teacher candidates teaching academic subjects in K-12 settings. In spite of the participants' various teaching backgrounds (i.e., years, levels, and subject areas of teaching), the sample was mostly homogeneous regarding, for instance, participants' race and gender, geographic locations, school contexts, and program training. Constrained by the small sample size and the limited background information sought from the participants, we could not conduct analyses to investigate nuances across subgroups beyond what we have done.

Instruments of this sort are not culturally neutral. In this development study a homogeneous sample was obtained, which reduced sources of construct irrelevant variance but reduced the generalizability of the scale. The comparative absence of teachers of color in the sample is problematic in that the scale can reinforce White interests/perspectives as normative and silence the voices and development of teachers of color (Sleeter, 2017). Item difficulties may be different for different groups of teachers, based on their knowledge and experience. Further work is needed to engage in reciprocal dialogue with teachers of color around the concepts behind the scale and with the items themselves.

Further, the survey completion rate (approximately 60%) was less than ideal for both phases. The high dropout rate was not surprising and was likely due to both the length and the novelty of the scenarios. Unlike typical survey instructions for Likert-type items that ask participants to rate a simple statement, the survey instructions in this study asked respondents to reflect on their practice, compare their practice against the practice described in a scenario, and give a rating. This task was undoubtedly atypical and even overwhelming for some participants. Our investigation of participants' dropout patterns also suggests that the majority of dropouts occurred at the beginning of the survey. This limitation points to a need to have a user-friendly survey instruction.

Lastly, although the current items are sufficiently spread to define the construct and participants' levels of enacting practice for equity can be reliably differentiated, the scenario-format items could be more "on target" with the participants' enactment levels. In other words, the current scenarios tended to be relatively easy for most participants to endorse.

Recommendations for Future Research

The following recommendations for future research include endeavors to improve the validity and utility of the TEES Scale. First, we are currently working on a series of enhancement and validation studies of the scale. The TEES Scale can be improved by making the critical and political aspects of practice for equity more explicit (and more difficult) in the high-level scenarios. For example, teachers enact practice that facilitates students' capacity to consume, investigate, challenge, and construct knowledge rather than accepting knowledge as a fixed, objective, and neutral intellectual product (Ladson-Billings, 2009; Sleeter & Carmona, 2017). Having a more user-friendly and graphic-oriented instruction could reduce confusion and increase accuracy in the response processes. The validation studies could examine the extent to which the item responses of the TEES Scale are related to other scales (e.g., the LTSJ-B Scale, the FEET) that measure a theoretically similar construct to teaching practice for equity. Additionally, the validation efforts could investigate the influence of social desirability on participants' responses. Furthermore, larger and more diverse samples are needed to test the generalizability and invariance of the current scale results.

Second, we suggest that with some revisions the self-report aspect of the instrument can be transformed into an observational tool. Mentors and collaborating teachers could then use the instrument to observe teacher candidates' and colleagues' classroom practices. Those same teacher candidates or teachers could use the self-report version, and various reliability, validity, and invariance analyses could be performed to better understand the relationships between the observational and self-report versions of the TEES Scale. In its present form, however, teacher candidates and their mentors or teachers and their colleagues can use the scenarios as a platform to discuss, elaborate, and reflect on their pedagogy, teaching philosophy, challenges, and confusions, and ways to respond to certain challenges in classroom and school contexts. The

TEES Scale is envisioned not only as a measure of status (i.e., one's location on the continuum) and a potential indicator of change in practice but also as a springboard for productive and lively dialogues among educators considering improving the measurement and implementation of the complex practice of equity-centered teaching.

Final draft before proofing

References

- Achinstein, B., & Ogawa, R. T. (2012). New teachers of color and culturally responsive teaching in an era of educational accountability: Caught in a double bind. *Journal of Educational Change*, 13(1), 1-39.
- Aitken, G., & Sinnema, C. (2008). *Effective pedagogy in social sciences/tikanga a iwi*. Wellington, NZ: Ministry of Education.
- Alton-Lee, A. (2003). *Quality teaching for diverse students in schooling: Best evidence synthesis*. Wellington, NZ: Ministry of Education.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in mathematics/pangarau*. Wellington, NZ: Ministry of Education.
- Arnold, J., Edwards, T., Hooley, N., & Williams, J. (2012). Theorising on-site teacher education: Philosophical project knowledge (PPK). *Asia-Pacific Journal of Teacher Education*, 40(1), 67-78.
- Au, W. (2016). Meritocracy 2.0: High-stakes, standardized testing as a racial project of neoliberal multiculturalism. *Educational Policy*, 30(1), 39-62.
- Ayers, W., Kumashiro, K., Meiners, E., Quinn, T., & Stovall, D. (2016). *Teaching toward democracy: educators as agents of change*. London, England: Routledge.
- Baluch, S., Greig, T., Ponterotto, J., & Rivera, L. (1998). Development and initial score validation of the teacher multicultural attitude survey. *Educational and Psychological Measurement*, 58(6), 1002-1015.
- Bishop, R., Berryman, M., Cavanagh, T., & Teddy, L. (2009). Te Kotahitanga: Addressing educational disparities facing Maori students in New Zealand. *Teaching and Teacher Education*, 25, 734-742.
- Borg, I., & Shye, S. (1995). *Facet theory*. Thousand Oaks, CA: Sage.
- Buehler, J., Gere, A. R., Dallavis, C., & Haviland, V. S. (2009). Normalizing the fraughtness. *Journal of Teacher Education*, 60(4), 408-418.
- Carmen Salazar, M. del. (2018). Interrogating teacher evaluation: Unveiling Whiteness as the normative center and moving the margins. *Journal of Teacher Education*, 69(5), 463-476.

