

International Journal of Learning, Teaching and Educational Research
Vol. 22, No. 4, pp. 494-511, April 2023
<https://doi.org/10.26803/ijlter.22.4.27>
Received Feb 6, 2023; Revised Apr 20, 2023; Accepted Apr 29, 2023

Assessment of Math Teachers' Dispositions to Improve Urban Teacher-Leaders' Growth and Effectiveness

L. L. Moore 

University of South Florida, St. Petersburg, Florida

W. Steve Lang* 

University of South Florida, St. Petersburg, Florida

Abstract. This paper focuses on the dispositions of in-service math teacher-leaders, as defined in the Interstate New Teacher Assessment and Support Consortium (InTASC) Standards by the Council of Chief State Officers (CCSSO) on 10 sets of critical dispositions. Teachers (N = 281) were assessed using versions of two instruments developed as part of the Disposition Assessment Aligned with Teacher Standards (DAATS) battery. One instrument is a self-report questionnaire called Beliefs About Teaching 2 (BATS2), and the second is a guided reflection, the Experiential Teaching Questionnaire (ETQ2). Item analysis and scoring were done utilizing the Rasch model of item response theory and scaled on a modification of the original Krathwohl affective taxonomy. A discussion of the validity and reliability of these versions of the instrumentation is included, along with a qualitative review of sample responses illustrating the scoring. Finally, there is a discussion of the standards, indicating that the participating math teacher-leaders were consistent with the InTASC Standards. One conclusion is that math teacher-leaders have a lower than expected ability to value diverse learners in valuing diverse learners.

Keywords: math teachers; teacher dispositions; teacher-leaders; Rasch model; urban education

1. Introduction

For national accreditation purposes in the United States, educator preparation programs (EPPs) are required to demonstrate that their graduates are competent in the Interstate New Teacher Assessment and Support Consortium (InTASC) Standards developed by the Council of Chief State School Officers (CCSSO, 2013). These standards establish the knowledge, skills, and dispositions found to be important across the 50 states, and include a set of critical dispositions. The

* Corresponding author: William Lang, lang@usf.edu

accreditation review is conducted by the Council for the Accreditation of Educator Preparation (CAEP, 2016), and the requirement to use the InTASC Standards is found in CAEP's first standard. This study is part of a series of studies (Lang, Moore et al., 2018) to investigate the assessment of teacher dispositions and the implications of the findings. In this case, the effort studies the dispositions of mathematics teacher-leaders, pilots a new instrument, and offers improvement to staff development.

2. Rationale and Background

2.1 The Importance of Educator Dispositions

Educators are in a position in which the expectation is to provide support to students that covers a wide range of diversity, abilities, beliefs, attitudes, and values. Culturally responsive pedagogy (CRP) is a method of focus towards not only instruction but interpersonal interactions with students to better gauge how to support them in a way that is culturally inclusive (Warren, 2017). In urban schools, there are challenges and barriers in which a traditional method of teaching would not particularly aid in a student's academic journey. The work of CRP is crucial in assisting educators by recognizing the potential negative outcomes present because of a standardized set of behaviors educators may abide by (Warren, 2017). With the knowledge that students in urban schools are typically subjected to trauma and otherwise toxic environments, empathy and acknowledgement of experiences go a long way in the efficacy of an educator. Empathy is the vehicle in which educators take part in understanding the sociocultural aspect of the students they are entrusted to help guide. In this way, educators are developing an interpersonal relationship with students that embodies trust, as opposed to punishment for atypical behavior. This method is important in constructing a more open and safe learning environment for students, particularly marginalized groups.

Educator dispositions can be explained as trends which encompass the local school setting (Warren, 2017). If students are subjected to more toxic environments, such as harsh verbal and non-verbal behaviors that disconnect them from engagement, coping may become more difficult and lead to a range of negative outcomes that affect the student and community alike. An educator's values, beliefs, and attitude are integral to the articulation of responses to culturally diverse students. In urban schooling, educators' values are uniquely important for the 80% of minority students in the United States who have majority teachers (Sleeter, 2016).

Educators that are not part of a specific culture tend to face obstacles in navigating responses and treatment to students that are diversely different from the experiences of the educator themselves. For instance, if a White educator is imagining themselves in the position of one of their students of color, the educator may rely on their experiences and simply attribute the blame for poor performance or behavior on the student rather than taking into consideration the unique experiences of the student. Evidence points to success when educators affirm racial diversity viewpoints, take the time to educate themselves on the

pathways in which students of color learned most effectively, and maintain knowledge of the students and community in which they teach (Warren, 2017).

2.2 Influence of Educator Disposition on Mathematics

Educator disposition is essential for the facilitation of a safe and beneficial learning environment, which is particularly true when considering students of color. As mentioned by Tan (2017), when educators' attitudes shift towards students with different learning preferences, mathematics becomes more inclusive and engaging. Mathematics curricula for students of color are most beneficial when there is a collaboration between educators and students and their parents to construct the most beneficial and understandable subject matter for everyone (Tan, 2017). Since educator dispositions largely rely on the personal experiences and beliefs of the individual educator (Warren, 2017), room is left for misaddressing obstacles, or behaviors, by students that do not have similar experiences to their teachers. Evidence has shown that educators with affirmative viewpoints on racial diversity educate themselves on pathways most effective for Black students to learn and continued to gain knowledge and understanding of students and the community, thereby helping students achieve more (Warren, 2017). The disconnect in relation to educator dispositions based on personal experience may become a large obstacle for educators in the field. Mathematics is often taught with low-cognition assignments for certain students, rather than assignments that highlight strengths and preferences of individual students. A collaborative effort by educators, students, and parents is essential when devising a plan for the neediest student population (Tan, 2017). The need for restructuring urban schools is rooted in the notion that all students, regardless of color, socioeconomic status, or disability, deserve to be provided with a safe, beneficial, and overall quality education.

