The DAATS Battery Short Form as a Measure of Teacher Dispositions

Judy R. Wilkerson*, Florida Gulf Coast University
jwilkers@fgcu.edu
W. Steve Lang, University of South Florida
LaSonya Moore, University of South Florida

*Corresponding author

The DAATS (Dispositions Assessments Aligned with Teacher Standards) battery is a series of five instruments of different item types that measure teachers' consistency with the critical dispositions embedded in the InTASC Standards. The purpose of this study was to continue a 20-year research project on the development and implementation of instruments in the battery that can be used for various purposes. Although a problem exists in defining and measuring teacher dispositions, this phase of the research works toward a shortened solution for building a composite single instrument, defined by the InTASC standards. Research questions centered on the validity and reliability of alternate forms of the instruments along with potential use of a shortened form of the battery. As a quantitative study, dichotomous and polytomous responses were scaled using the Rasch model of item response theory. Results indicate that the DAATS short form maintains evidence of validity and reliability and provides promise for a variety of short forms of the battery. In conclusion the DAATS battery can be useful for making decisions about teachers' dispositions that can lead to opportunities for improvement at the student, course, and program levels. Continued testing of other short forms and improvement opportunities are recommended.

Keywords: Affective; Measurement; Teacher education

Introduction

The research presented herein argues that teacher educators are responsible not only for ensuring that teachers have both the knowledge and skills required of effective teachers but also that teachers are equally committed to using those skills in practice in their own classrooms. While assessments abound for teaching and testing knowledge and skills to diverse learners, far less attention is paid to the assurance of teacher commitments, often referred to as teacher dispositions. As a result, teachers are typically well trained and assessed in writing lessons and other critical job-related skills, but assessments of their commitment to continuing what they have learned in their own classrooms are far less common. This research explores the problem of defining and measuring teacher affect, also known as teacher dispositions, with a particular focus on the nationally accepted teacher standards used in accreditation, the use of a variety of measurement methods, and scaling results based on a taxonomy using modern measurement techniques. The twofold purpose of this study is to demonstrate that a shortened version of the multi-instrument DAATS battery maintains evidence of validity and reliability while providing potential for use as a pre/post test to demonstrate changes in teacher dispositions upon completion of coursework aimed at cognitive goals.

Literature Review

Problems in Defining Teachers' Affect or Dispositions

Given the strong correlation between teacher dispositions and the quality of their students' learning, measurement of pre-service teacher dispositions is a crucial part of teacher preparation programs and, as a result, has become an important part of their preparation programs (Bradley, 2020; Choi et al., 2016; Phelps, 2006).

Part of the problem in assessing teacher dispositions is the long-term general difficulty in operationalizing the affective domain; there is no commonly accepted predominant definition of the term "dispositions" (Johnston et al., 2018). As one might expect, this lack of definitional clarity leads to ambiguity (Kinderwater, 2013), subjectivity (Meidl & Baumann, 2019), and varying approaches to measurement across teacher preparation programs (Choi et al., 2008; Pufpaff et al., 2017). More recently, Fonseca-Chacana (2019) attempted a definition of dispositions as the set of nurtured academic, internal, and social qualities that influence preservice teacher knowledge and skills, which contribute to a professional community of colleagues, students, and families, but this definition does not appear widely in the literature.

Various researchers have described dispositions as internal attributes or psychological characteristics that motivate action, or as a tendency to act in a certain manner that is predictive of future action (Borko et al., 2007; Villegas, 2007). There does appear to be consensus that certain situations result in teachers making a specific choice to act or react in a certain way, and they often make choices among various action options that are driven by their own internal beliefs (Hollon et al., 2010).

Consensus issues aside, all teacher preparation programs in the U.S. that seek national accreditation by the Council for Accreditation of Educator Preparation programs (CAEP, 2022) are required to demonstrate that their graduates have demonstrated effective dispositions. In the CAEP accreditation standards, these dispositions, as well as correlated knowledge and performances, are defined through the InTASC (Interstate Teacher Assessment Support Consortium) Standards promulgated by the Council of Chief State School Officers (CCSSO, 2013).

The InTASC standards provide the content standards embedded in the CAEP accreditation standards. The standards were written by a consortium (the InterState Teacher Assessment Support Consortium or INTASC) formed by the Council of Chief State School Officers (CCSSO)– the secretaries and commissioners of education of all 50 states. The Council charged the consortium with developing the standards, which are referred to as the InTASC Standards. The InTASC Standards include equal attention to performances (also called skills), essential knowledge, and critical dispositions (CCSSO, 2013). There are four categories and ten standards:

Category 1: The Learner and Learning
- #1: Learning Development
- #2: Learning differences
- #3: Learning Environments

Category 2: Content
- #4: Content Knowledge
- #5: Application of Content

Category 1: The Learner and Learning
- #6: Learning Development
- #7: Learning differences
- #8: Learning Environments

Category 2: Content
- #9: Content Knowledge
- #10: Application of Content

The InTASC definition of dispositions includes the "habits of professional action and moral commitments that underlie the performances and play a key role in how teachers do, in fact, act in practice" (CCSSO, 2013, p. 6).

Another key point is that these standards maintain the delineation of knowledge, dispositions, and performances as a way to probe the complexity of the teacher's practice. The relationships among the three have been reframed, however, putting performance first—as the aspect that can be observed and assessed in teaching practice. The others were renamed. "Essential knowledge" signals the role of declarative and procedural knowledge as necessary for

effective practice and "critical dispositions" indicates that habits of professional action and moral commitments that underlie the performances play a key role in how teachers do, in fact, act in practice (p. 8).