- Chang, W-C C. (2017). *Measuring the complexity of teachers' enactment of practice for equity: A Rasch model and facet theory-based approach*. Retrieved from ProQuest Dissertations & Theses Global. (ProQuest ID 1891348451)
- Cochran-Smith, M., Barnatt, J., Lahann, R., Shakman, K., & Terrell, D. (2009). Teacher education for social justice: Critiquing the critiques. In W. Ayers, T. Quinn & D. Stovall (Eds.), *The handbook of social justice in education* (pp. 625-639). Philadelphia: Taylor and Francis.
- Cochran-Smith, M. (2010). Toward a theory of teacher education for social justice. In M. Fullan, A. Hargreaves, D. Hopkins, & A. Lieberman (Eds.), *Second international handbook of educational change* (pp. 445-467). Dordrecht: Springer.
- Cochran-Smith, M., Ludlow, L., Ell, F., O'Leary, M., & Enterline, S. (2012). Learning to teach for social justice as a cross-cultural concept: Findings from three countries. *European Journal of Educational Research*, 1(2), 171-198.
- Cochran-Smith, M., Ell, F., Grudnoff, L., Ludlow, L., Haigh, M., & Hill, M. (2014). When complexity theory meets critical realism: A platform for research on initial teacher education. *Teacher Education Quarterly*, 41(1), 105-122.
- Cochran-Smith, M., Ell, F., Ludlow, L., Grudnoff, L., & Aitken, G. (2014). The challenge and promise of complexity theory for teacher education research. *Teachers College Record*, 116(5), 1-38.
- Cochran-Smith, M., Ell, F., Grudnoff, L., Haigh, M., Hill, M., & Ludlow, L. (2016). Initial teacher education: What does it take to put equity at the center? *Teaching and Teacher Education*, 57, 67-78.
- Cochran-Smith, M., Carney, M.C., Keefe, E.S., Burton, S., Chang, W-C., Fernández, M.B., Miller, A., Sánchez, J.G., & Baker, M. (2018). *Reclaiming accountability in teacher education*. New York: Teachers College Press.
- Dalton, S. S. (2007). *Five standards for effective teaching: How to succeed with all learners*. San Francisco, CA: Wiley.
- D'Andrea, M., Daniels, J., & Noonan, M. (2003). New developments in the assessment of multicultural competence: The multicultural awareness-knowledge-skills survey-teachers form. In D. B. Pope-Davis (Ed.), *Multicultural competencies in counseling and psychology* (pp. 154-167). Thousand Oaks, CA: Sage.
- Danielson, C. (2014). *The framework for teaching: Evaluation instrument (2013 Edition)*. Retrieved on 22 April, 2018 from: <https://www.danielsongroup.org/download/?download=448>

- Fraser, N. (2003). Social justice in an age of identity politics: Redistribution, recognition and participation. In N. Fraser & A. Honneth (Eds.), *Redistribution or recognition: A political-philosophical debate* (pp. 7-109). London: Versos.
- Fraser, N., & Honneth, A. (Eds.). (2003). *Redistribution or recognition: A political-philosophical debate*. London: Versos.
- Fraser, N. (2009). *Scales of justice. reimagining political space in a globalizing world*. New York: Columbia University Press.
- Gay, G. (2000). *Culturally responsive teaching: Theory, research & practice*. New York, NY: Teacher College Press.
- Gore, J. (2001). Beyond our differences: A reassembling of what matters in teacher education. *Journal of Teacher Education*, 52(2), 124-135.
- Grant, C., & Agosto, V. (2008). Teacher capacity and social justice in teacher education. In M. Cochran-Smith, S. Feiman Nemser, J. McIntyre, & K. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed.). Philadelphia: Taylor and Francis.
- Grudnoff, L., Haigh, M., Hill, M., Cochran-Smith, M., Ell, F., & Ludlow, L. (2017a). Rethinking initial teacher education: Preparing teachers for schools in low socio-economic communities in New Zealand. *Journal of Education for Teaching*, 42(4), 451-467.
- Grudnoff, L., Haigh, M., Hill, M., Cochran-Smith, M., Ell, F., & Ludlow, L. (2017b). Teaching for equity: Insights from international evidence with implications for a teacher education curriculum. *The Curriculum Journal*, 28(3), 1-22.
- Guttman, R., & Greenbaum, C. W. (1998). Facet theory: Its development and current status. *European Psychologist*, 3(1), 13-36.
- Henry, G. B. (1986). Cultural diversity awareness inventory. Virginia: Hampton University Mainstreaming Outreach Project. (ERIC Document Reproduction Service No. ED282657)
- Hollins, E. R., & Torres Guzman, M. (2005). Research on preparing teachers for diverse population. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 477-544). Mahwah, NJ: Lawrence Erlbaum.
- James, M., & Pollard, A. (2011). TLRP's ten principles for effective pedagogy: Rationale, development, evidence, argument and impact. *Research Papers in Education*, 26(3), 275-328.

- Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education, 18*, 73-85.
- Kretchmar, K., & Zeichner, K. (2016). Teacher prep 3.0: A vision for teacher education to impact social transformation. *Journal of Education for Teaching, 42*(4), 417-433.
- Ladson-Billings, G. J. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465-491
- Ladson-Billings, G. (2009). *The dream-keeper: Successful teachers of African American children*. San Francisco, CA: Jossey-Bass.
- Lampert, J., & Burnett, B. (Eds.). (2015). *Teacher education for high poverty schools* (Vol. 2). New York, NY: Springer.
- Linacre, J. M. (2015). WINSTEPS (Version 3.91.2) [Computer program]. Beaverton, OR: Winsteps.com. Retrieved January 1, 2016. Available from <http://www.winsteps.com/>
- Linacre, J. M. (2016). Winsteps® Rasch measurement computer program User's Guide. Beaverton, Oregon: Winsteps.com
- Lord, F. M., & Novick, M. R. (1968). *Statistical test theories of mental test scores*. Reading, MA: Addison-Wesley.
- Ludlow, L. H., & Haley, S. M. (1995). Rasch model logits: Interpretation, use, and transformation. *Educational and Psychological Measurement, 55*(6), 967-975.
- Ludlow, L. H., Enterline, S. E., & Cochran-Smith, M. (2008). Learning to teach for social justice-beliefs scale: An application of Rasch measurement principles. *Measurement and Evaluation in Counseling and Development, 40*(4), 194-214.
- Ludlow, L. H., Matz-Costa, C., Johnson, C., Brown, M., Besen, E., & James, J. B. (2014). Measuring engagement in later life activities: Rasch-based scenario scales for work, caregiving, informal helping, and volunteering. *Measurement and Evaluation in Counseling and Development, 47*(2), 127-149.
- Ludlow, L. H., Matz-Costa, C., & Klein, K. (2018). Enhancement and validation of the productive engagement portfolio-scenario (PEP-S8) scales. *Measurement and Evaluation in Counseling and Development, 52*(1), 15-37.
- McDonald, M., & Zeichner, K. (2009). Social justice teacher education. In W. Ayers, T. Quinn, & D. Stovall (Eds.), *The handbook of social justice in education* (pp. 595-610). Philadelphia: Taylor and Francis.

- MET [Measures of Effective Teaching] Project. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study*. Retrieved on 22 April, 2018 from: <http://files.eric.ed.gov/fulltext/ED540958.pdf>
- Piburn, M., & Sawada, D. (2000). Reformed Teaching Observation Protocol (RTOP): Reference Manual. Arizona Collaborative for Excellence in the Preparation of Teachers. ACEPT Technical Report No. IN00-3
- Quartz, K. H. (2003). "Too Angry To Leave": Supporting new teachers' commitment to transform urban schools. *Journal of Teacher Education*, 54(2), 99–111.
- Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. (Copenhagen, Danish Institute for Educational Research), expanded edition (1980) with foreword and afterword by B. D. Wright. Chicago, IL: The University of Chicago Press.
- Sleeter, C. E., & Grant, C. (1987). An analysis of multicultural research in the United States. *Harvard Educational Review*, 57(4), 421–445.
- Sleeter, C. E. (2001). Preparing teachers for culturally diverse schools: Research and the overwhelming presence of Whiteness. *Journal of Teacher Education*, 52(2), 94-106.
- Sleeter, C. E. (2008). Equity, democracy, and neoliberal assaults on teacher education. *Teaching and Teacher Education*, 24(8), 1947-1957.
- Sleeter, C. (2009). Teacher education, neoliberalism and social justice. In W. Ayers, T. Quinn, & D. Stovall (Eds.), *Handbook of social justice in education* (pp. 611-624). New York: Routledge.
- Sleeter, C. E. (2014). Toward teacher education research that informs policy. *Educational Researcher*, 43(3), 146-153.
- Sleeter, C. E. (2017). Critical race theory and the Whiteness of teacher education. *Urban Education*, 52(2), 155-169.
- Sleeter, C. E., & Carmona, J. F. (2017). *Unstandardizing curriculum: Multicultural teaching in the standards-based classroom*. New York, NY: Teacher College Press.
- Smith, R. M., Linacre, J. M., & Smith, Jr., E. V. (2003). Guidelines for manuscripts. *Journal of Applied Measurement*, 4, 198-204.
- Siwatu, K. O. (2007). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education*, 23(7), 1086-1101.
- Tan, E., & Barton, A. C. (2012). *Empowering science and mathematics education in urban schools*. Chicago: University of Chicago Press.

- Villegas, A. M., & Lucas, T. (2002). *Educating culturally responsive teachers: A coherent approach*. Albany, NY: State University of New York Press.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.
- Young, I. M. (2011). *Justice and the politics of difference*. Princeton, NJ: Princeton University Press.
- Zeichner, K. (1993). Connecting genuine teacher development to the struggle for social justice. *Journal of Education for Teaching*, 19(1), 5–21.
- Zeichner, K. (2005). A research agenda for teacher education. In M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 737–759). Mahwah, NJ: Erlbaum and Associates.
- Zeichner, K. M., & Lipton, D. P. (1987). Teaching student teachers to reflect. *Harvard Educational Review*, 57(1), 23-48
- Zeichner, K. M., & Liston, D. P. (1996). *Reflective teaching: An introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Zygmunt, E., & Clark, P. (2016). *Transforming teacher preparation for social justice*. New York, NY: Teachers College Press.