2.3 Teacher Preparation is Pivotal

An original sociology study by Barone (2006) indicated that teacher quality is integral to the students' academic success. The U.S. Department of Education (Snyder & Dillow, 2015) revealed that Black/African American students continue to exhibit learning gaps that require intensive intervention. Schools should also incorporate stronger strategies to close the racial-achievement gap in mathematical computations. Unfortunately, the prominent issue of teacher persistence remains and will continue to remain until specific attention is given to enhance preparation programs and schoolboard onboarding processes that evaluate individual teacher dispositions. Researchers have argued that when preparation programs focus on dispositions, schools and districts are more likely to sustain and retain such teachers (Williams et al., 2016); those teachers are more committed and more likely to persist (Lang et.al, 2016). When teachers feel a connection to their students, families, and community in which they teach, they are more likely to continue to review, revise, and develop their dispositions as teacher-leaders and learners, thus reducing the massive exodus experienced in most urban schools (Lang, Moore et al., 2018). According to Williams et al. (2016), there remains a disparity between theory and practice. Therefore, deliberate attention must be placed on teacher dispositions. Williams et al. (2016) found that *"[c]ritical to becoming an effective mathematics teacher for diverse learners is a combination of content knowledge, dispositions, and pedagogical knowledge that will*

support student' mathematical thinking and learning as well as their home culture" (p. 18).

2.4 The Importance of Teacher Dispositions in Urban Schools and Classrooms

A disposition is a habitual tendency to behave or react in a certain way to a certain stimulus due to a certain belief. A disposition may also be thought of as someone's values or character. Many teachers, or people for that matter, would tell you that context does not matter and that they would reach the same conclusion to a situation no matter what circumstances are in play. However, countless studies over time have shown that people's intentions and their actions often differ (Truscott & Obiwo, 2021). A recent study has shown how context is an important defining factor in how future preservice teachers and teachers alike react to situations in the urban school classroom.

Through CRP, the Math Teacher Leader Institute (PCS Teaching, 2018) partnered teacher-leaders and school-based administrators with University of South Florida (USF) and trained school district experts. These USF and trained district personnel guided the teacher-leaders and administrators through a dispositional survey process, as captured by Lang, Moore et al. (2018). This study has yielded very promising results that suggest that CRP can be used as a tool that could guide future generations into a more understanding and respectful society through education about culture and understanding. CRP acts as a lens for preservice and in-service teachers, teacher-leaders, and administrators which examines their beliefs and dispositions about practice. Between mentorship and real-world application, pre- and in-service teachers, teacher-leaders, and administrators can learn and immediately apply the information. One of the greatest takeaways from the Math Teacher Leadership Initiative (MTLI) participants consisted of awareness of their personal dispositions and how to use CRP as a mathematical learning tool. Teacher-leaders and administrators also reported that they were able to recognize that not all urban schools are the same and that their students do not respond the same way to the same types of practices and professional development previously learned (Truscott & Obiwo, 2021).

2.5 Math Teacher Dispositions Related to Students with Disabilities

Math teacher dispositions are important when instructing Black students in urban educational environments (specifically Black students with disabilities). A negative attitude can lead to anxiety about the subject, convincing the student that math may not be meant for them (Paul, 2021). The National Center for Learning Disabilities (2020) conducted an online study to gain an idea of what learning disabilities the general population knows about and the learning disabilities the population had not heard of as much. Ninety-one percent of the respondents indicated the belief that a learning disability makes it difficult for the individual learner. However, 66% of these respondents had never heard of such challenges such as dyscalculia (Paul, 2021). Dyscalculia is a learning disability that makes it difficult for the individual to solve arithmetic problems. It is unique to arithmetic and is like dyslexia in the sense that arithmetic operations become impossible for the individual to complete. Research in this field has suggested that students are able to be successful in mathematics when their needs are accounted for in the lesson planning.

Teacher-leaders that are unclear of how their personal dispositions, beliefs, and attitudes affect student learning may see the use of a calculator as a crutch or cheating mechanism (PCS Teaching, 2018). They may believe that students need to know the computations without the use of a calculator. Most importantly for students with dyscalculia, it is important to break down the computations into smaller steps and provide individual feedback, helping the students and guiding them exactly where they went wrong. For students in urban schools, this process may take more time and effort from the teacher-leader. Teacher-leaders who understand their students' beliefs and values also understand that giving partial credit is more beneficial than no credit for equations that were partially correct but failed to reach the end number. This explains that success comes in steps and with practice.

Students in urban schools often become frustrated with teachers that do not understand or attempt to understand their cultural beliefs, values, and attitudes towards math (National Center for Learning Disabilities, 2020). Students with dyscalculia may get frustrated after trying for some time and give up, so it is important for math teachers to model each step exactly as well as each step where the student went wrong. All these practices benefit not only individuals with dyscalculia but also those with other learning exceptionalities and mathematical anxieties. Therefore, math teacher dispositions are crucial in instructing students with exceptionalities.

As it relates to urban schooling, there remains an urgent need to acknowledge and address the role of teacher and student dispositions, and the inequities experienced by urban schools that ultimately affect the students and communities' job market, social statuses, and income (Carlson, 2017). In addition to disparities in relation to race, there are also disparities in the way urban schools educate Black students in general education and exceptional education. No longer can stakeholders allow the school district and policy-makers to require and implement professional development that typically restricts and segregates teachers from their beliefs, values, and attitudes. These historical dispositional disconnects between teachers and students only perpetuate school failure and an increase in the likelihood of prison (Annamma et al., 2018).