The multi-decade debate about the definition of teacher dispositions continues to this day. It includes whether the definition of dispositions should be standards-based or morality-based (Bullough, 2023, Dottin & Sockett, 2006, Lang & Wilkerson, 2024; Wilkerson, 2006; Wilkerson & Lang, 2007). Much of this debate has occurred in the literature sponsored by the American Association of Colleges of Teacher Education, with the most recent non-standards-based argument appearing in their pre-eminent journal, *The Journal of Teacher Education*. In that journal, Bullough (2023) suggested that dispositions be rethought as "Virtue and the Manners of Democracy." While these are clearly important aspects of character, they do not address the national standards promulgated by the Council of Chief State School Officers, the standards that drive national accreditation (CCSSO, 2013; CAEP, 2023). The research described herein (and over the years) applies the standards-based definition of dispositions -- the InTASC Standards.

Whatever definition one applies to dispositions, the question of measuring them remains a significant problem. Even if one accepts the standards-based definition, there is a large number of critical dispositions identified, and they are detailed, spanning behaviors for both beginning and advanced teachers, organized into four categories with ten standards. A developmental structure that demonstrates increasing levels of commitment is needed to build a measurement process.

Problems in Measuring Teachers' Affect or Dispositions

To build a process to measure teacher dispositions effectively, three things are necessary: a definition of construct (e.g., the InTASC Standards), a progression of demonstration (e.g.,

Krathwohl's 1956 Affective Taxonomy, an integral part of the original Bloom Taxonomy), and one or more measurement methods. In terms of the progression of demonstration, the Taxonomy remains a neglected framework (Lang & Wilkerson, 2024).

The Taxonomy is comprised of three domains -- cognitive, affective, and psychomotor. Affective objectives describe values, appreciations, adjustments, attitudes, and interests (Krathwohl et al., 1956). The Taxonomy was designed as a communication tool for teachers to use to support organization and implementation of classroom teaching objectives, and it classifies objectives in each domain along a continuum. As such, it provides the basis for sorting teachers into levels of commitment that can lead to a planned progression of increasing commitment. However, of the three domains, Bloom (1956) noted that the affective domain was the most difficult to complete, given that the objectives are less precise and more difficult to operationalize, since they represent internal or covert feelings that are difficult to describe and observe.

While the standards can provide the content and the Taxonomy can provide the levels, there remains yet another need – the need for tools or instruments that withstand the tests of validity and reliability. As Byrd (2023) noted, significant gaps exist in understanding a teacher's professional dispositions in practice and definition as well as the best method to assess dispositions, with applications differing from one teacher preparation program to the next and with success frequently questionable. The approaches of using observations, Likert items, and confirmatory factor analysis have proven highly unsatisfactory (Niu et al., 2017).

The literature does provide a few examples of single assessments of teacher affect such as surveys, indices, observations, or interviews, (Richardson & Onwuegbuzie, 2003; Lund et al.; Schulte et al., 2004; Wasicsko, 2004; Jung & Vogt, 2006; Singh & Stoloff, 2008). Two decades

ago, one of the first dispositions measures was proposed (Schulte et al., 2004). More recently, West et al. (2020) validated a *Teacher Disposition Scale*. However, each of these instruments is a stand-alone instrument, leading to a very narrow assessment of dispositions that is subject to problems associated with the instrument type applied. For example, observational instruments typically suffer from a halo effect, while self-reports often are influenced by respondents' desire to put themselves in the best possible light, leading to conclusions that may not be borne out in reality. Surprisingly, the use of multiple measures with solid psychometric properties remains limited despite the inclusion of dispositions measurement as a fundamental component of teacher accreditation (CAEP, 2023) and despite the common use of multiple measures of cognitive competence within teacher preparation programs and the accompanying emphasis on using cognitive measures to foster program improvement.

As a result, there is a paucity of well-developed and tested measurement devices (Katz & Raths, 1986; Lang & Wilkerson, 2024; LaPaglia, 2020; Wilkerson, 2012; Wilkerson & Lang, 2007). Teacher preparation programs continue to attempt to use what is available with a focus on observation instruments (Griffin, 2022; Seay, 2021), but they actually need to assess dispositions based on a successful (and even mandated) construct definition (i.e., InTASC Standards), time-tested levels of demonstration (i.e., the Bloom Taxonomy), and modern measurement theory. The instruments developed and described here represent a practical and theoretical attempt to follow a basic set of conditions appropriate for modern measurement with tools such as the Rasch models of Item Response Theory (Mauri et al., 2023).

Without well-developed instruments, meticulously developed based on a well-defined construct, it is difficult to use assessment results in meaningful ways. These include making determinations about what and how much students learned in the affective domain and what

relationship affective and cognitive development have with each other. LaPaglia (2020) determined that preservice teacher dispositions, as defined by the InTASC Standards, improved during undergraduate students' program of cognitive study even when there was no defined curriculum dedicated to dispositions training and development. LaPaglia (2020) used the DAATS battery to reach that conclusion, but it is limited to undergraduate students.

Research Purpose

The DAATS battery has demonstrated multiple uses in teacher education including international comparisons (Wilkerson et al., 2020); program improvement (Englehart, et. al., 2012); program admissions (Carter, et. al, 2011); teacher performance (Wilkerson & Lang, 2008); and accreditation (Wilkerson & Lang, 2009). The research described herein continues the development and application of a battery of assessments of teacher dispositions. The long-term research design is driven by a four-part conceptual framework:

- The dispositions assessed are those that are defined by national professional standards.
- Use of an accepted taxonomy to frame levels of construct attainment is crucial to decision-making.
- No single measure or item type is sufficient, so multiple measures (or at least a composite form of condensed multiple measures) are required.
- The measurement process should use modern measurement techniques.

In addition to the above four elements of the conceptual framework, as in many tests, there is a need for multiple forms of the test. In this current phase of the research, the development of multiple forms of the Beliefs About Teaching Scale (BATS) were tested for comparability, allowing for future use of the scale as a pre-post test. Second, since scoring the instruments that require students to write a response to a prompt takes time, a shortened form of

the test that maintains psychometric quality is needed, so a short form of one of the instruments, Student Reflection Assessment (SRA), combined with the BATS forms, was tested.