Appendix. Survey instrument

Section I. Enactment of Practice for Equity

Instruction: Each of the scenarios below captures some aspects of teachers' practice for equity. Equity-centered teaching recognizes and challenges social and educational inequities and promotes students' learning, broadly defined to include academic, social, emotional, civic and critical learning

To respond to the scenario-style items:

- Consider each scenario holistically, reflect on your own practice, and compare your own practice against the individual teachers' practice described in each scenario.
- Based on the 5-point scale, choose one of these:
 - *About the same* means that your practice is similar to the practice of the teacher in the specific scenario;
 - *Slightly lower* or *Much lower* means that you consider your practice to be at a lower enactment level than the practice of the person in the scenario;
 - *Slightly higher* or *Much higher* means that you consider your practice to be at a higher enactment level than the practice of the person in the scenario.

Practice Item:

Joe is a teacher who has positive relationships with the parents/caregivers of some of the students in his class. He connects with some students and generally appreciates the diverse experiences they bring to school with them. Joe sometimes involves his students in helping to design a lesson or choosing a topic of their interests. Although he usually has students concentrate on their own assignments, he sometimes encourages students to work together as groups. Joe's classroom is welcoming and comfortable to some students.

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Joe's level?

Much lower Slightly lower About the same Slightly higher Much higher

For the remainder of these items, consider each scenario holistically, reflect on your own practice, and compare your level of enactment of practice for equity against the level of the teacher described in the scenario. Based on the 5-point scale, choose one of these:

- *About the same:* Your practice is similar to the practice of the teacher in the specific scenario;
- *Slightly lower or Much lower:* You do less or much less well than the practice of the person in the scenario;
- *Slightly higher or Much higher:* You do better or much better than the practice of the person in the scenario.

1. Tim holds high expectations for some students in his class and mostly communicates these expectations clearly. He generally sees students' home culture as a strength and collaborates with some parents/caregivers. He sometimes lets his students choose a topic consistent with their interests to further their learning. Tim sometimes draws on cultural examples to design learning experiences that are relevant to students. He utilizes a selected number of approaches to explain key concepts. His explanations are clear and interesting to all students. (F126M)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Tim's level?

Much lower Slightly lower About the same Slightly higher Much higher

2. Katherine cares for and respects her students. She encourages students to be independent learners and to investigate and build understandings of their own, and she involves them in setting criteria and goals for their learning. She constructs her teaching practice to be engaging for all students and integrates a variety of assessment approaches into her teaching. Katherine interacts with students to provide constructive feedback and adjusts her practice appropriately. She monitors and facilitates collaborative learning among her students. (F346H)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Katherine's level?

Much lower Slightly lower About the same Slightly higher Much higher

3. Adrian considers her role as a teacher primarily as transmitting knowledge to students. Adrian sets attainable goals for students but struggles to engage them. She adheres to standards and curriculum documents to design her lessons and makes sure that students memorize the content. She often uses the same teaching strategies although she is unsure whether other approaches would be more or less effective for student learning. Adrian tends to work alone and sticks with what she knows. (F156L)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Adrian's level?

Much lower Slightly lower About the same Slightly higher Much higher

4. Ryan involves parents in school activities and draws on students' cultures as valued resources to design their learning experiences. His interaction with students is genuinely warm and caring. Ryan involves students in making decisions and setting classroom expectations that are relevant to all of them. He constantly encourages and monitors supportive interactions among students so that they help each other and take responsibility for each other's learning. Ryan effectively arranges the classroom space to be inviting, safe, and accessible to all students. (F236H)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Ryan's level?

Much lower Slightly lower About the same Slightly higher Much higher

5. Dave believes good teaching is executing a set of techniques to ensure that students attain curriculum expectations for their grade/year levels. He relies solely on standards or curriculum documents to identify learning priorities and teaching approaches. He designs assessments on his own and generally uses them to check whether students meet the minimum academic standards. Dave reviews his students' test results, sometimes altering his practice to boost their scores. (F456L)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Dave's level?

Much lower Slightly lower About the same Slightly higher Much higher

6. Kevin sets achievable goals for his students but finds it hard to communicate them. He sees students' home culture as challenging and doesn't expect parents to be his partners in teaching. Kevin sets out lessons for his students so they know what they need to do. He uses textbooks and self-designed learning experiences that he believes deliver the appropriate curriculum. He utilizes a few different teaching approaches to explain key concepts. At times he feels he does not understand the concepts he is teaching well and this affects his ability to capture students' interests. (F126L)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Kevin's level?

Much lower Slightly lower About the same Slightly higher Much higher

7. Kim generally cares for and respects her students. Overall, she believes in students' capacity to take initiative regarding their learning, and sometimes involves them in designing assessments and setting classroom expectations. Her teaching practice engages some students and assessments are generally integrated into her teaching. Kim sometimes circulates among students to provide feedback and modify her practice. Accordingly, she sometimes monitors and facilitates classroom interactions among students. (F346M)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Kim's level?

Much lower Slightly lower About the same Slightly higher Much higher

8. Megan fully embraces her responsibility to identify and challenge classroom and school practices that promote inequities for students. Megan sets cognitively challenging goals and communicates to her students clearly and consistently. She purposefully draws upon a variety of sources to cultivate their conceptual understanding and encourages students to challenge information in textbooks. She deliberately uses various pedagogical strategies to capture students' interests. Megan also works with others in a professional community to pose questions, reflect on her own assumptions, and proactively respond to student needs. (F156H)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Megan's level?