2.6 Teacher Accreditation and Dispositions

The Council for Accreditation of Educator Preparation (CAEP, 2022) dictated assessment that *"relies on relevant, verifiable, representative, cumulative, and actionable measures to ensure interpretations of data are valid and consistent"* (Standard 5.2, Data Quality). Five CAEP Standards include critical dispositions. Standard 1 includes the InTASC Standards adopted by the CCSO (CCSSO, 2013). InTASC has 10 standards, and each standard includes a list of critical dispositions. CAEP Standards 2, 3, and 4 demand assessment of dispositions as part of accreditation and require evidence of valid and reliable measures. The early development of assessments of dispositions indicated that validity required more than one instrument and typically involve different item types (Wilkerson & Lang, 2007). Originally, institutions assessed dispositions with an observation form. Some older, simple observation instruments include the Eastern Teacher Dispositions

Index (Singh & Stoloff, 2008) and the Clinical Experience Rubric (Flowers, 2006). These efforts suffered psychometrically.

The CAEP Standards validation of the use of multiple measures combined into a single calibrated ruler for assessing teacher candidates has been demonstrated (Wilkerson, 2012). Brindle (2012) also recommended the use of different assessment devices to measure dispositions, such as giving students feedback, using self-reflection, and prescriptive remediation. The Disposition Assessment Aligned with Teacher Standards (DAATS) battery has always included multiple measures (Lang et al., 2016; Lang, Moore et al., 2018; Lang, Wilkerson, Gilbert et al., 2018; Lang, Wilkerson, Moore, & Fields, 2018; Wilkerson & Lang, 2006), noting that in a parallel to knowledge and skills, one should not depend on a single moment in time or score.

The current effort pilots a new edition of disposition assessments originally called DAATS. The DAATS battery measures the consistency of teachers and teacher candidates to the InTASC-based dispositions of teaching and was originally proposed in a book by Wilkerson and Lang (2007). The original version of DAATS was built on an earlier edition of the InTASC Standards but without the emphasis on a taxonomy. The earlier DAATS battery had well-established construct and predictive validity and excellent reliability (Wilkerson & Lang, 2006) but only one form of each instrument. Later studies included rater accuracy when judges (scorers) use the Rasch model (Rasch, 1960/1980) of item response theory (IRT) (Lang et al., 2014) to estimate judge error.

3. Disposition Assessment Aligned with Teacher Standards v2 (DAATS2)

The implications of dispositions for teacher education (Wadlington & Wadlington, 2002) have not changed from conceptualization decades ago, but assessment is still lacking. The first DAATS was a battery of five instruments developed based on the 1992 InTASC Standards (Wilkerson & Lang, 2007). The five item types include: 1) a Thurstone scale (Thurstone, 1928) called Beliefs About Teaching (BATS), 2) an apperception projective called the Situational Reflection Assessment (SRA), 3) a children's focus group called the K-12 Impact Dispositions Survey (KIDS), 4) an observation report form called Candidate Dispositions Checklist (CDC), and 5) a guided reflection called Experiential Teaching Questionnaire (ETQ). All instruments in the DAATS battery measure the consistency between teachers and the critical dispositions of the InTASC Standards. When the InTASC Standards were revised (CCSSO, 2013), the DAATS battery became a target for revision. In addition to standards alignment, the new DAATS2 instruments have the same item types, but the revised instruments have at least two forms for each instrument and short forms for time-utility. As may be expected, the short forms sacrifice some psychometric power in return for utility. In this study, BATS2 Form B was combined with the short form of ETQ2 Form B for the assessment of math teacher-leaders.

3.1 Beliefs About Teaching Scale 2 (BATS2)

BATS2 is a Thurstone agreement scale of 50 items, constructed to measure each of the 10 InTASC Standards along a continuum of commitment based on a modification of the original affective taxonomy (Krathwohl et al., 1973; Wilkerson & Lang, 2011). Responses are scored using the Rasch model with Winsteps software (Linacre, 2023). Evidence of the validity and reliability has been presented previously for both versions of the scale. (See, for example, Englehart et al., 2012; Lang et al., 2014; Lang, Wilkerson, Gilbert et al., 2018; Lang, Wilkerson, Moore, & Fields, 2018; Wilkerson, 2012.)

The two forms allow multiple measurements without repeating the same items. All items were developed from the InTASC critical dispositions statements. Each item was aligned with both Krathwohl levels and critical dispositions. For example, Standard 2, Learning Differences, includes: *“The teacher respects learners as individuals with differing personal and family backgrounds and various skills, abilities, perspectives, talents, and interests.”* The BATS2 items ask a student to *agree/disagree* with statements such as:

- I usually think about children’s home life and environment so that I can tell if something is wrong. (Taxonomy level: valuing)
- I have a rule in my classroom: “We all speak proper English and ignore gestures, slang, or foreign languages.” (Taxonomy level: unaware)

BATS1 and BATS2 use a Thurstone (Linacre, 1998) format of agree/disagree items. Thurstone’s items are a forced-choice, dichotomous decision (agree/disagree only), while Likert items are a continuum creating a rating scale, typically five points, from *strongly agree* to *strongly disagree*, with a neutral midpoint. Roberts et al. (1999) studied the connection between Likert and Thurstone agreement scaling, recommending the Thurstone scale model when extreme positions (e.g., high/low levels of commitment) are of interest.

Every item is coded for analysis by InTASC Standards (10), Core Area (4), and Krathwohl Taxonomy (6). A modification of the taxonomy (Wilkerson & Lang, 2011) classifies student affect into six levels: unaware, receiving, responding, valuing, organizing, and characterizing, since the original taxonomy was designed for instruction and not assessment, omitting the possibility that respondents might have no commitment. Examples of the levels are presented in Table 1.

