Design, Validity and Reliability

Instrument Background and Design

The current national policy context demands a more nuanced understanding of the association between teaching and student learning. Federal laws such as No Child Left Behind and federal competitive incentive programs such as Race to the Top, the Teacher Incentive Fund and School Improvement Grants drive the need to identify the conditions under which teachers best contribute to student learning (Steele, Hamilton & Stecher, 2010). Additionally, research from private organizations, such as the Bill and Melinda Gates Foundation's Measuring Effective Teachers (MET) project, increase the visibility and support of efforts to explain the relationship between teaching and learning.

This brief furthers the discussion by contributing to a growing body of research that describes how teaching and learning conditions theoretically and empirically link to important outcomes, including teacher retention and student learning. The purpose of this brief is to provide an overview of the research base documenting the association between teaching and learning conditions and outcomes of interest, and to summarize the design and psychometric properties of the New Teacher Center (NTC) Teaching, Empowering, Leading and Learning (TELL) survey instrument. The information provided in this brief serves as the technical basis for additional analyses and reporting, and will be referenced in future briefs.

Research Base

Why do teaching and learning conditions matter? Teaching and learning conditions impact two significant areas of national interest: teacher retention and student learning. The following section summarizes the quantitative relationship between teaching and learning conditions and student learning and teacher retention. It is not intended as an exhaustive review.

Teacher Retention

Large-scale empirical studies present evidence that contextual factors matter for teachers' decisions about staying and leaving schools. In a meta-analysis of 34 studies, researchers suggest that teaching and learning conditions influence teachers' career paths more than previously documented (Borman & Dowling, 2008). Boyd et al. (2011) demonstrate that teachers' perceptions of the school administration have the greatest influence on teacher retention decisions. Other work finds similar effects (Pogodzinski et al., 2012). Studies also find statistically significant relationships between teachers' perception of school facilities and their plans to stay or leave (Loeb, Darling-Hammond & Luczak, 2005; Buckley, Schneider & Shang, 2004).

Researchers using TELL data from various states confirm that teaching and learning conditions influence teachers' plans to stay. Johnson, Kraft and Papay (2011) demonstrate that the conditions that matter most in deciding to stay include the school's culture, the principal's leadership and relationships among peers. Ladd (2009), also using TELL data, documents that teaching and learning conditions predict plans to leave a school, independent of school demographics.

Student Learning

Far fewer large-scale empirical studies explore the association between teaching and learning conditions and student achievement.

The analysis by Ladd (2009) shows that teaching and learning conditions predict student achievement in mathematics, and to a lesser degree, in reading. The Johnson, Kraft and Papay (2011) research indicates that positive conditions contribute to improved student achievement. Both of these efforts use the TELL survey data from various states to estimate the impact of teaching and learning conditions on student learning. Additional work by Kraft and Papay (2014) also uses student-teacher linked data and school-level teaching conditions as measured by the TELL survey to find that teachers who work in more supportive environments become more effective at raising student achievement on standardized tests over time than do teachers who work in less supportive environments, after controlling for student characteristics, prior test scores, and teacher and school characteristics.

In a forthcoming book featuring research from the Bill & Melinda Gates Foundation's MET project, Ferguson with Hirsch (2014) demonstrate significant connections between teaching conditions and student value-added gains. In particular, the authors find that four areas assessed by the NTC's TELL survey—student conduct management, demands on time, professional autonomy and professional development—are significant predictors of student learning gains and student perceptions of rigor and support.

TELL Background and Structure

The TELL survey originates from the Governor's Teacher Working Conditions Initiative in North Carolina (2002–2009). As part of this work, the North Carolina Professional Teaching Standards Commission (NCPTSC) conducted a literature review and analyses of state and national survey data from the National Center for Education Statistics' School and Staffing Survey to better understand the factors contributing to teacher satisfaction and employment trajectories. Based on these efforts, the NCPTSC identified the following areas as related to teachers' future employment plans: time, empowerment, leadership, decision-making, and facilities and resources. The commission created standards aligned with these areas and administered a statewide survey in 2002 to assess whether the standards were in place in schools.