Much lower Slightly lower About the same Slightly higher Much higher

9. Christine occasionally engages with parents but generally she sees this as unnecessary. She has a quiet, reserved manner with her students, approaching all students the same way. Christine, rather than involving the students, makes most of the decisions in the classroom. Because she believes that students should work individually, Christine does not see the need to facilitate interactive skills and usually assigns work for students to carry out on their own. Her classroom is arranged to be a generally quiet, non-interactive space. (F236L)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Christine's level?

Much lower Slightly lower About the same Slightly higher Much higher

10. Michael generally has a positive sense of professional identity as a teacher, but sees advocacy on behalf of students as a peripheral role. Michael supports student learning but relies on standards and personal experiences with some inputs from students to identify learning priorities and teaching strategies. He usually designs assessments on his own and sometimes integrates them into his teaching to evaluate his own practice and to give feedback to students. Michael sometimes reflects on and adjusts his practice based on students' needs to better motivate their interests. (F456M)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Michael's level?

Much lower Slightly lower About the same Slightly higher Much higher

11. Maria holds high expectations for all students and clearly communicates challenging and meaningful learning goals. Maria sees students' home cultures as assets and collaborates closely with parents/caregivers. She encourages students to explore topics that connect to their lives. She consistently draws on students' prior knowledge and cultures and purposefully designs relevant learning experiences for all. Maria skillfully uses a variety of instructional approaches to motivate students' learning. Her explanations are clear, compelling, and accurate. (F126H)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Maria's level?

Much lower Slightly lower About the same Slightly higher Much higher

12. Tracey cooperates with some parents/community members and draws on some students' culture as examples to design their learning experiences. Overall, she genuinely cares for and respects her students, although she sometimes engages in stereotypical thinking. Tracey sometimes involves her students in designing a lesson or setting classroom rules. Although she often has students concentrate on their own work, she sometimes encourages collaboration among students. Tracey's classroom is inviting and safe for some students. (F236M)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Tracey's level?

Much lower Slightly lower About the same Slightly higher Much higher

13. Juan is deeply committed to supporting the learning and life of diverse students, advocating on behalf of them, and contributing to the profession. Juan builds on students' perspectives and draws on a variety of sources to identify learning priorities and teaching strategies. He involves students in designing assessments and fully integrates assessment into his instruction to provide constructive and timely feedback to students. Juan takes charge of his professional learning through continuous reflection on his practice and experimenting with new approaches to motivate and respond to students' learning needs. (F456H)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Juan's level?

Much lower Slightly lower About the same Slightly higher Much higher

14. Sarah focuses her teaching on the academic side of things rather than seeing the focus of her job as involving caring for her students. She sees herself as the authority in her classroom and decides how assessments will be carried out and how students should behave. Her teaching practice covers the curriculum, and she carries out assessments to check up of student learning. Sarah uses tried-and-true learning activities and expects her students to complete them independently. She runs a quiet classroom in which students learn individually most of the time. (F346L)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Sarah's level?

Much lower Slightly lower About the same Slightly higher Much higher

15. Erin generally takes responsibility for supporting her students' learning. Her lesson planning is mostly guided by curriculum documents and textbooks, and she sometimes invites students' ideas and opinions. Erin relies on a familiar repertoire of teaching approaches to capture students' interests. Erin sometimes reflects on and checks to see whether certain approaches are more effective than others in responding to student needs. (F156M)

Considering this scenario holistically and reflecting on your own practice, how would you describe your level of enactment of practice for equity in comparison to Erin's level?

Much lower Slightly lower About the same Slightly higher Much higher

Table 1. Linkages between the Five Syntheses and the Six Principles of Practice

Theme	Characteristics/principles/domains on which the theme draws				
	BES*	TLRP	MET	CREDE	TK
1. Selecting worthwhile content and designing and implementing learning opportunities aligned with valued outcomes	1.1, 1.5, 1.6, 1.7, 2.2 3.2, 3.4, 3.7, 3.8, 3.9	1, 2	3b, 3a, 3c	4	
2. Connecting to students' lives and experiences	1.3, 2.1, 2.4 3.3, 3.5	3, 8	2a, 2b, 2c, 2d, 2e	2, 3	
3. Creating learning-focused, respectful and supportive learning environment	1.2, 2.3 3.3, 3.5	6, 7	3d, 3e,	1, 5	
4. Using evidence to scaffold learning and improve teaching	1.4, 1.8, 1.9, 1.10, 3.6	4, 5			
5. Taking an inquiry stance for professional engagement and learning	3.10	9, 10			
6. Recognizing and challenging inequities					Overall

*BES – 1= General (Alton-Lee, 2003); 2= Social sciences (Aitken & Sinnema, 2008); 3= Mathematics (Anthony & Walshaw, 2007)

Final draft before proofing

Table 2. Structure for Scenario Development

Level of Enactment	Scenario #	Facet 1	Facet 2	Facet 3	Facet 4	Facet 5	Facet 6
High Level of Enactment	Scenario F126H	X	X				X
	Scenario F236H		X	X			X
	Scenario F346H			X	X		X
	Scenario F456H				X	X	X
	Scenario F156H	X				X	X
Moderate Level of Enactment	Scenario F126M	X	X				X
	Scenario F236M		X	X			X
	Scenario F346M			X	X		X
	Scenario F456M				X	X	X
	Scenario F156M	X				X	X
Low Level of Enactment	Scenario F126L	X	X				X
	Scenario F236L		X	X			X
	Scenario F346L			X	X		X
	Scenario F456L				X	X	X
	Scenario F156L	X				X	X

*Note: X indicates that the specific facet is included in the scenario.