Table 1: Modified Krathwohl taxonomic levels used by DAATS instruments

DAATS taxonomic level	Typical teaching behaviors at each level
Unaware	Has not considered the value in any meaningful way
	May be opposed to the value
Receiving	Recognizes or is aware of the importance
	Is beginning to consider the value, but not acting on the value
	May promise to apply the value without evidence of doing so
Responding	Tries to act on values
	Expresses commitment or satisfaction to acting on the value

Valuing	Accepts the value and expresses personal or collective positivity for related actions
	Commits to the value despite interferences or required effort
Organizing	Plans and schedules to ensure application of the value
	Expresses the relationship between knowledge/skill and the value
	Takes a leadership role with others concerning the value
Characterizing	The value is at the center or driving force for many life efforts
	Expresses or lobbies for the value even in the face of social or institutional opposition

3.2 Experiences in Teaching Questionnaire 2 (ETQ2)

The ETQ2 is a guided reflection like those often used in clinical settings in education. The ETQ2 includes short-answer questions aimed at the critical dispositions listed in the InTASC Standards. The items can also be grouped into four categories of standards identified by InTASC: 1) the learner and learning, 2) content knowledge, 3) instructional practice, and 4) professional responsibility. There are two parallel forms of the revised ETQ2 (A and B) with ten items in each form representing the 10 InTASC standards. The ETQ2 can also be given as a short form by choosing one item from each of the four categories of InTASC for a four-item instrument. The ETQ2 is scored by judges, and it is harder to fabricate responses compared to the Thurstone agreement BATS2, which is self-report. Responses are scored using a rating scale based on the Krathwohl affective taxonomy. Each response to items is assessed from low (unaware or receiving) to high as to the quality level of values, with the expectation that average practicing teachers will score at the valuing level (3). Ratings of 4 and 5 (organizing and characterization) are rare among preservice teachers.

The ETQ2 takes more time to score than BATS2 or the CDC2, but, as a constructed response assessment it provides assessment of dispositions that is hard to fake, and a view of what the teachers really believe. Individual questions can be used to finetune the evaluation of a teacher on an individual InTASC standard (Wilkerson & Lang, 2012). In this case, we used a short form of ETQ2B with four reflective items with one item representing each of the four InTASC categories.

For example, an InTASC critical disposition is: "1(h) The teacher respects learners' differing strengths and needs and is committed to using this information to further each learner's development." The associated ETQ2 reflective question would be "Think about a lesson that did not work as you had hoped it would. Did the students perform lower than you expected as a group or just several? Did you deal with some students individually? If it happens again, would you do something differently?" Typical scores for responses to this item are shown in Table 2.

Table 2: Example scoring for an item in ETQ2, Form A

Taxonomic level	Paraphrased sample response from teacher-leaders
Unaware	My students performed lower than expected because they weren't paying attention, so I had to repeat the instructions several times
Receiving	I did not set goals or expectations for the class at the beginning of the lesson, so that's something that I should do differently for next time
Responding	Students performed lower than wanted. As remediation for these students, I would work individually with them, and I would aim my instruction towards their individual learning styles.
Valuing	I remember a lesson that did not go over as I expected. I concluded that it was a subset of students who struggled with some ideas. To remediate, I created groups based on their problems and worked with these students individually and in small groups to ensure mastery. When this occurs again in the future, I will use the strategy of prescriptive learning and monitor progression closely.
Organizing	My students always enjoy group projects. Working together with other teachers and having students see this and creating lesson plans for them that integrate learning is always something that I do in my classroom, whether it's a small project or a big project like a PBL [Performance Based Learning]. Collaboration with others brings in a positive vibe in my classroom. Then students push aside their differences, and it shows through their work.
Characterizing	My lesson plans are very much geared toward different types of learners, because I teach very high-level learners, and I also teach a lot of students with learning challenges. So my lessons are almost always geared toward all types of learners, with extra help provided during study hall for some of the students. We do team planning at my school, and I have altered lessons for my class to be small group and differentiated to better support individual groups of my students. My school district provides models for each lesson and standard. I will read over each module and lesson accompanying the provided material. If differentiation or modification is necessary, I will do so to meet the needs of my student. Additionally, I will work to create more engaging lessons.

4. Math Teacher-Leaders

The teacher-leader project described in this paper was developed as a strategic partnership between an institution of higher education and a southern urban school district. Thirty-two passionate educators in one large urban school district applied and were to participate in a new district-wide initiative based on readiness to fully dedicate their energy to the program. Administrators applied and were expected to fully participate in every school. Three teachers were selected from each school by the principal based on the given criteria. Our target population was all elementary schools that do not receive regular mathematics coaching. MTLI members are expected to act as on-site math support from the

classroom for schools that typically do not receive much coaching support. Schools were selected based on their request to participate each year and on administrator enthusiasm and school readiness. Teacher-leaders were selected by principals based on these criteria: respected by peers, demonstration of high-quality teaching practices, use of data to inform instruction, and interest in participation.

The MTLI (PCS Teaching, 2018) was developed to support and replicate strategies of current school-based reformers. These educational reformers act from a set of values that guide their professional decision-making. As passionate individuals who persist, they place high professional development expectations on themselves, often finding ways, outside the school district, to get what they need. The teacher-leaders in the MTLI provide mentoring to others, while staying focused on students and their learning. The MTLI focuses on teachers who demonstrate resilience, are able to take charge, solve problems, and find opportunities which may positively add to the school in important ways that support student achievement, increase teacher leadership, and strengthen the school's overall success.