The TELL survey incorporates these constructs and includes others logically and empirically linked to outcomes of interest, such as teacher retention and student learning. These constructs include student behavior support, community support, and instructional practices and support. Based on the NCPTSC-identified areas and an external validation study, the TELL survey currently includes the eight constructs identified in Exhibit 1.

EXHIBIT 1. TELL SURVEY CONSTRUCTS

Time—Available time to plan, collaborate, provide instruction, and eliminate barriers to maximize instructional time during the school day

Facilities and Resources—Availability of instructional, technology, office, communication, and school resources to teachers

Community Support and Involvement—Community and parent/guardian communication and influence in the school

Managing Student Conduct—Policies and practices to address student conduct issues and ensure a safe school environment

Teacher Leadership—Teacher involvement in decisions that impact classroom and school practices

School Leadership—The Ability of school leadership to create trusting, supportive environments and address teacher concerns

Professional Development—Availability and quality of learning opportunities for educators to enhance their teaching

Instructional Practices and Support—Data and support available to teachers to improve instruction and student learning

919.806.2200

On current TELL survey administrations, the NTC adds questions to these eight core constructs about general demographic information, beginning teacher support and client-specific information. Survey responses related to the eight constructs are scored using Likert-type ratings ranging from strongly disagree (1) to strongly agree (4), with a "Don't Know" option.

External Analyses of Validity and Reliability

This section describes the methods used by an external analyst to verify that the structure and items included in the TELL survey result in meaningful and useful information. This work is part of the MET project supported through the Bill and Melinda Gates Foundation (Swanlund, 2011). The Swanlund analyses use data from 286,835 educators from 11 states across the U.S. The external survey review examines both validity and reliability. These analyses identify patterns in the data that provide a clear structure for the survey and confidence for interpreting the results.

Validity

Validity generally refers to the process of ensuring that a survey accurately measures what it is intended to measure, in this case teaching and learning conditions. There are several approaches to testing validity. The external validity testing conducted for the TELL survey assesses the structure of the response scale and the alignment between survey items and the broader survey constructs identified in Exhibit 1. The review uses the Rasch rating scale to examine the item-measure correlations, item fit, rating scale functioning, unidimensionality and generalizability of the instrument.

Results from the external validity testing prompted several edits to increase the statistical stability of the TELL survey. For example, a four-point rating scale replaced the original six-point scale to ensure appropriate scoring for both individual-level responses and school-level responses. Based on the external study finding that some survey constructs are more stable if broken into multiple constructs, an additional construct was added, resulting in eight constructs. Additionally, the results indicate that some individual items

overlap across survey constructs. For example, items found in the teacher leadership construct overlap with the school leadership construct and should be reviewed for each analysis.

Reliability

Reliability testing ensures the survey instrument produces the same results across repeated measures, either within the same population or with a similar population. A reliable survey is generalizable and is therefore expected to reproduce similar results across settings. The external review analyzes reliability using both the Rasch model and Cronbach's alpha. The Swanlund (2011) study concludes the survey is capable of producing consistent results across participant groups.

In summary, the external analyses confirm that the TELL survey offers a robust and statistically sound approach for measuring teaching and learning conditions. For a detailed review of the methods and results from the external analyses, consult Swanlund (2011).

Internal Analyses of Validity and Reliability

In addition to the external analyses, the NTC conducts internal analyses of validity and reliability to verify the stability of the instrument across survey populations as promoted by industry standards found in Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association and National Council on Measurement in Education, 1999). Statistical tests of validity include factor analyses and reliability tests that generate internal consistency estimates.

In early 2014, the NTC administered the anonymous survey to all the reported 80,200 school-based licensed educators in Massachusetts. Over 38,000 (48 percent) in the state responded (Exhibit 2). Of the 38,217 responding, 33,854 are licensed educators (89 percent), 2 percent are administrators and 9 percent are other licensed professionals, such as librarians and school psychologists. Exhibit 2 provides response rates by participant type. These analyses are based on responses provided by the school-based licensed educators.