Research Questions

1. Is there evidence that the Forms A and B of the *Beliefs About Teaching Scale* (BATS2) are equally valid and reliable and potentially useful for pre-post testing?
2. Can BATS2 be used diagnostically for individual InTASC Standards and Categories?
3. Is the Krathwohl Taxonomy evident in the *Situational Reflection Assessment* (SRA2) item analysis?
4. What are the psychometric properties of a short form of SRA2 (4 items only), when analyzed separately as a measure of teacher dispositions?

Method

Instrumentation: The DAATS Battery: *Dispositions Aligned with Teacher Standards*

The *Dispositions Assessments Aligned with Teacher Standards* (DAATS) battery (Lang & Wilkerson, 2006) was developed and described in a five-step standards-based model describing a process for affective instrument design (Wilkerson & Lang, 2007). In 2024, the model was expanded to nine steps that incorporated additional attention to measurement issues while summarizing previous validation studies (Lang & Wilkerson, 2024). The research presented herein contributes to the long-term study of a systematic, standards-based approach to assessing teacher dispositions.

The DAATS battery is comprised of the following five instruments, all using different item types and all aligned with the InTASC Standards. The instruments requiring a narrative response are scored using the Krathwohl Taxonomy, and the dichotomous items are written to reflect different levels of the Taxonomy. Modern measurement theory, the Rasch Model of item

response theory (Rasch, 1960), is used to build the reporting scale. For item development and candidate rating purposes, the taxonomic levels are defined in Figure 1, with responding to valuing serving as the expected levels for entry into the profession and valuing to organizing serving as the expected levels for practicing teachers. Note that the original taxonomy was designed for learning and did not provide for a total lack of learning, so the "unaware" level was added to the version used in this work. The interpretation applied is represented in Figure 1.

Figure 1.

Definition of Taxonomic Levels.

Taxonomic Levels	Typical Teaching Behaviors at Each Taxonomic Level
Unaware	Has not considered the skill in any meaningful way. May be opposed to the skill.
Receiving	Recognizes (is aware of) importance. Is beginning to think about it. May provide a promise to use it without evidence of having used it.
Responding	Is emotionally ready to do something and makes an attempt. Gives a little extra effort, as time permits, to comply. Can easily be distracted from application. Has a beginning level of commitment or satisfaction.
Valuing	Accepts worth and derives definite satisfaction from it. Feels a need and would commit continuing time and effort. Tolerates and may expect interferences.
Organization	Plans, organizes, and schedules to ensure success with it. Determines inter-relationships among knowledge and skills. Adapts other aspects to fit it. Is uncomfortable with interferences or lack of time to finish.
Characterization	Sees the skill as the center or driving force of all work. Helps others to see the skill's importance, lobbying for it. Integrates everything with it.

The DAATS Battery is comprised of five instruments, the first two of which were used in this research:

- BATS2: Beliefs About Teaching Skills (Thurstone Agreement Scale)
- SRA2: Situational Reflection Assessment (Projective Apperception)

- ETQ2: Experiential Teaching Questionnaire (Constructed Response – Reflective Essays)
- CBC: Candidate Behavior Checklist (Performance Observation)
- KIDS: K-12 Impact Dispositions Scale (Focus Group)

BATS2 is a 50-item Thurstone agreement scale, with two response options for each item (agree or disagree). Items are scored as consistent or inconsistent with the InTASC Standards, and there are ten items for each standard. The items tap varying levels of the Krathwohl taxonomy, predominantly at the unaware and valuing levels. Two examples are provided here and are drawn from one of the critical dispositions for InTASC Standard 2, *Learner Differences,* which reads: "The teacher respects learners as individuals with differing personal and family backgrounds and various skills, abilities, perspectives, talents, and interests. One item with an expected "agree" response and another with an expected "disagree" response follow:

- I usually think about the children's home life and environment so that I can tell if something is wrong ("Agree" is the expected response, so if a student agrees, we would classify the student as "valuing" in the Krathwohl Taxonomy).
- I have a rule in my classroom: We all speak proper English (The expected response is "disagree," so if the student agrees with this item, we would classify the student as "unaware.")

SRA2 is a 20-item projective, also called an apperception instrument, which asks students to describe what they see in a picture. The pictures were drawn intentionally to be ambiguous by a professional artist (Barbara Slitkin, NYC) under our direction. They are intended to elicit a response showing either negative or positive perceptions about students, teachers, or learning. The pictures are accompanied by question prompts; respondents describe their perceptions,

typically about one paragraph long. Responses are scored from zero to five, representing unaware to characterizing, based on the Krathwohl Taxonomy. Although each picture is intended to target one of the ten InTASC Standards. In the results section, two prompts along with representative responses are provided.

Sample

For BATS2, graduate students were tested at one public university in Florida over five years (Form A n=947; Form B n=575). For SRA2 graduate students (n = 47) were administered the four item short form.

Scoring

BATS uses a Thurstone (1928) format for items. Thurstone's technique requires a dichotomous decision (agree/disagree only), while Likert provides for a rating scale, typically five-points, from strongly agree to strongly disagree with a neutral midpoint. Roberts et al. (1999) examined the relationship between Likert and Thurstone agreement scaling, recommending the Thurstone scale when extreme positions (e.g., high/low levels of commitment) are of interest. In the case of teacher dispositions, high levels of commitment are the norm, but low levels are of particular interest for diagnostic purposes. Scores are interpreted as "consistent" or "inconsistent" with the InTASC Standards. SRA is scored using a rating scale based on the Krathwohl Taxonomy as explained above, with scores ranging from zero to 5.