The goal of the MTLI is to develop strategic mathematics leadership teams at elementary schools in a southeastern (USA) school district to improve the knowledge, skills, dispositions, teaching, and communication of teacher-leaders and school-based administrators who have demonstrated the ability to work effectively with students and colleagues in their own classroom and beyond. University professors, district math supervisors, and math coaches collected data while investigating strategies used by urban teachers to build their personal leadership dispositions. Thirty-two persistent teacher-leaders and administrators from one urban district reported student achievement equal to or higher than the state average on standard tests of mathematics.

The administrative leaders focused on the importance of their teacher-leaders' professional development experiences. This was done while cultivating each participant's expertise and desire to communicate, collaborate, and develop a consensus while co-creating new learning experiences - which would be interwoven into their teaching and reflective skills. Administrative leaders (principals and assistant principals) provide ongoing, direct, facilitated support to ensure that the role of the teacher-leader is not downplayed within the overall school setting. When teacher-leaders and school-based administrative leaders have clear converging views about the MTLI goals and the initiative's connection to the school vision and mission, the MTLI would be more efficacious.

5. Research Design and Analysis

This research design has two parts that are connected. The first involves the use and psychometric analysis of new versions of two DAATS instruments aligned with the InTASC Standards. This part of the research examines the validity and reliability of the instruments through IRT. The second part comprises examination of the resulting scores from in-service math teachers to validate responses but to

also plan improvement in the programming of staff training for MTLI participants.

5.1 Instrument Analysis

The ability of individuals (in this case the consistency of teacher-leaders with the InTASC dispositions) and the difficulty of items (the value level of each item) influence each other. This is called conjoint measurement. IRT models envision a ruler of knowledge, skills, or dispositions that can be measured concurrently (or conjointly) without regard to a particular sample. The units of measurement are called “logits” and the scores are called “measures”. One of the most well-known uses of Rasch modeling is the Lexile Reading Scale (Stenner & Wright, 2004). Bond and Fox (2007) summarized the model and uses of this type of item analysis.

This analysis using Winsteps software was completed in a single calibration for both people and items. This is described by Linacre (2003, p. 18) as:

$$\text{Log} \{ P_{nij} / P_{ni(j-1)} \} = B_n - D_i - F_{j-1}$$

Where

P_{nij} is the probability that person n encountering item i is observed in category j ,

B_n is the “ability” measure of person n ,

D_i is the “difficulty” measure of item i ,

F_j is the “calibration” measure of category J relative to category $j-1$.

In this study, the analysis followed the guidelines provided in Smith and Wind (2018). The raw data were calibrated using the Rasch model rating scale and incorporating Winsteps software. IRT differs from classical test theory, which is sample dependent, while the Rasch model is not.

5.2 Sample

Two hundred and eighty-one ($N = 281$) math teacher-leaders were administered the BATS2 Form B and the ETQ2 short form simultaneously. All the participants were in-service teachers or administrators in a large urban school district in the southeastern United States. The sample was selected through convenience sampling, with the participants representing a wide span of experience, prior education, and demographics. The participants agreed to participate, and the research met all relevant human subject considerations.

6. Findings

6.1 Quantitative Results

All analyses were generated using Winsteps software (Linacre, 2023) (see Tables 3 and 4). In the initial calibration of a modest sample size ($N = 281$), the real item separation = .98, which suggests that the scale discriminates between the persons adequately. The real person separation = .69 (Cronbach alpha = .71), indicating a reasonably defined variable. The outfit MNSQ = .99 (expected value = 1.0) and outfit ZSTD = -.11 (expected value = 0.0), indicating that the data variability fit the Rasch model.

Table 3: Initial Rasch model calibration: Summary of 281 measured persons

	Total score	Count	Measure	Standard error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
Mean	42.0	54	66.96	3.58	.95	-.18	.99	-.11
SEM	.3	0	.41	.01	.03	.07	.04	.06
P. SD	5.5	0	6.84	.13	.50	1.22	.71	1.02
S. SD	5.5	0	6.85	.13	.51	1.22	.71	1.02
Max.	58.0	54	86.77	3.98	3.53	4.45	5.21	4.04
Min.	21.0	54	42.96	3.23	.42	-1.68	.19	-1.54

Model RMSE = 3.58; true SD = 5.83; separation = 1.63; person reliability = .73
Cronbach alpha (KR-20) person raw score "Test" reliability = .71
MNSQ is the mean-square fit statistic with expectation = 1
ZSTD is the MNSQ standardized to a theoretical normal with mean = 0 and mean = 1

Table 4: Initial Rasch model calibration: Summary of 54 measured items

	Total score	Count	Measure	Standard error	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD
Mean	218	281	50.0	2.02	.98	.10	.99	.27
SEM	7.5	0	2.15	.14	.01	.18	.04	.23
P. SD	54.6	0	15.67	1.04	.11	1.34	.27	1.66
S. SD	55.1	0	15.82	1.05	.11	1.35	.27	1.68
Max.	278.0	281	86.96	5.84	1.23	6.33	1.85	6.01
Min.	81.0	281	19.21	.81	.64	-2.73	.36	-2.20

An examination of the Krathwohl scoring scale indicates ordered thresholds as expected. There were no scores in the highest rating (5; characterizing) in this sample. The results are shown in Table 5, confirming the expected category structure and thresholds.

Table 5: Summary of category structure for judged ETQ2 items

Category label	Score	Average expected	Infit MNSQ	Outfit MNSQ	Andrich threshold	Category measure
0	0	-20.3	.63	.80	None	-17.41
1	1	-16.6	.61	.41	10.48	-8.99
2	2	-13.0	.77	.55	-20.77	-2.58
3	3	-9.42	.75	.56	-1.88	6.83
4	4	-6.39	1.15	1.07	12.17	24.70

An examination of the mean scores by InTASC Standards (see Table 6) indicate that the math teacher-leaders scored least consistent with Standard 2 (Learning Differences), with a scale value of 34.08, and most consistent with Standard 10 (Leadership and Collaboration), with a scale value of 65.84.

