

Your full name: Ridwan Maulana

Affiliated authors with institutions: Michelle Helms-Lorenz, Xiangyuan Feng, James Ko, Seyeoung Chun, Abid Shahzad, Yulia Irnidayanti, Okhwa Lee, Thelma de Jager, Thys Coetzee, Nurul Fadhillah, Carmen – Maria Fernandez – Garcia, Mercedes Inda – Caro, Alison Kington, Amarjargal Adiyasuren, Ulziisaikhan Galindev

Affiliation: Department of Teacher Education, University of Groningen, the Netherlands Department of Education Policy and Leadership, The Education University of Hong Kong Department of Education, Chungnam National University, Daejeon, South Korea Department of Education, The Islamia University of Bahawalpur, Bahawalpur, Pakistan Department of Biology and Biology Education, State University of Jakarta, Indonesia Department of Education, Chungbuk National University, Cheongju, South Korea Department Educational Foundation, Tshwane University of Technology South Africa Department of Biostatistic and Population, University of Indonesia, Indonesia, University of Oviedo, Spain, University of Worcester, The United Kingdom, Mongolian National University of Education, Mongolia

Current position: Associate Professor

Title of your paper: Testing the validity of the ICALT observation instrument for measuring effective teaching behaviour in 10 countries

Abstract (300 words)

The present study aims to test the validity of the ICALT observation instrument, based on the effective teaching behaviour model, to measure secondary school teachers' effective teaching behaviour in 10 contrasting countries including the Netherlands, South Korea, Indonesia, South Africa, Hong Kong - China, Pakistan, Mongolia, Spain, England, and the USA. In total, 3411 observed teachers were included ($N_{\text{Netherlands}} = 605$, $N_{\text{Indonesia}} = 335$, $N_{\text{Pakistan}} = 400$, $N_{\text{SouthAfrica}} = 304$, $N_{\text{SouthKorea}} = 207$, $N_{\text{HongKong-China}} = 284$, $N_{\text{Spain}} = 344$, $N_{\text{USA}} = 320$, $N_{\text{Mongolia}} = 403$, $N_{\text{England}} = 209$). Categorical Confirmatory Factor Analyses based on the Structural Equation Modelling (SEM) framework was applied. Results show that the full measurement model of the six-factor structure for the ICALT instrument with complete items is supported in six countries including The Netherlands, Indonesia, South Africa, South Korea, Pakistan, and England. In Mongolia, Spain, and Hong Kong – China, removing a limited number of items was necessary to obtain adequate model fit. The hypothesized factor structure was not confirmed in the USA even after large model modifications. Implications of the findings for theory and practice will be discussed, including the promises and challenges of using the ICALT instrument in various cultural contexts and how to potentially improve the measurement of effective teaching behaviour in future research.

Extended summary (1000 words, excluding reference list)

Introduction

The International Comparative Analysis of Learning and Teaching (ICALT) instrument has been validated to measure teaching quality across primary schools in various European countries (van de Grift, 2007; 2014). Since then, the observation instrument has gained popular attention internationally. However, little was known regarding the validity of the measure for use in secondary

education internationally. In 2015, an ambitious research project to validate the ICALT instrument in secondary schools across 10 diverse countries started¹. Currently, the project team has successfully collected observation data of teaching behaviour, which allows us to test the validity of the ICALT instrument for measuring effective teaching behaviour in the Netherlands, South Korea, Indonesia, South Africa, Hong Kong - China, Pakistan, Mongolia, Spain, England and the USA.

Theoretical background

Research on teaching and teacher effectiveness has consistently shown the significant role of teaching behavior for pupil learning and outcomes (Creemers & Kyriakides, 2008; Hattie, 2012). Based on synthesis of evidence-based effective teachers' behaviour research, van de Grift (2007) identified six observable domains of teaching behaviour. The six domains domains are: 1) *Safe and Stimulating Learning Climates*, 2) *Efficient Classroom Management*, 3) *Clarity of instruction*, 4) *Activating Teaching*, 5) *Differentiated Instruction*, and 6) *Teaching Learning Strategies*. The hypothesized model of effective teaching behaviour for secondary education was tested in the Netherlands using a pre-service teachers sample (Maulana et al., 2017, Figure 1). The study found that the six domains were distinguishable and strongly related to student engagement. The six domains of effective teaching behaviour are also in accordance with other models and empirical findings of effective teaching (Danielson, 2013; Pianta & Hamre, 2009; Ko et al., 2013).

Aims

The aim of the present study is to test the validity of the ICALT observation instrument to measure secondary school teachers' effective teaching behaviour in the Netherlands, South Korea, Indonesia, South Africa, Hong Kong - China, Pakistan, Mongolia, Spain, England and the USA.

Methods

In total, 3411 observed teachers participated in this study ($N_{\text{Netherlands}} = 605$, $N_{\text{Indonesia}} = 335$, $N_{\text{Pakistan}} = 400$, $N_{\text{SouthAfrica}} = 304$, $N_{\text{SouthKorea}} = 207$, $N_{\text{HongKong-China}} = 284$, $N_{\text{Spain}} = 344$, $N_{\text{USA}} = 320$, $N_{\text{Mongolia}} = 403$, $N_{\text{England}} = 209$). The ICALT instrument was used for measuring the six domains of teaching behaviour across the 10 countries, which was constructed based on the effective teaching behaviour model (Maulana et al., 2017; van de Grift, 2007). The instrument consists of 32 high-inference indicators (4 response categories ranging from weak to strong), followed by 114 corresponding low-inferences indicators (2 response categories, yes/no). The instrument was adapted to each specific country's context using the translation-back translation procedure (Hambleton, 1994). Classroom observations were conducted by trained observers. Inter-observer reliability in the 10 countries was adequate (> 70%).