Results

BATS2: Research Question 1

The first research question asked: Is there evidence that BATS2 (Forms A&B) are equally valid and reliable and potentially useful for pre-post testing?

The variable (or Wright) map from Winsteps (Linacre, 2024) software is provided in Figure 2 and illustrates the distribution of person commitment (left) and item difficulty (right). At the bottom are the least committed persons and the easiest items. At the top are the most committed persons and the most difficult items. Items are coded by the item number, form number, pre- or post-test, standard, "C", category, and response (A or D). Notice that our results are normally distributed and not highly skewed, as is commonly found in many observation assessments of distributions. The top item difficulty is approximately even with the highest scoring students.

BATS items are relatively easy to guess. The large number of BATS items at the bottom of the map (the easiest items) confirms that expectation. The preponderance of items cluster near the center with a few being particularly difficult. BATS item 5 was the most difficult item. It reads: "In today's classroom, every single lesson should have an assessment, and every assessment result must be recorded and analyzed." The expected response is "agree," although they may have been expressing their concern with the notion of recording and analyzing the data. Another difficult item was number 22, "Usually, the teacher has to plan and practice questions, making changes on the spot, to get the most learning out of students." These teachers may have been worried about practicing at their career stage.

By contrast, the easiest items on the test supported the commitment of these students to assessment and planning. It read "It is important that I assess my students with lots of different types of measures and use the results to plan instruction." Another easy item was number 4: "After I've taught a lesson several times, I don't need to plan that topic anymore," and they tended to disagree with this item, as expected, having learned the value of continued planning.

Items are coded by the item number, form number, pre or post-test, standard, "C" for category, category number, and response (A or D). The most difficult item was number 12 from Standard 9, category 4: "Planning for improvements is the most important part of my job." The expected response is "agree," and, if a student answers "agree," then that student is likely at the organizing or characterizing levels in the Krathwohl Taxonomy, which is not expected because it is so extreme. In the middle is item 2, "It would be best for experts to provide lesson plans that teachers could follow instead of wasting time with all the teachers reinventing the wheel." A teacher who values planning would want some input into lessons.

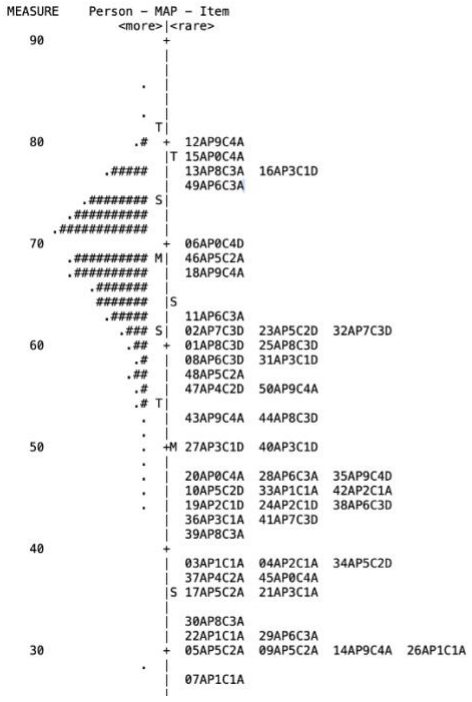
Theoretically, when candidates and items are next to each other on the map, the difficulty of the item and the ability of the candidate are comparable, so the candidate has approximately a 50% probability of answering the item correctly. The items at the top of the map were probably answered correctly by about 30% of the candidates who are the most able. The items at the bottom of the map are the very easy items and were probably answered correctly by over 90% of the candidates. Those items are well below the ability of the least able candidate, indicating that all candidates have a **greater** than 50% probability of answering the items correctly. For dichotomous items, there should be a nearly even spread of items along the variable (the y-axis) with no gaps. Gaps can indicate poorly defined or poorly tested regions of the variable. Good tests usually have the items targeted (lined up with) the persons (Linacre, 2024).

In Figures 2 and 3, the two variable maps presented for Forms A and B, the respondents are lined up with the items and there are no large gaps between groups of items, so it can be concluded that the construct of teacher dispositions was well measured in this sample of relatively homogeneous respondents, thereby supporting a claim to validity. Form A was

administered to students at the beginning of their program; Form B to another set of students at the end of their program. The two sets of students are different.

Figure 2

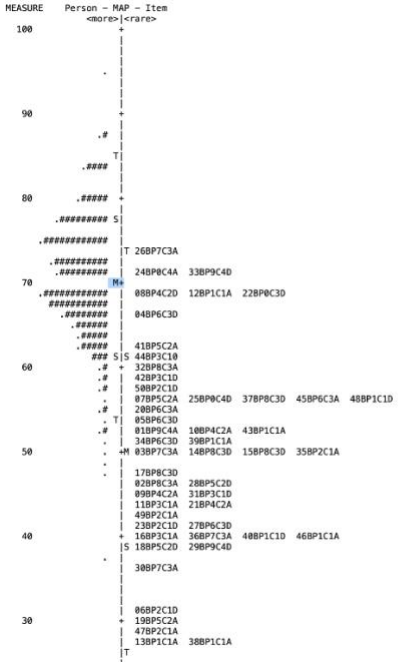
Rasch Variable Map BATS Form A Pre-Test



In Figure 3, administered at the end of students' programs, performance is higher and even more normally distributed, as reflected on the left side of the map. The group dispositions have become more consistent with InTASC.

Figure 3

BATS Form B – Post-Test



Figures 4 and 5 provide the separation tables for the two forms of the test. The item reliability (similar to Cronbach's alpha) on both is .67 and .74, respectively, supporting the reliability of the two forms. This indicates that the scale discriminates between persons well, even though the expected sample is homogeneous. Note that the Real Item Separation reliability of .94 and .99 (tables not shown) respectively indicate a well-defined variable (Smith & Wind, 2018).