Table 6: Comparison of math teacher-leaders on InTASC Standards: Extreme and non-extreme item scores

Item count	Mean measure	S.E. mean	P. SD	Median	Model separation	Model reliability	RSM E	True SD	Mean outfit	Code
54	49.25	2.25	16.36	16.51	48.05	3.09	.91	5.04	15.56	*
9	50.50	5.55	15.71	16.66	50.51	3.65	.93	4.15	15.15	1
6	34.08	5.14	11.49	12.59	33.57	1.26	.61	7.14	9.01	2

5	45.20	4.83	9.67	10.81	45.12	2.13	.82	4.11	8.75	3
5	58.81	5.34	10.68	11.94	57.25	4.02	.94	2.58	10.36	4
5	45.50	8.38	16.75	18.73	40.48	3.08	.90	5.18	15.93	5
6	52.42	4.25	9.51	10.42	50.81	2.79	.89	3.21	8.95	6
5	57.56	6.93	13.87	15.51	60.25	4.61	.96	2.94	13.55	7
6	38.56	6.75	15.10	16.54	40.58	1.50	.69	8.39	12.55	8
3	51.89	15.20	21.50	26.33	65.57	3.36	.92	6.12	20.61	9
4	65.84	6.61	11.44	13.22	63.66	4.94	.96	2.27	11.22	10

A pairwise test of differences between InTASC Standards indicate significant contrasts between InTASC Standards 2 and 10; 8 and 10; 2 and 4; 2 and 6; 2 and 7; 4 and 8. The results of this analysis are shown in Table 7.

Table 7: Comparison of math teacher-leaders on InTASC Standards

In TASC Standard contrast	Mean diff	SE	t	df	Prob.
1 2	16.42	7.57	2.17	12	.051
1 3	5.30	7.36	.72	11	.487
1 4	-8.31	7.70	-1.08	10	.306
1 5	5.00	10.05	.50	7	.634
1 6	-1.92	7.00	-.27	12	.788
1 7	-7.06	8.88	-.79	8	.450
1 8	11.94	8.74	1.37	10	.202
1 9	-1.39	16.18	-.09	2	.939
2 3	-11.12	7.06	-1.58	8	.154
2 4	-24.73	7.41	-3.34	8	.010*
2 5	-11.41	9.83	-1.16	6	.290
2 6	-18.34	6.67	-2.75	9	.023*
2 7	-23.48	8.63	-2.72	7	.030*
2 8	-4.48	8.49	-.53	9	.610
2 9	-17.81	16.05	-1.11	2	.383
3 4	-13.61	7.20	-1.89	7	.101
3 5	-.30	9.67	-.03	6	.977
3 6	-7.22	6.44	-1.12	8	.295
3 7	-12.36	8.45	-1.46	7	.187
3 8	6.64	8.30	.80	8	.447
3 9	-6.69	15.95	-.42	2	.716
4 5	13.32	9.93	1.34	6	.229
4 6	6.39	6.83	.94	8	.377
4 7	1.25	8.75	.14	7	.891
4 8	20.25	8.61	2.35	8	.047
4 9	6.92	16.11	.43	2	.709
5 6	-6.93	9.39	-.74	6	.489
5 7	-12.07	10.87	-1.11	7	.304
5 8	6.93	10.76	.64	8	.537
5 9	-6.39	17.36	.37	3	.737
6 7	-5.14	8.14	-.63	6	.551
6 8	13.86	7.89	1.74	8	.121
6 9	.54	15.78	.03	2	.976
7 8	19.00	9.68	1.96	8	.085
7 9	5.68	16.71	.34	2	.766
8 9	-13.32	16.63	-.80	2	.507

10	1	15.33	8.63	1.78	7	.119
10	2	31.75	8.37	3.79	6	.009*
10	3	20.64	8.19	2.52	5	.053
10	4	7.02	8.50	.83	6	.440
10	5	20.34	10.67	1.91	6	.105
10	6	13.41	7.86	1.71	5	.149
10	7	8.27	9.58	.86	6	.421
10	8	27.27	9.45	2.89	7	.023*
10	9	13.95	16.57	.84	2	.489

* $p < .05$

6.2 Qualitative Results

Below are four contrasting responses as examples with a subset of teacher-leader comments taken from their ETQ2 reflections (edited to protect identity) that illustrate the differences from lower consistency to higher consistency with InTASC. These examples come from Category 4, Standard 10. The ETQ2 prompt in this case was: "Describe your participation in your school and professional community. Can you name a time that you worked with families of your students? Did different beliefs or backgrounds create challenges? How do you interact with colleagues and families to enhance practice or support students?"

"All the time, daily. I make it a point to know all my students and their families. I am aware of their homelife situations and their needs. It comes with the territory of my job and the population of the students I work with. I must build positive relationships with my students and their families in order to be successful, otherwise I will get nowhere. Relationships are key with my students and families." (Rating 3, valuing)

"I message and meet with families as needed. Every year, I have one or two families that need a lot of interaction, which I always do. Struggled a little with one student whose mother didn't speak English this year." (Rating 1, receiving)

"I hosted a multi-cultural fair at our STEAM night. I had several families volunteer to work at one of the countries represented. I didn't have any challenges, except I didn't plan for an adopted student from China to work at the China station. I felt foolish for not automatically placing him there, but made a quick change so he could. I have a multi-cultural club at my school that meets twice a month during lunch. We plan culturally/socially significant school-wide activities and events to promote tolerance. This year, I invited ALL of our ESOL students to participate in an effort to get them involved in a club to promote leadership." (Rating 4, organizing)

"Attending community events or sports events in my neighborhood or the neighborhood of my families." (Rating 0, unaware)

7. Conclusions

7.1 Disposition Instruments

One purpose of this research effort was to pilot a new forms of the BATS2 self-report scale (Thurstone items) and a short form of the ETQ2 (reflection). The Rasch item analysis estimated separation reliability and fit statistics within expected parameters, and no items in this sample appeared misfitting. This supports the scale use for consistency with the InTASC Standards as measures. In addition, the item estimates ranged in difficulty, while the person scores appeared normally distributed. This reinforced use of the BATS2 with samples of in-service teachers as DAATS was intended. As such, there is a strong conclusion of support for the instrument's psychometric qualities for similar samples.