EXHIBIT 2. RESPONSE RATE BY PARTICIPANT TYPE	
Respondents*	**Response Rate (N)** **Spring 2014**
Teachers	88.5% (33,854)
Administrators	2.4% (922)
Other Education Professionals	9.0% (3,441)

***Note.** The respondent category "teachers" includes instructional coaches, department heads, literacy specialists, etc. The respondent category "administrators" includes principals and assistant principals. The respondent category "Other Education Professionals" includes school counselors, school psychologists, social workers, etc.
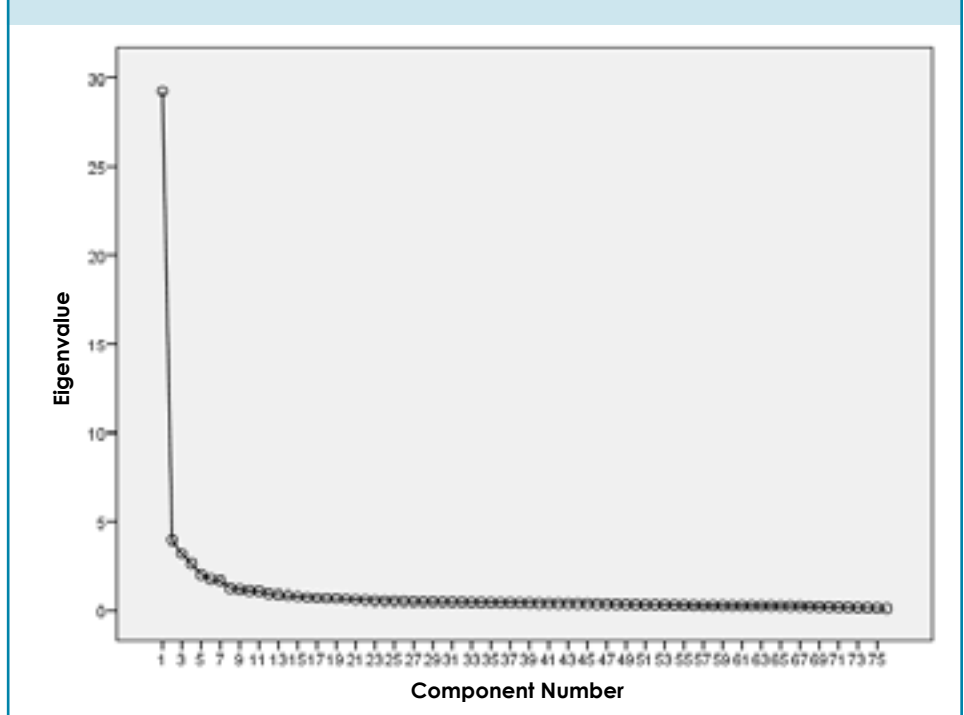
Validity

The validity analyses assess the degree to which the 2014 TELL Massachusetts Survey measures the eight theoretical constructs it is intended to capture. See Exhibit 1 for descriptions of the constructs. The NTC conducts factor analyses to group variables with similar characteristics. The NTC also performs confirmatory factor analysis (CFA), using principal components analysis and varimax rotation procedures, to verify that the actual structure of the data reflects the structure expected from the external validation study.

Researchers suggest several empirical criteria for determining which orthogonal or correlated factors to retain in a stable instrument. These criteria are based on an eigenvalue. Eigenvalues indicate how much variation each factor or component can explain. The criteria include: scree plot, Kaiser criterion and variance explained (Cortina, 2002). Additionally, Hair et al. (2006) suggest including construct correlations. However, the final decision about which factors should be retained should be based on judgments of interpretability and consistency of the factors with sound theory (Bandalos & Boehm-Kaufman, 2009). Next, the NTC provides information about each of the recommended empirical criteria.