Figure 4

Rasch Separation Table for BATS2, Form A

SUMMARY OF 947 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	39.0	50.0	68.19	4.17	.99	.04	.92	.00
SEM	.1	.0	.22	.02	.01	.04	.02	.03
P.SD	4.3	.0	6.79	.48	.24	1.10	.53	1.05
S.SD	4.3	.0	6.80	.48	.24	1.10	.53	1.06
MAX.	47.0	50.0	86.42	10.23	2.04	5.34	5.96	5.65
MIN.	1.0	50.0	3.32	3.38	.54	-2.42	.19	-1.86

```
| REAL RMSE   4.35 TRUE SD    5.22  SEPARATION  1.20  Person RELIABILITY  .59 |
|MODEL RMSE   4.20 TRUE SD    5.34  SEPARATION  1.27  Person RELIABILITY  .62 |
| S.E. OF Person MEAN = .22                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .67  SEM = 2.49
```

Figure 5

Rasch Separation Table for BATS2, Form B

```
     SUMMARY OF 575 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                          MODEL          INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      40.5      50.0       69.59    4.30      1.00    .11    .92    .01 |
|  SEM        .2        .0         .32     .04       .01    .04    .02    .04 |
| P.SD       4.9        .0        7.76     .89       .18    .88    .42    .90 |
| S.SD       4.9        .0        7.77     .89       .18    .88    .42    .90 |
| MAX.      49.0      50.0       95.26   10.26      1.75   4.75   2.95   4.42 |
| MIN.       5.0      50.0       23.02    3.22       .58  -2.50    .20  -1.94 |
|-----------------------------------------------------------------------------|
| REAL RMSE   4.51 TRUE SD    6.32  SEPARATION  1.40  Person RELIABILITY  .66 |
|MODEL RMSE   4.39 TRUE SD    6.40  SEPARATION  1.46  Person RELIABILITY  .68 |
| S.E. OF Person MEAN = .32                                                   |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = .97 (approximate due to missing data)
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .74  SEM = 2.48
```

BATS2: Research Question 2

The second research question asked: Can BATS2 be used diagnostically for individual InTASC Standards and Categories? Tables 1 and 2 present the measures for each of the InTASC Standards and categories. The easiest InTASC standards, with the lowest mean measures of 34.39 and 43.25 respectively, were Learner Development and Learning Differences (Standards 1 and 2). These two standards are a significant focus in the curriculum, making the results of high commitment to these two standards reasonable. Conversely, the most difficult standards were in Professional Learning/Ethical Practice and Leadership/Collaboration, two standards for which little instruction or opportunity to demonstrate are provided. Teamwork is generally an issue among the students. When analyzed separately, Category 4 yields similar results as the most difficult category. Experience supports the quantitative results and the claim to validity.

Table 1

Mean Measures for Each InTASC Standard

Standard #	Standard Name	Mean Measure
1	Learner Development	34.39
2	Learning Differences	43.25
3	Learning Environments	52.81
4	Content Knowledge	46.53
5	Application of Content	46.06
6	Assessment	53.43
7	Planning for Instruction	55.30
8	Instructional Strategies	54.36
9	Professional Learning and Ethical Practice	55.56
10	Leadership and Collaboration	57.62

Figure 6

Winsteps Output for BATS 2, Form A

Item COUNT	MEAN SCORE	MEAN COUNT	MEAN MEASURE	S.E. MEAN	P.SD	S.SD	MEDIAN	MODEL SEPARATION	MODEL RELIABILITY	RMSE	TRUE SD	MEAN OUTFIT	CODE
50	738.3	947.0	50.00	2.07	14.48	14.63	46.58	11.69	.99	1.23	14.43	.92	*
4	615.3	947.0	57.65	9.63	16.67	19.25	57.95	15.84	1.00	1.05	16.64	1.08	0
5	900.2	947.0	34.39	3.41	6.81	7.62	31.18	3.58	.93	1.83	6.56	.80	1
4	859.5	947.0	43.25	1.74	3.02	3.49	44.28	2.33	.84	1.19	2.78	.74	2
6	717.2	947.0	52.81	5.78	12.93	14.16	50.58	12.36	.99	1.04	12.89	.88	3
2	807.5	947.0	46.53	9.34	9.34	13.21	46.53	7.75	.98	1.20	9.26	.87	4
8	779.8	947.0	46.06	5.31	14.06	15.03	42.04	10.03	.99	1.39	13.99	.85	5
6	696.8	947.0	53.43	6.31	14.10	15.45	52.77	12.68	.99	1.11	14.06	1.00	6
3	698.7	947.0	55.30	6.31	8.93	10.93	61.46	9.83	.99	.90	8.88	.86	7
6	685.8	947.0	54.36	6.30	14.10	15.44	56.63	13.09	.99	1.07	14.05	1.07	8
6	661.2	947.0	55.56	6.92	15.46	16.94	54.44	13.71	.99	1.12	15.42	1.01	9

Table 2

Mean Measures for Each InTASC Category

Category #	Standard Name	Mean Measure
1	The Learner and Learning	44.12
2	Content Knowledge	44.16
3	Instructional Practice	54.17
4	Professional Responsibility	56.40

Figure 7

Winsteps Output for BATS 2, Form A

Item COUNT	MEAN SCORE	MEAN COUNT	MEAN MEASURE	S.E. MEAN	P.SD	S.SD	MEDIAN	MODEL SEPARATION	MODEL RELIABILITY	RMSE	TRUE SD	MEAN OUTFIT	CODE
50	738.3	947.0	50.00	2.07	14.48	14.63	46.58	11.69	.99	1.23	14.43	.92	*
15	816.1	947.0	44.12	3.24	12.11	12.54	43.65	8.65	.99	1.39	12.03	.82	1
10	785.3	947.0	46.15	4.42	13.25	13.97	42.04	9.71	.99	1.36	13.18	.86	2
15	692.8	947.0	54.17	3.54	13.25	13.71	58.92	12.50	.99	1.06	13.20	1.00	3
10	642.8	947.0	56.40	5.33	15.99	16.85	54.44	14.56	1.00	1.10	15.95	1.04	4

SRA2: Research Question 3

Research question 3 asked: Is the Krathwohl Taxonomy evident in SRA2 scoring? For Category 1, The Learner and Learning, a prompt entitled *Walking to School* was used and is presented in Figure 8, along with selected responses representing all levels in the Taxonomy.