The qualitative examination of responses in the ETQ2 reflection revealed several conclusions. The judges were able to score responses on the modified Krathwohl taxonomy. This pilot did not perform an analysis of judge error, which would be a suggestion for future research. Nonetheless, nothing in the results from this sample indicated misfitting or disordered categories due to rating confusion. At worst, an analysis might reveal that some judges may have been harsher or more lenient than others. Again, the conclusion would be that the ETQ2 worked as expected to assess teacher consistency with the InTASC critical dispositions.

7.2 Math Teacher Leadership Initiative (MTLI) Dispositions

There are provocative results in the assessment of the math teacher-leaders in that there was lower consistency of dispositions related to Learning Differences and higher consistency with Leadership and Collaboration. The math teacher-leaders were in-service teachers, principals, and assistant principals. The higher scores for Standard 10 (Leadership) might be expected, but the lower ratings for Standard 2 (Learning Differences) were problematic. The critical dispositions of interest are indicated below:

- 2(l) The teacher believes that all learners can achieve at high levels and persists in helping each learner reach his/her full potential.
- 2(m) The teacher respects learners as individuals with differing personal and family backgrounds and various skills, abilities, perspectives, talents, and interests.
- 2(n) The teacher makes learners feel valued and helps them learn to value each other.
- 2(o) The teacher values diverse languages and dialects and seeks to integrate them into his/her instructional practice to engage students in learning.

It is likely that an additional follow up with more detailed assessments would be useful, not to mention some thinking about interventions. Clearly, there needs to be some thinking about the learner and differing learners as a target for development or intervention on the dispositions of math teacher-leaders. Generally, Standard 2 on learners would be a strong point for experienced teachers. The math teacher-leaders were surprisingly less consistent with InTASC on this standard than desirable. It is unknown whether this is a characteristic of

the content specialty (math), the sample of teachers from an urban setting, or some other variable. The results offered an immediate target for discussion in the MTLI as the results were revealed to participants.

7.3 Implications

The assessment of teacher dispositions is still relatively new in teacher preparation. It remains to be seen if university programs, accreditation, or certification will emphasize such measures in the future, but any proposals will clearly require valid and reliable instruments. That is one of the primary reasons that creating and piloting such devices is a precursor to training changes or research on effective teaching regarding dispositions. Even in this one pilot with a homogeneous sample, there were findings that could alter the planned training and served as eye-opener to participants.

8. References

- Annamma, S., Ferri, B., & Connor, D. (2018). Disability critical race theory: Exploring the intersectional lineage, emergence, and potential futures of DisCrit in education. *Review of Research in Education*, 42(1), 46–71. <https://doi.org/10.3102/0091732X18759041>
- Barone, C. (2006). Cultural capital, ambition and the explanation of inequalities in learning outcomes: A comparative analysis. *Sociology*, 40(6), 1039–1058. <https://doi.org/10.1177/0038038506069843>
- Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). LEA.
- Brindle, S. (2012). *An exploratory study on the assessment of pre-service teacher dispositions by teacher education programs in Iowa* [Doctoral dissertation, Drake University, Iowa].
- Carlson, D. (2017). *Teachers and crisis: Urban school reform and teachers' work culture*. Routledge.
- Council for the Accreditation of Educator Preparation (CAEP). (2022). *The CAEP Standards*. Washington, DC. <https://caepnet.org/standards/2022-itp/introduction>
- Council of Chief State School Officers (CCSSO). (2013). *InTASC model core teaching standards and learning progressions for teachers 1.0*. <https://ccsso.org/resource-library/intasc-model-core-teaching-standards-and-learning-progressions-teachers-10>
- Englehart, D. S., Batchelder, H. L., Jennings, K. L., Wilkerson, J. R., Lang, W. S., & Quinn, D. (2012). Teacher dispositions: Moving from assessment to improvement. *The International Journal of Educational and Psychological Assessment*, 9(2), 26–44. http://tijepa.books.officelive.com/Documents/V9.2_TIJEPA.pdf
- Flowers, C. (2006). Confirmatory factor analysis of scores on the Clinical Experience Rubric: A measure of dispositions of preservice teachers. *Educational and Psychological Measurement*, 66, n3, p 478–488.
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1973). *Taxonomy of educational objectives: The classification of educational goals. Affective domain* (handbook II, vol. 1). David McKay Co., Inc.
- Lang, W. S., Moore, L., Wilkerson, J. R., Parfitt, C. M., Greene, J., Kratt, D., Martelli, C. D., LaPaglia, K., Johnston, V., Gilbert, S., & Zhang, J. (2018). Beliefs About Teaching (BATS2): Construction and validation of an instrument based on InTASC critical dispositions. *International Journal of Learning, Teaching and Educational Research*, 17(8), 56–77. <https://doi.org/10.26803/ijlter.17.8.4>
- Lang, W. S., Wilkerson, J. W., Rea, D. C., Quinn, D., Batchelder, H. L., Englehart, D. S., & Jennings, K. J. (2014). Measuring teacher dispositions using the DAATS battery:

- A multifaceted Rasch analysis of rater effect. *Journal of Applied Measurement*, 15(3), 240–215.
- Lang, W., Wilkerson, J., Gilbert, S., Wang, C., Kratt, D., Parfitt, C., Martelli, M., Zhang, J., Fields, L., LaPaglia, K., Greene, J., & Johnston, V. (2018). *Measuring teacher dispositions systematically using multiple measures* [Panel session]. Annual meeting of the Eastern Educational Research Association, Clearwater, Florida, February 9, 2018.
- Lang, W., Wilkerson, J., Moore, L., & Fields, L. (2018). *Beliefs About Teaching (BATS2): Assessment of InTASC dispositions* [Meeting paper]. Annual meeting of the American Association of Colleges of Teacher Education, Baltimore, MD.
- Lang, W., Wilkerson, J., Moore, L., Fields, L., Parfitt, C., Greene, J., Kratt, D., Martell, D., LaPaglia, K., Johnston, V., Gilbert, S., & Zhang, J. (2016). *Beliefs about teaching (BATS2): Construction and validation of an instrument based on InTASC critical dispositions* [Research paper]. Annual meeting of the Eastern Educational Research Association, Clearwater, Florida, February 9, 2018.
- Linacre, J. M. (1998). Thurstone thresholds and the Rasch model. *Rasch Measurement Transactions*, 12(2), 634–635. <https://www.rasch.org/rmt/rmt122j.htm>
- Linacre, J. M. (2003). *A user's guide to Winsteps/Ministep Rasch-Model computer programs*. Winsteps.com
- Linacre, J. M. (2023). Winsteps [Computer program Ver. 3.93.2]. winsteps.com
- National Center for Learning Disabilities. (2020). *Significant disproportionality in special education: Current trends and actions for impact*. <https://www.nclld.org/sigdispro/>
- Paul, M. (2021). Teaching elementary mathematics to students with disabilities: Strategies for instruction. *Central College 1853*. <https://central.edu/writing-anthology/2019/04/30/473/>
- PCS Teaching. (2018, February). *Math Teacher Leader Institute* [YouTube video]. https://www.youtube.com/watch?v=UeG_kSqt0WA
- Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests* (Copenhagen, Danish Institute for Educational Research) (expanded ed. [1980; foreword & afterword Wright, B. D.]). The University of Chicago Press.
- Roberts, J. S., Laughlin, J. E., & Wedel, D. H. (1999). Validity issues in the Likert and Thurstone approaches to attitude measurement. *Educational and Psychological Measurement*, 59(2), 211–233. <https://doi.org/10.1177/00131649921969>
- Singh, D. K., & Stoloff, D. L. (2008). Assessment of teacher dispositions. *College Student Journal*, 42(4), 1169–1180.
- Sleeter, C. (2016). Critical race theory and the Whiteness of teacher education. *Urban Education*, 52(2), 155–169. <https://doi.org/10.1177/0042085916668957>
- Smith, R. M., & Wind, S. A. (2018). *Rasch measurement models: Interpreting Winsteps and Facets output* (2nd ed.). JAM Press.
- Stenner, A. J., & Wright, B. (2004). Uniform reading and readability measures. In B. D. Wright, & M. H. Stone (Eds.), *Making measures* (pp. 161-181). Phaneron.
- Tan, P. (2017). Building inclusive mathematics classrooms for students with disabilities. *For the Learning of Mathematics*, 37(3), 21–24. [file:///C:/Users/User/Downloads/Building_Inclusive_Mathematics_Classroom%20\(1\).pdf](file:///C:/Users/User/Downloads/Building_Inclusive_Mathematics_Classroom%20(1).pdf)
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33, 529–554. <https://doi.org/10.1086/214483>
- Truscott, D. M., & Obiwo, S. M. (2021). School-based practices and preservice teacher beliefs about urban teaching and learning. *Peabody Journal of Education*, 96(1), 54–64. <https://doi.org/10.1080/0161956X.2020.1864246>
- Snyder, T.D., and Dillow, S.A. (2015). *Digest of Education Statistics 2013* (NCES 2015-011). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

- Wadlington, E., & Wadlington, P. (2011) Teacher dispositions: Implications for teacher education. *Childhood Education, 87*(5), 323–326.
<https://doi.org/10.1080/00094056.2011.10523206>
- Warren, C. (2017). Empathy, teacher dispositions, and preparation for culturally responsive pedagogy. *Journal of Teacher Education, 69*(2), 169–183.
<https://doi.org/10.1177/0022487117712487>
- Wilkerson, J. R. (2012). Measurement and evaluation perspectives on scaling teacher affect with multiple measures. *The International Journal of Educational and Psychological Assessment 9*(2), 165–191.
http://tijepa.books.officelive.com/Documents/V9.2_TIJEPA.pdf
- Wilkerson, J. R., & Lang, W. S. (2006). Measuring teaching ability with the Rasch model by scaling a series of product and performance tasks. *Journal of Applied Measurement, 7*(3), 239–259.
- Wilkerson, J. R., & Lang, W. S. (2007). *Assessing teacher dispositions: Five standards-based steps to valid measurement using the DAATS model*. Corwin Press.
- Wilkerson, J. R., & Lang, W. S. (2011). Standards-based teacher dispositions as a necessary and measurable construct. *The International Journal of Educational and Psychological Assessment, 7*, 34–54.
- Williams, D. L., Edwards, B., Kuhel, K. A., & Lim, W. (2016). Culturally responsive dispositions in prospective mathematics teachers. *Discourse and Communication for Sustainable Education, 7*(2), 17–33. <https://doi.org/10.1515/dcse-2016-0013>