Final draft before

Table 3. An Example of Sentence Mapping Structure and Scenario-format Item for Facets 1, 2, and 6

Sentence mapping structure	A scenario capturing the high-level enactment	A scenario-format item
<p>(Person) holds (Facet 1, Characteristic 1). He/she sees students' home culture (Facet 2, Characteristic 1), and (Facet 2, Characteristic 1). (Person) (Facet 6, Characteristic 1). In the classroom, he/she (Facet 2, Characteristic 2) and (Face 1, Characteristic 2). (Person) (Facet 1, Characteristic 3). His/her (Facet 1, Characteristic 3)</p>	<p>Maria holds high expectations for all students and clearly communicates challenging and meaningful learning goals. She sees students' home culture as an asset and collaborates closely with parents/caregivers as partners. Maria encourages students to take initiative regarding their learning. In the classroom, she consistently draws on students' prior knowledge and culture and purposefully designs relevant learning experiences for all. Maria skillfully uses a variety of instructional approaches to motivate students' learning. Her explanations are clear, compelling, and accurate.</p>	<p><u>Prompt:</u> What is your assessment of your own level of enactment compared to that teacher's practice described in that scenario?</p> <p><u>Scenario:</u> Maria holds high expectations for all students and clearly communicates challenging and meaningful learning goals. She sees students' home culture as an asset and collaborates closely with parents/caregivers as partners. Maria encourages students to take initiative regarding their learning. In the classroom, she consistently draws on students' prior knowledge and culture and purposefully designs relevant learning experiences for all. Maria skillfully uses a variety of instructional approaches to motivate students' learning. Her explanations are clear, compelling, and accurate.</p> <p><u>Response option:</u> Much lower Slightly lower About the same Slightly higher Much higher</p>

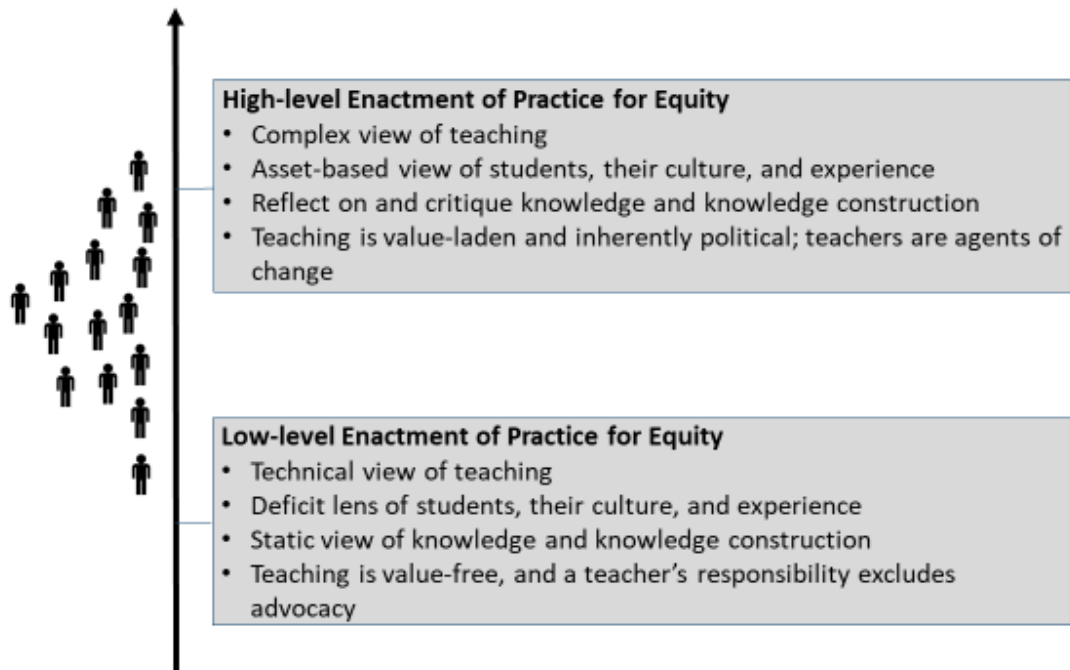
*Note: For the sentence mapping structure (first column), the bolded parts in the parentheses are the formal elements that must be filled in with descriptions of practice drawn from selected facets. The unbolded parts are the informal elements in the sentence mapping structure that remain constant to link the facet descriptions together and to provide a meaningful context.