Figure 8

Sample #1 of an SRA Prompt and Responses

The questions asked were:

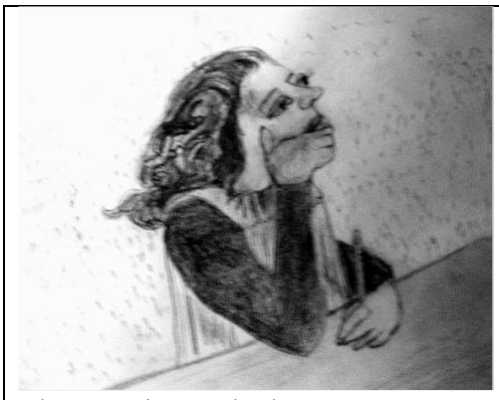
- What kind of teacher would be best to teach this child?
- What would you do if this child were in your class?

- Unaware: … An ESOL teacher would be best.
- Receiving: …I would allow this child in my class with great caution.
- Responding: …I want this child to stay in my class so I can make sure they get help and an education.
- Valuing: … I would want this child in my class. There is no child I would not want in my class It would make my work as a teacher more rewarding.
- Organizing: …I would welcome any child with any home situation, behavioral issue, or disability to my classroom. They're all unique and we could help each other as the school year progresses. These students challenge you for the better and make for better teachers and they also need someone understanding, patient, and ready to stand in the gap for them and make sure they get the help they need.

For Category 3, Instructional Practice, a prompt entitled *Lost in Thought* was used and is presented in Figure 9, along with selected responses representing two levels of the Taxonomy.

Figure 9

Sample #2 of an SRA Prompt and Responses



The questions asked were:

- What lesson did you just teach?
- What strategies did you use? What did you just ask the students to do?
- Is this child doing what you expect?
- How can you tell? What would you say to her?
- Would you want her in your class? Explain.

Unaware: I just taught a mathematics lesson on fractions and wanted my students to do a worksheet after the lesson so they would have less homework. I used questions that were brainteasers, so the students felt more challenged than usual. *This child is not following my directions and is spacing off due to being bored most likely.* I can tell she is bored because she is not engaged in the lesson being taught. I would want her in my class because *I was the same way as a student, and I could give her advice.*

Valuing: This image shows a student that looks to be lost in thought. The lesson that was taught, was mostly likely a history lesson. Based on the picture, I can see that the student is thinking about the lesson, so it was most likely to be a heavy topic. Teaching strategies I used was to question students on the topic and to imagine themselves in the historical event. *Questioning the students like this would further their thinking beyond facts learned in class.* I would ask a question like how they would feel in this event and questions like that. I also would have students discuss with groups to compare how different people feel. *The child picture would be doing exactly what I expect. The child is independently thinking about what it would be like to be in the situation and how that would affect her, her family, and other people around or related to her.* I can tell this because she looks to be deep in thought. *I would ask her what she is thinking and ask her to share her thoughts.* I would want this student in my class because she is doing what is asked of her.

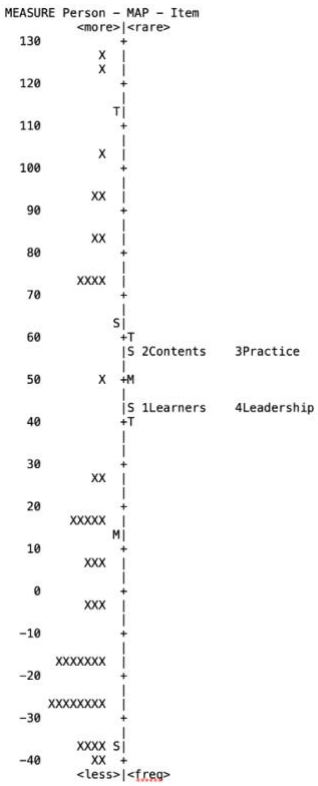
SRA: Research Question 4

Research Question 4 asked: What are the psychometric properties of a short form of SRA2 (4 items only) when analyzed separately as a measure of teacher dispositions?

Figure 10 provides a Wright map for the four SRA items, which are scaled in the middle of the map on the right with no gaps, indicating no gaps in the construct with this small number of items. The students are spread throughout the range of abilities on the left. The open-ended items representing the four InTASC categories are all at a neutral and similar difficulty, providing an opportunity for the range of scores seen on the left.

