

### **Purpose**

The purpose of the B-MTL instrument is to measure beliefs about mathematics teaching and learning through teacher self-report. Specifically, B-MTL is designed to measure the strength of conviction as it relates to three distinct beliefs: Transmissionist, Facts First, and Fixed Instructional Plan. Those three constructs are described in [Schoen and LaVenia \(2019\)](#).

The 2019 data were used in a multiyear randomized controlled trial that was designed to study of the effect of a teacher professional development program called Cognitively Guided Instruction (CGI; [Schoen et al., 2022](#)) on teachers, teaching, and students. Subsequent waves of data collection were planned for spring 2020, spring 2021, and spring 2022.

### **Previous Versions of the Instrument**

A total of 21 items were used on the B-MTL instrument. These items were identical to those described by [Schoen and LaVenia \(2019\)](#). A five-point, Likert-type scale was used for each item. The response categories were labeled: strongly agree, agree, undecided, disagree, and strongly disagree. Five of the 21 items were reverse-coded; four of which (BMTL18, BMTL06, BMTL17, and BMTL10) correspond to the Transmissionist scale, and one of which (BMTL05) corresponds to the Facts First scale.

### **Sample and Setting**

Data were collected in 2019 through the Foundations for Success project, which includes 1197 teachers before cleaning the missing data. Analytic sample used for psychometric analysis included 1196 teachers.

Approximately 1,200 elementary educators in Florida completed the B-MTL in 2019 in two waves of data collection. Approximately 160 of those educators participated in 8 days of a CGI professional development program between January and April 2019. That group of educators was asked to complete the B-MTL (and other assessments, such as K-TEEM) two times during that period in early 2019: one time before they participated in the 8-day CGI program in winter/spring 2019, and one time after they participated in those 8 days of the CGI program. A much larger sample of Florida educators also contributed data by completing B-MTL (and other assessments) in that second wave of data collection that occurred in spring 2019. Randomization of schools and notification of treatment condition occurred in late spring 2019, after that first group had participated in some of the CGI PD program, so the second time those teachers completed the assessments is considered to be the baseline scores for the purpose of the RCT. The data corresponding to that set of approximately 160 teachers who had participated in the CGI PD between January–April 2019 and completed the B-MTL (and other assessments) in spring 2019 are in a separate location and are called the 2019 Second Administration.

Because it is a relatively large sample, and the participants approximately represent the general population of teachers in Florida, the FS B-MTL 2019 First Administration data set was used to establish a baseline for equating scores across those subsequent waves of data collection. The methods for vertical scaling are not described here and will be described separately.

Demographic information for the participants were not available at the time of publication and will be added in the future.

### **Data Collection and Management**

The B-MTL questionnaire was administered through an online survey using Qualtrics software. Response data were exported from Qualtrics and cleaned using syntax in SPSS. Participant identification numbers were replaced with de-identified data before psychometric data analysis occurred. Only the deidentified ID numbers are included in these files.

### **Data Analysis and Findings**

The B-MTL 2019 first administration final sample is composed of 1,196 teachers' responses to 21 items. Dimensionality of the B-MTL 2019 scales were investigated using both parallel analysis (PA) and exploratory factor analysis (EFA). We calculated polychoric correlations of the items to understand the extent to which the items are related. Classical test theory (CTT) and item response theory (IRT) based statistics, and EFA factor loadings were used to examine the item performances. The reliabilities of summed scores (i.e., coefficient alpha and non-linear Structural Equation Model reliability) and response pattern scores (i.e., marginal reliability) are also calculated.

#### **Results of Analyses of Dimensionality**

PA results suggested one dominant component for each B-MTL 2019 first administration scale. The results from one-factor and two-factors solutions (three-factors solution if available) were meticulously examined. The content experts and psychometric team found the one-factor solution more interpretable. Root mean square error of approximations from one-factor EFA solution exceeded the traditional cutoffs of 0.06 for each scale. Comparative fit index values are close or larger than the value of 0.95. Likewise, standardized root mean square residuals are less than the value of 0.08. We note that we evaluated model fit using the thresholds recommended in the oft-cited study by Hu and Bentler (1999), but the situation in Hu and Bentler's study was not a directly applicable match for these data.

#### **Results of Analyses Based on Classical Test Theory**

Corrected item-total correlations (which are also called item-rest correlations) for items in the Transmissionist, Facts First, and Fixed Instructional Plan scales ranged from 0.447 to 0.695, 0.424 to 0.635, and 0.476 to 0.661, respectively. The range of item difficulty for the polytomous items in the three scales was 0.538 to 0.795, 0.365 to 0.698, and 0.608 to 0.833 for the three respective scales.

#### **Results of Analyses Based on Item-Response Theory**

IRT results showed that the item discrimination index of the items in the Transmissionist, Facts First, and Fixed Instructional Plan scales ranged from 1.404 to 3.004, 1.532 to 2.642, and 1.350 to 3.181, respectively.

#### **Reliability**

For the raw sum scores, coefficient alphas of Transmissionist, Fact First, and Fixed Instructional Plan were calculated to be 0.851, 0.757, and 0.772 with SEMs of 2.438, 1.924, and 1.538, respectively. In addition, non-linear Structural Equation Model reliability coefficient estimates were 0.873, 0.790, and 0.795, respectively. Marginal reliabilities for response pattern scores of the Transmissionist, Facts First, and Fixed Instructional Plan scales were 0.890, 0.820, and 0.830, respectively.

### Scoring

Both summed scores and IRT person ability estimates (i.e., theta scores) are provided in the output files containing the person-ability estimates. Summed scores were calculated as a sum of the responses to the items in each scale. Maximum Likelihood Estimation method and Expected A Priori method were used to estimate person locations on the latent continuum.

### Funding Source

Schoen, Robert (PI), Levi, Linda (Co-PI), & Secada, Walter (Co-PI). (Oct 2018–Sept 2021). *Foundations for Success: Developing Effective Mathematics Educators through Cognitively Guided Instruction*. Funded by the United States Department of Education, Office of Innovation and Improvement. (U423A180115). Total award \$9,733,865.

### Bibliographic contributors

[Robert C Schoen,](#)

[Ahmet Guven,](#)

[Gizem Solmaz,](#)

[Amanda M. Tazaz](#)

### Non-bibliographic contributors

[Nancy Thiemann Donahue](#) assisted with data cleaning and associated documentation.

### References

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi-org.proxy.lib.fsu.edu/10.1080/10705519909540118>

Schoen, R.C., & LaVenia, M. (2019) Teacher beliefs about mathematics teaching and learning: Identifying and clarifying three constructs, *Cogent Education*, 6, (1). <https://doi.org/10.1080/2331186X.2019.1599488>

Schoen, R. C., Bray, W. S., Tazaz, A. M., & Buntin, C. K. (2022). *A Description of the Cognitively Guided Instruction Professional Development Program in Florida: 2013–2020*. Florida State University. <https://doi.org/10.33009/fsu.1643828800>