Table 4. Teaching Equity Enactment Scenario Scale Score Interpretations

Scenario Score	Enactment Level	Description of score	Example of Scenario
75	Extremely high level of enactment	You are “much higher” than all scenarios presented	
60 - 74	Very high level of enactment	On average, you are “slightly higher” than Scenario F156H and “much higher” than the scenarios below this section	F156H (Facet score = 9) Megan fully embraces her responsibility to identify and challenge classroom and school practices that promote inequities for students. Megan sets cognitively challenging goals and communicates to her students clearly and consistently. She purposefully draws upon a variety of sources to cultivate their conceptual understanding and encourages students to challenge information in textbooks. She deliberately uses various pedagogical strategies to capture students’ interests. Megan also works with others in a professional community to pose questions, reflect on her own assumptions, and proactively respond to student needs.
55 - 59	High level of enactment	On average, you are “about the same” as Scenario F346H and “slightly higher” than the scenarios below this section	F346H (Facet score = 8) Katherine cares for and respects her students. She encourages students to be independent learners and to investigate and build understandings of their own, and she involves them in setting criteria and goals for their learning. She constructs her teaching practice to be engaging to all students, and integrates a variety of assessment approaches into her teaching. Katherine interacts with students to provide constructive feedback and adjusts her practice appropriately. She monitors and facilitates collaborative learning among her students.
45 - 54	Moderately high level of enactment	On average, you are “about the same” as Scenario F126M and “slightly higher” than	F126M (Facet score = 7) Tim holds high expectations for some students in his class and mostly communicates these expectations clearly. He generally sees students’ home culture as a strength and collaborates with some parents/caregivers. He sometimes lets his students choose a topic consistent with their interests to further their learning. Tim sometimes

		the scenarios below this section	draws on cultural examples to design learning experiences that are relevant to students. He utilizes a selected number of approaches to explain key concepts. His explanations are clear and interesting to all students.
40 - 44	Moderate level of enactment	On average, you are “slightly lower” than Scenario F236M and “about the same” as the scenarios below this section	F236M (Facet score = 6) Tracey cooperates with some parents/community members and draws on some students’ culture as examples to design their learning experiences. Overall, she genuinely cares for and respects her students, though sometimes engages in stereotypical thinking. Tracey sometimes involves her students in designing a lesson or setting classroom rules. Although she often has students concentrate on their own work, she sometimes encourages collaboration among students. Tracey’s classroom is inviting and safe for some students.
30 - 39	Low level of enactment	On average, you are “slightly lower” than Scenario F456L and “about the same” as the scenarios below this section	F456L (Facet score = 3) Dave believes good teaching is executing a set of techniques to ensure that students attain curriculum expectations for their grade/year levels. He solely relies on standards or curriculum documents to identify learning priorities and teaching approaches. He designs assessments on his own and generally uses them to check whether students meet the minimum academic standards. Dave reviews his students’ test results, sometimes altering his practice to boost their scores.
15 - 29	Extremely low level of enactment	On average, you are “much lower” than F236L and all other scenarios presented	F236L (Facet score = 3) Christine occasionally engages with parents but generally she sees this as unnecessary. She has a quiet, reserved manner with her students, approaching all students the same way. Christine, rather than involving the students, makes most of the decisions in the classroom. Because she believes that students should work individually, Christine does not see the need to facilitate interactive skills and usually assigns work for students to carry out on their own. Her classroom is arranged to be a generally quiet, non-interactive space.

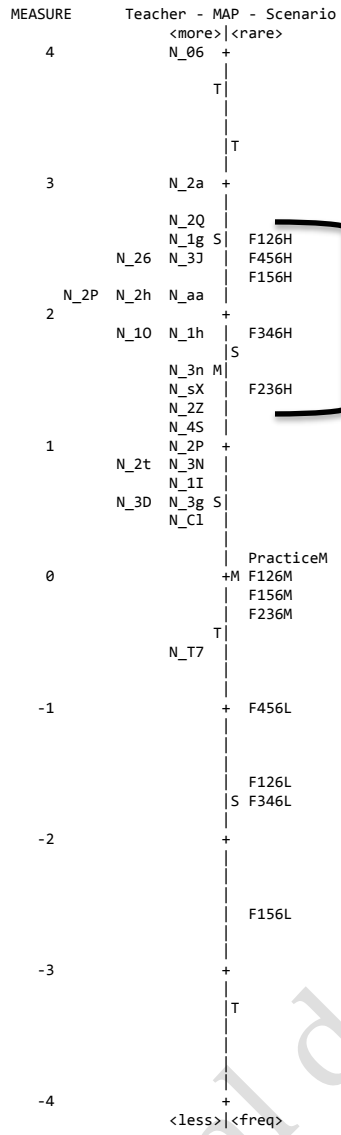
Figure 1. Difficulty progression of equity-centered teaching practice