The scree plot graphically represents the eigenvalues in descending order and connects them with a line. Researchers suggest examining the line for where it levels off. Exhibit 3 indicates an "elbow" beginning with factor two, continuing through factors seven or nine, and then smoothing or showing that each additional factor beyond that accounts for smaller amounts of the total variance (Ledesma & Vlero-Mora, 2007). Therefore, the scree plot shown in Exhibit 3 would suggest approximately a seven- to nine-factor solution.

EXHIBIT 3. SCREE PLOT

The Kaiser criterion (K1) suggests only including factors where eigenvalues are greater than one (as a theoretical lower bound). The initial eigenvalues displayed in Exhibit 4 show that at least eight factors have a value of more than one and therefore meet minimal variance-explained thresholds.

Another recommended approach for deciding which factors to retain suggests examining the variance explained and retaining factors contributing 10 percent or more. Exhibit 4 shows that the eight factors each contribute at least 10 percent of the variance, and together explain 60 percent of the variance.

EXHIBIT 4. EIGENVALUES AND VARIANCE EXPLAINED			
	Initial Eigenvalues		
Component	**Total**	**Percent of Variance**	**Cumulative Percent**
1	29.222	38.450	38.450
2	3.966	5.218	43.668
3	3.184	4.189	47.857
4	2.630	3.461	51.318
5	2.015	2.651	53.968
6	1.751	2.304	56.272
7	1.657	2.181	58.452
8	1.259	1.656	60.109

THE INITIAL EIGENVALUES *displayed in Exhibit 4 show that at least eight factors have a value of more than one and therefore meet minimal variance-explained thresholds. The eight factors each contribute at least 10 percent of the variance, and together explain 60 percent of the variance.*

The empirical criteria reviewed together indicate at least an eight-factor solution, including a general leadership category and a general community support and involvement category with a subcategory. Due to the findings of the external validation study and the theoretical framework the TELL survey is based on, the NTC conducted further analyses.

The construct correlations are presented to examine if factors are correlated above the professional standard of 0.70 (Hair et al., 2006). Factor correlations above 0.70 indicate that the constructs overlap and do not capture distinct areas of teaching and learning conditions. Exhibit 5 suggests that the factors for these analyses are distinct and have limited overlap.

EXHIBIT 5. COMPONENT CORRELATION MATRIX								
Component	1	2	3	4	5	6	7	8
1	1.000	.653	.550	.515	.554	.642	.537	.308
2	.653	1.000	.494	.410	.550	.441	.508	.317
3	.550	.494	1.000	.502	.548	.513	.382	.346
4	.515	.410	.502	1.000	.387	.541	.365	.318
5	.554	.550	.548	.387	1.000	.428	.305	.344
6	.642	.441	.513	.541	.428	1.000	.426	.288
7	.537	.508	.382	.365	.305	.426	1.000	.010
8	.308	.317	.346	.318	.344	.288	.010	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalization.

FACTOR CORRELATIONS *above 0.70 indicate that the constructs overlap and do not capture distinct areas of teaching and learning conditions. Exhibit 5 suggests that the factors for these analyses are distinct and have limited overlap.*

The NTC conducted an additional factor analysis that defined the variables associated with each of the eight pre-identified constructs listed in Exhibit 1. The factor loadings across these constructs indicate that the school leadership and teacher leadership constructs include items with high factor loadings (all above 0.694), do not cross load and could stand as independent constructs. However, the community support and involvement construct included items that cross loaded and will be explored further in outcome analyses.

Empirically and theoretically, the factor analysis for the TELL Massachusetts Survey supports eight factors that are similar to the external validity work.

Confirmatory factor analyses of the data set suggest that the TELL Massachusetts Survey eight-construct structure provides stable and generalizable measures of teaching and learning conditions, and is consistent with the theoretical framework supporting this work.

Reliability

The internal reliability testing for the TELL Massachusetts Survey confirms that the survey is generalizable and will produce similar results with similar populations. The reliability analyses for the survey produce Cronbach's alpha coefficients ranging from 0.80 to 0.94. Alphas normally range between 0.00 and 1.00. The closer the Cronbach's alpha coefficient is to 1.00, the greater the internal consistency of the items in the scale. Alpha coefficients above 0.70 are considered acceptable (George & Mallery, 2003).

