University of Kentucky

A Methodological Breakthrough in Instrument Validation
(Instrument Validation = 测量工具验证)

Xin Ma

Reference

Ma, X., Shen, J., Reeves, P., & Yuan, J. (2020). A multilevel examination of an instrument measuring school renewal via teachers. *Studies in Educational Evaluation, 65,* XXX-XXX.

(*Studies in Educational Evaluation* is a SSCI Journal)

A Critical Measurement Issue

For a given construct, what is the best source of data to measure it?

For educational data:

- Students
- Teachers
- Parents
- Principals
- Superintendents

Measuring School Renewal (学校更新)

Get data to measure this school-level leadership construct from:

- Students: No
- Parents: No
- Superintendents: No
- Principals: Maybe
- Teachers: Yes

A Serious Analytical Problem

The Problem of Unit-of-Analysis: When use teacher data to measure a school-level construct, what is the unit of analysis?

Sirotnik (1980) distinguished between two phases in data analysis: The psychometric phase and the study phase.

"The psychometric implications of unit-of-analysis issues have been almost universally ignored in the organizational climate literature" (Sirotnik, 1980, p. 158).

Our Position

Sirotnik (1980) position:

- Consider multilevel analysis in the study phase

Ma et al. (2020) position:

- Consider multilevel analysis both in the study phase and in the psychometric phase

Our Plan

Goal:

- Take into consideration the data hierarchy with teachers nested within schools (principals)

Strategy:

- Multilevel confirmatory factor analysis
  - Exploratory factor analysis (EFA) (探索性): Data driven, exploratory
  - Confirmation factor analysis (CFA) (确认性): Theory driven, confirmatory

Our Instrument

Table 1

Intra-Class Correlation (ICC) and Descriptive Statistics of Items Measuring School Renewal via Teachers.

		ICC	M	SD
F1	Focus on students and their achievement			
01	Our school improvement process is guided strongly by the goal of improving student achievement.	.979	5.160	1.004
02	Our school truly has high expectations for all students.	.977	4.880	1.081
03	All teachers have a clear, shared vision about expectations for all students.	.974	4.430	1.149
F2	Continuous school improvement			
04	Our school has a continuous focus on teaching and learning.	.980	5.120	.974
05	All our teachers continuously seek ways to enhance the teaching and learning processes.	.976	4.850	1.011
06	Our school consistently uses a continuous improvement process/strategy, rather than starting from scratch for each initiative.	.968	4.270	1.282
F3	Balance between the internal and external influences			
07	We openly welcome ideas and input on school improvement from all stakeholders.	.967	4.500	1.259
08	We successfully balance external pressure and internal initiative for school improvement.	.974	4.160	1.121
09	We successfully prioritize our school improvement efforts despite competing priorities.	.974	4.280	1.131
F4	The dialogue, decision, action and evaluation (DDAE) process			
10	We consistently dialogue in our school about our school improvement priorities.	.969	4.270	1.283
11	Our school improvement strategies are well coordinated within the school.	.972	4.110	1.254
12	Our school successfully monitors the progress of our school improvement initiatives with data.	.761	4.610	1.207
F5	Implementation integrity			
13	We consistently monitor our data and develop school improvement initiatives accordingly.	.974	4.520	1.185
14	We have a clear process in place to continuously generate new ideas for school improvement.	.970	3.920	1.256
15	We consistently re-prioritize school improvement efforts based on continuous data updates.	.972	4.160	1.242
F6	Implementers as active developers			
16	Our school really decides our school improvement priorities.	.966	4.090	1.307
17	We usually develop our own programs for school improvement (rather than buying from an external vendor).	.694	4.020	1.322
18	We consistently adapt and adjust existing programs based on our outcome data.	.298	4.210	1.203
F7	Internal responsibility and professionalism			
19	We all hold ourselves and each other accountable.	.683	4.500	1.221
20	We all hold our students accountable for their own achievement.	.589	4.430	1.191
21	Continuous reflection on school improvement is part of our professional culture.	.106	4.350	1.265

Note. All items are measured on a Likert scale of 1–6 (strongly disagree = 1, moderately disagree = 2, slightly disagree = 3, slightly agree = 4, moderately agree = 5, strongly agree = 6).