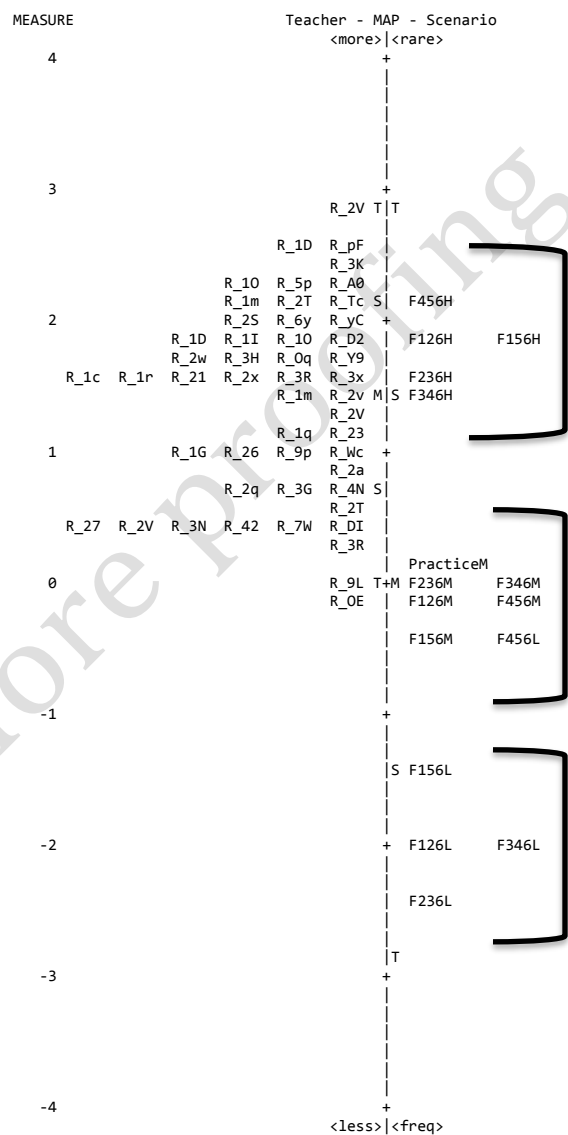
Final draft before

Figure 2. Variable maps: Comparison between New Zealand and the US participants

NZ Participants (N = 23)



US Participants (N = 50)



Final draft before printing

Figure 3. Variable map: Final study (57 cases)

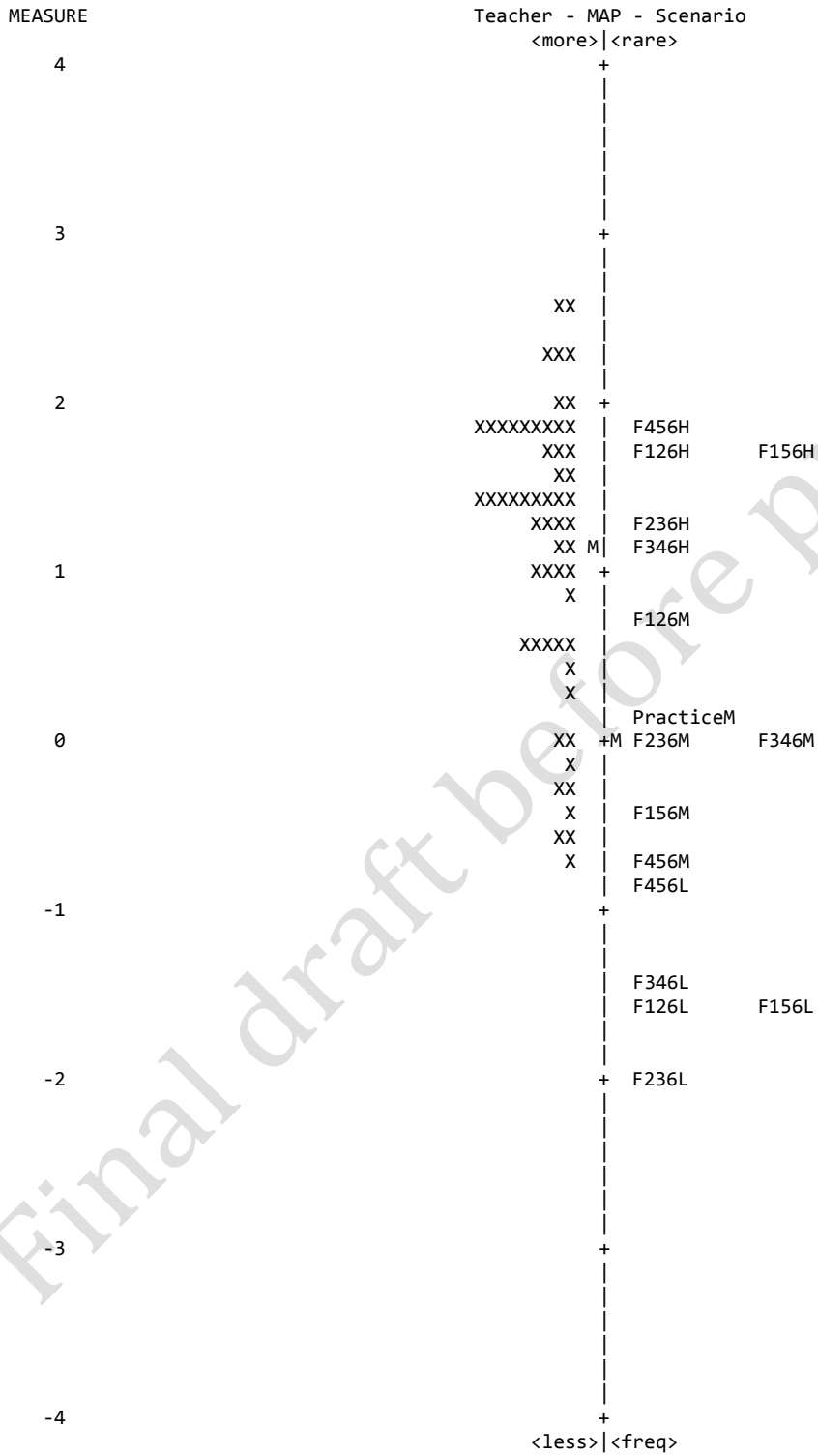


Figure 4. Construct definition map (52 cases)