Figure 10

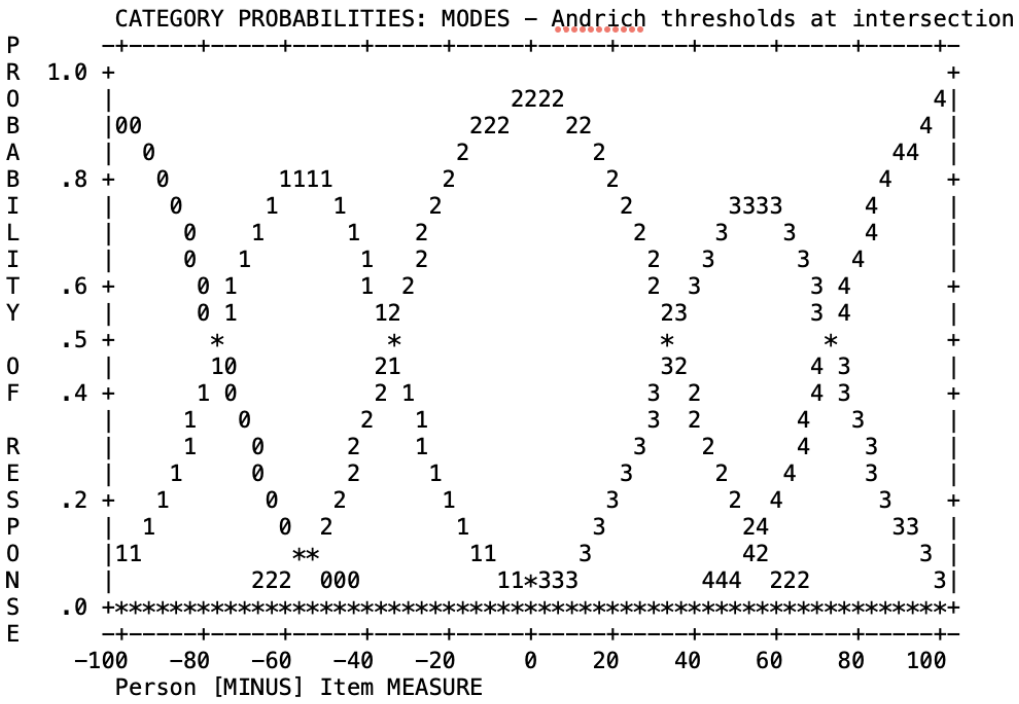
Map of Persons and Items for SRA 2, Short Form



In Figure 11, a Threshold analysis is provided. The four SRA2 prompts each yielded a range of responses that were scorable on the Krathwohl Taxonomy and were ordered as expected.

Figure 11

Threshold Analysis of SRA 2, Short Form



Conclusions, Implications, Recommendations, and Limitations

The results presented above indicate that the continued testing of two instruments in the DAATS battery, the Beliefs About Teaching Scale (BATS) and the Situational Reflection Assessment (SRA) continue to support claims of validity and reliability, while providing promise for use of the instruments in a pre- and post-test context through viable alternate forms. The Krathwohl Taxonomy is clearly visible in student responses to open-ended questions, as in SRA, and BATS items can clearly be aligned with the Taxonomy during the test construction phase. Variable maps provide support for construct validity through distributions that span the scale range, and students show higher measures after instruction than before.

This study supports prior claims that standards-based teacher dispositions can be measured effectively, with evidence of validity and reliability, even with a smaller number of items that maintain the commitment to the use of national standards, a recognized taxonomy, and the incorporation of multiple item types. Perhaps even more practically important, for those who

believe that teacher dispositions are important and can be influenced during teacher training, this study provides psychometric support that such influence is possible – even when not necessarily planned as an instructional goal.

Continued testing of the battery is recommended, including similar testing with ETQ as a short form. There also appears to be a potential for a short form that contains a sample of items from at least three of the instruments – BATS, SRA, and ETQ. As in past studies of the DAATS instruments, the homogeneity of the sample presents a scaling problem, since there is limited opportunity to measure students with no interest in teaching. Rasch measurement typically requires larger samples, but, given the context for these instruments, they perform well despite that expectation. Nonetheless, a more diverse sample would increase confidence.

References

Bloom, B. S., and Krathwohl, D. R., et al. (1956). *Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners.* Handbook I: Cognitive domain: Longmans, Green.

Borko, H., Liston, D., & Whitcomb, J. A. (2007). Apples and fishes: The debate over dispositions in teacher education. *Journal of Teacher Education*, *58*(5), 359-364. doi:10.1177/0022487107309977.

Bradley, E. & Isaac, P., King, J. (2020). Assessment of pre-service teacher dispositions. *Excelsior, Leadership in Teaching and Learning 13*(1). 50-62.

Bullough, R. V. (2023). Rethinking dispositions in teaching and teacher education: Virtue and the manners of democracy as a way of life. *Journal of Teacher Education (74)*4. https://doi.org/10.1177/00224871231154900.

Byrd, C. M. (2023). *Defining teacher dispositions: A phenomenological investigation of teacher evaluators.* Proquest (2901456326; ED633791).

Carter, C., Rea, D., Valesky, T., Wilkerson, J., & Lang, W. S. (2011). Development and analysis of survey instruments to assess education leadership candidates' dispositions. In *Leading in the decade of challenges and opportunities.* Southern Regional Council on Educational Administration 2011 Yearbook: Leading in the Decade of Challenges and Opportunities (E. Reames, Ed.). Auburn University.

Choi, H., Benson, N. F., & Shudak, N. J. (2016). Assessment of teacher candidate dispositions: Evidence of reliability and validity. *Teacher Education Quarterly*, *43*(3), 71-89.

Council of Chief State School Officers (2013). *InTASC Model Core Teaching Standards and Learning Progressions for Teachers 1.0*. Author.

Creasy, K. (2008). Teacher candidate disposition development and the concerns-based adoption model. *International Journal of Learning*, *15*(4), 277-284. doi:10.18848/1447-9494/CGP/v15i04/45698.

Dottin, E. S., & Sockett, H. (2006). A Deweyan approach to the development of moral dispositions in professional teacher education communities. *Teacher dispositions: Building a teacher education framework of moral standards*, 27-47.

Englehart, D., Batchelder, H., Jennings, K., Wilkerson, J., & Lang, W. S. (2012). Teacher Dispositions: Moving from assessment to improvement. *The International Journal of Educational and Psychological Assessment. 9*(2), 26-44.