ICC = intra-class correlation

ICC measures the proportion of variance attributable to schools in our case.
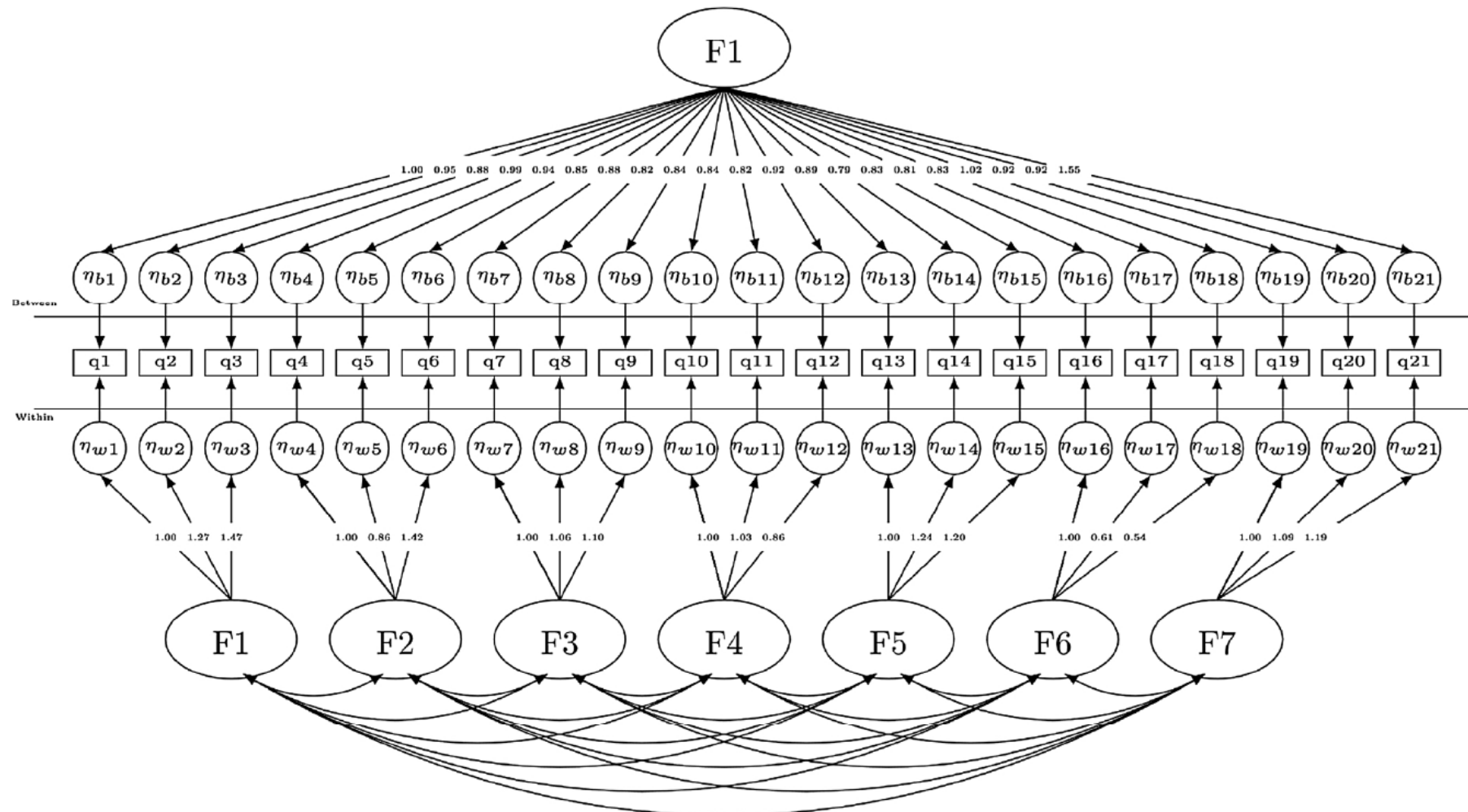
Our Model

Fig. 1. Multilevel Factor Structure of An Instrument Measuring School Renewal via Teachers.

Our Results: Structural Validity

- Model-data-fit statistics:
    - RMSEA=.038,
    - TLI=.954
    - CFI=.957
- Standard:
    - RMSEA≤.08
    - TLI≥.95
    - CFI≥.95
- Conclusion: Our data support our multilevel CFA model (i.e., structural validity)

Our Results: Multilevel Reliability

The results on multilevel reliability:

- ω=.319 among teachers
- ω=.998 among schools

Different reliability measures:

- MacDonald's ω
- Cronbach's α
- As a reliability measure (for internal consistency), ω is similar to α.
- Major advantage of ω over α: ω takes into consideration the strength of the relationship between items and factors as well as specific item-level measurement errors.

Our Conclusion (Caution)

- When teachers provide responses as indicators of school renewal, the instrument can generate a valid and highly reliable measure or estimate of school renewal at the school level.

- However, data analysis at the teacher level attempting to use teacher perceptions of school renewal as either dependent variables or independent variables should be avoided because of the low reliability at the teacher level.

Implications for Research in Math Education

- The use of instruments or scales that collect data from students to measure teacher attributes is questionable without validation evidence from multilevel CFA
  - Example: Use student responses to measure classroom practice of math teachers
- The use of instruments or scales that collect data from teachers to measure school (principal) attributes is questionable without validation evidence multilevel CFA
  - Example: Use teacher responses to measure principal support for math education