As Exhibit 6 indicates, all eight alpha coefficients are above 0.70, confirming internal consistency of the TELL Massachusetts Survey constructs.

EXHIBIT 6. RELIABILITY BY CONSTRUCT	
Construct	**Cronbach's Alpha**
Time	.815
Facilities and Resources	.882
Community Support and Involvement	.885
Managing Student Conduct	.897
Teacher Leadership	.932
School Leadership	.944
Professional Development	.946
Instructional Practices and Support	.795

As Exhibit 6 indicates, *all eight alpha coefficients are above 0.70, confirming internal consistency of the TELL Massachusetts Survey constructs.*

Summary of TELL Massachusetts Validity and Reliability

Based on external and internal analyses of TELL Massachusetts Survey data, the most appropriate structure of the survey includes eight factors. These eight constructs will be the basis for other reports presenting general trends across the data.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.

Bandalos, D. & Boehm-Kaufman, M. (2009). Four common misconceptions in exploratory factor analysis. In C. Lance & R. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*. (pp. 61–87). UK:Taylor & Francis.

Borman, G. & Dowling, N. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, 78(3).

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The influence of school administrators on teacher retention decisions. *American Educational Research Journal*, 48(2).

Buckley, J., Schneider, M., & Shang, Y. (2004). *The effects of school facility quality on teacher retention in urban school districts*. Chestnut Hill, MA: National Clearinghouse for Educational Facilities.

Cortina, J. (2002). Big things have small beginnings: An assortment of "minor" methodological misunderstandings. *Journal of Management*, 28(3).

Ferguson, R. & Hirsch, E. (2013). *Using teacher and student surveys to link school context, classroom learning conditions and achievement*. Forthcoming.

George, D. & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston, MA: Allyn & Bacon.

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). Multivariate data analysis (6th ed.). Upper Saddle River, N.J.: Pearson Prentice Hall.

Johnson, S., Kraft, M., & Papay, J. (2012). How context matters in high-need schools: The effects of teachers' working conditions on their professional satisfaction and their students' achievement. *Teachers College Record*, 114(10).

Kraft, M.A. & Papay, J.P. (2014, Jan. 30). Do supportive professional environments promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*.

Ladd, H. (2009). Teachers' perceptions of their working conditions: How predictive of policy relevant outcomes? CALDER Working Paper 33. Washington, DC: National Center for Analysis of Longitudinal Data in Education.

Ledesma, R., & Valero-Mora, P. (2007). Exploratory factor analysis. *Practical Assessment, Research & Evaluation*, 12(22).

Loeb, S., Darling-Hammond, L., & Luczak, J. (2005). How teaching conditions predict teacher turnover in California schools. *Peabody Journal of Education*, 80(3).

Pogodzinski, B., Youngs, P., Frank, K., & Belman, D. (2012). Administrative climate and novices' intent to remain teaching. *The Elementary School Journal*, 113(2).

Steele, J., Hamilton, L., & Stecher, B. (2010). *Incorporating student performance measures into teacher evaluation systems*. Santa Monica, CA: RAND Corporation.

Swanlund, A. (2011). *Identifying working conditions that enhance teacher effectiveness: The psychometric evaluation of the Teacher Working Conditions Survey*. Chicago. IL: American Institutes for Research.

UCLA: Statistical Consulting Group. Factor Analysis: from http://www.ats.ucla.edu/stat/spss/output/factor1.htm (accessed November 20, 2012).

About the New Teacher Center

New Teacher Center focuses on improving student learning by accelerating the effectiveness of new teachers. NTC partners with states, school districts, and policymakers to design and implement systems that create sustainable, high-quality mentoring and professional development; build leadership capacity; work to enhance teaching conditions; improve retention; and transform schools in vibrant learning communities where all students succeed.

725 Front Street, Suite 400, Santa Cruz, CA 95060
831-600-2200 I Fax: 831-427-9017 I info@newteachercenter.org
www.newteachercenter.org