Fonseca-Chacana, J. (2019). Making teacher dispositions explicit: A participatory approach. *Teaching & Teacher Education*, *77*, 266-276. doi:10.1016/j.tate.2018.10.018

Griffin, D. (2022). The assessment of preservice teachers' dispositions. *Excellence in Education Journal, (11)*1, 69-87.

Johnston, P., Wilson, A., & Almerico, G. M. (2018). Meeting psychometric requirements for disposition assessment: Valid and reliable indicators of teacher dispositions. *Journal of Instructional Pedagogies*, *21,* 1-12. ERIC Number: EJ1346078

Jung, E., Rhodes, D., & Vogt, W.P. (2006). Disposition assessment in teacher education: A framework and instrument for assessing technology disposition. *The Teacher Educator, 41*(4), 207-233.

Kinderwater, W. A. (2013). *The role of dispositions in teacher candidate education* (Publication No. 3568113) [Doctoral dissertation, University of Montana]. ProQuest Dissertations & Theses Global.

Krathwohl, D., Bloom, B. & Masia, B. (1956). *Taxonomy of educational objectives. Handbook II: Affective domain.* McKay.

LaPaglia, K. E. (2020). *Preservice teacher dispositions: A case study of changes during a teacher preparation program.* [Doctoral dissertation, Florida Gulf Coast University]. Proquest. Dissertations & Theses Global.

Lang, W. S., & Wilkerson, J. R., (2024a, February 8-9). *Analysis of two forms of the BATS2 (Beliefs About Teaching2) assessment of dispositions based on InTASC*. [Conference presentation]. Eastern Educational Research Association 2025 Conference, Clearwater, FL, United States.

Lang, W. S., Moore, L. S., & Wilkerson, J. R. (2024b, February 8-9). *Application and analysis of the SRA2 (Situational Reflection Assessment2) Short form for assessment of dispositions using a modified Krathwohl scale.* [Conference presentation]. Eastern Educational Research Association 2025 Conference, Clearwater, FL, United States.

Lang, W. S., & Wilkerson, J. R. (2006). *Measuring teacher dispositions systematically using InTASC Principles: Building progressive measures of dispositions.* [Conference

presentation]. American Association of Colleges of Teacher Education 2006 Conference, San Diego, CA, United States.

Lang, W. S., & Wilkerson, J. R. (2024). Measuring teacher dispositions: Steps in an innovative journey in affective assessment. In Fisher, Jr., William P. & Pendrill, Leslie (Eds.), *Models, measurement, and metrology extending the SI: Trust and quality assured knowledge infrastructures* (pp. 305-346), De Gruyter.

Lund, J., Wayda, V., Woodard, R. & Buck, M. (2007). Professional dispositions: What are we teaching prospective physical education teachers? *The Physical Educator, 12*, 38-47.

Mari, L., Wilson, M., & Maul, A. (2023). Fundamental concepts in measurement. In: Measurement Across the Sciences. *Springer Series in Measurement Science and Technology* (pp. 19-48). Springer. https://doi.org/10.1007/978-3-031-22448-5_2

Meidl, T., & Baumann, B. (2019). Extreme make over: Disposition development of preservice teachers. *Journal of Community Engagement & Scholarship*, *8*(1), 90-97.

Pufpaff, L., Sciuchetti, M., & Taylor, L. (2017). Dispositions by committee for educator preparation programs. *Teacher Educator*, *52*(3), 268-283. doi:10.1080/08878730.2017.1313923

Richardson, D., & Onwuegbuzie, A. (2003). *Attitudes toward dispositions related to the teaching of pre-service teachers, in-service teachers, administrators, and college/university professors*. (ED482689). ERIC. https://files.eric.ed.gov/fulltext/ED482689.pdf

Roberts, J. S., Laughlin, J. E., & Wedell, 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/10.1177/00131649921969811

Schulte, L., Edick, N., Edwards, S., & Mackiel, D. (2004). The development and validation of the Teacher Dispositions Index. *Essays in Education*, *12*, 1-16.

Seay, D. (2021). An examination of the correlation between a teacher candidate's disposition assessment from the college supervisor and the cooperating teaching during student teaching *Educational Research: Theory and Practice (32)*1, 48-53.

Singh, D., & Stoloff, D. (2008). Assessment of teacher dispositions. *College Student Journal, 42*(4), 1169-1180.

Villegas, A. M. (2007). Dispositions in teacher education: A look at social justice. *Journal of Teacher Education, (58)*5, 370-380. DOI: 10.1177/0022487107308419

Wasicsko, M. M. (2004). The 20-minute hiring assessment: How to ensure you're hiring the best by gauging educator dispositions. *The School Administrator, 61(9), 40.*

West, C., Baker, A., Ehrich, J., Woodcock, S., Bokosmaty, S., Howard, S. & Eady, M. (2020). Teacher Disposition Scale (TDS): Construction and psychometric validation, *Journal of Further and Higher Education*, 44:2, 185-200, DOI: 10.1080/0309877X.2018.1527022

Wilkerson, J. R. & Lang, W. S. (2004). Measuring teacher dispositions: An application of the rasch model to a complex accreditation requirement. [Conference presentation]. International Objective Measurement Workshop (IOMW) Cairns, Australia https://files.eric.ed.gov/fulltext/ED502872.pdf

Wilkerson, J. R. (2006). Measuring teacher dispositions: Standards-based or morality-based? *Teachers' College Record*. http://www.tcrecord.org/content.asp?contentid=12493.

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., (2009). *Report to the Louisiana Board of Regents Qualitative Research Team on measuring teacher dispositions with Wilkerson and Lang DAATS instruments.* Baton Rouge, LA, USA.

Wilkerson, J. R., Moore, L. L., Lang, W. S., & Zhang, J. (2020). Comparison of students in teacher education from China and the USA: An assessment of dispositions. *International Journal of Learning, Teaching and Educational Research.* 19 (11), 109-126, doi.org/10.26803/ijlter.19.11.7