To test the validity of the ICALT instrument in each country, Categorical Confirmatory Factor Analyses based on Structural Equation Modelling (SEM) framework was performed separately for each country data using Mplus 8.1 (Muthen & Muthen, 2017). The six correlated factor solution was compared with the competing one-factor solution where all items loaded on only one factor. The common model-data goodness of fit indices including the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI) were used. The criteria

¹ <https://www.rug.nl/gmw/lerarenopleiding/onderzoek/psychometrisch/>

include: CFI and RMSEA and lower and upper for 90 % confidence interval of RMSEA < 0.08; CFI > 0.90; TLI > 0.90 for an acceptable model fit (Brown, 2014).

Results

In general, the six-factor solution model resulted in better model fit compared to the one-factor solution in all countries. The full measurement model of the six-factor C-CFA for the ICALT instrument with complete items is supported in six countries including The Netherlands, Indonesia, South Africa, South Korea, Pakistan, and England see Table 1). In Mongolia, Spain, and Hong Kong – China, the measurement model is supported after excluding one item, four items, and eight items respectively. In the USA, the measurement model is not supported even after significant model modifications.

Theoretical and education significance

The present study contributes to theory and measurement of teaching quality by linking the effective teaching behaviour model to testing the validity of the instrument measuring effective teaching behaviour in secondary education across diverse national contexts. The empirical testing of the measure across various cultural contexts provides insights into the relevance of the effective teaching behaviour model in other cultural contexts. For educational practice, results provide evidence regarding the promise and challenges of using the ICALT measure as diagnostic tool for formative assessments in schools across countries. Findings that certain number of items do not function well in certain countries suggest that there are specific (unique) aspects of teaching behaviour operating at the specific context, which cannot be tapped generally and requires an integration of etic and emic approaches in studying teaching behaviour across countries.

Relevance to the QUINT ambition

Our study fits the QUINT ambition “theorizing teaching quality and testing out a framework of effective teaching behaviour” very well.

References

- Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research. Methodology in the Social Sciences*. London: Guilford.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge
- Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. Princeton: The Danielson Group.
- Hambleton, R. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*, 229-244.
- Hattie, J. A. C. (2012). *Visible learning for teachers: maximizing impact on learning*. Abingdon and New York: Routledge.

Ko, J., Sammons, P., & Bakkum, L. (2013). *Effective teaching: A review of research and evidence*. CIBT Education Trust.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide*. (Eighth Edition ed.). Los Angeles, CA: Muthén & Muthén.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109-119.

Van de Grift, W. J. C. M. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educational Research*, 49(2), 127-152. doi:10.1080/00131880701369651.

Van de Grift, W. J. C. M. (2014) Measuring teaching quality in several European countries. *School Effectiveness and School Improvement*, 25, 295-311, DOI: 10.1080/09243453.2013.794845

Table 1. Categorical Confirmatory Factor Analysis for ten countries

Country	Model	Valid N	χ^2 (df)	RMSEA with 90% CI	SRMR	CFI	TLI
1. The Netherlands	All items	606	1761.164*(449)	0.069 [0.06, 0.073]	0.065	0.917	0.908
2. Indonesia	All items	335	1346.169*(449)	0.077 [0.073, 0.082]	0.071	0.938	0.931
3. South Africa	All items	304	1372.800*(449)	0.082 [0.077, 0.087]	0.053	0.966	0.962
4. South Korea	All items	208	815.571*(449)	0.063 [0.056, 0.069]	0.067	0.962	0.958
5. Pakistan	All items	400	1362.984*(449)	0.071 [0.067, 0.076]	0.075	0.926	0.919
6. England	All items	209	748.371*(449)	0.056 [0.049, 0.064]	0.086	0.969	0.965
7. Mongolia	All items	403	1660.461*(449)	0.082 [0.078, 0.086]	0.078	0.893	0.881
	Item 23 deleted	403	1364.907*(419)	0.075 [0.070, 0.079]	0.070	0.914	0.904
8. Spain	All items	344	1606.119*(449)	0.087 [0.082, 0.091]	0.086	0.865	0.851
	Items 13, 17, 23, 31 removed	344	977.737*(335)	0.075 [0.069, 0.080]	0.072	0.914	0.903
9. HK-China	All items	284	2449.374*(449)	0.125 [0.120, 0.130]	0.137	0.752	0.726
	Items 3, 4, 10, 13, 16, 18, 19, 20 removed	284	587.683*(237)	0.072 [0.065, 0.080]	0.080	0.944	0.934
10. USA	All items	320	3580.603*(449)	0.148 [0.143, 0.152]	0.189	0.580	0.537
	Items 1, 14, 16, 18, 19, 20, 24, 25 removed	320	930.073*(237)	0.096 [0.089, 0.102]	0.120	0.827	0.799
	Items 1, 14, 16, 18, 19, 20, 24, 25, 32 removed	320	793.046*(215)	0.092 [0.085, 0.099]	0.110	0.847	0.820

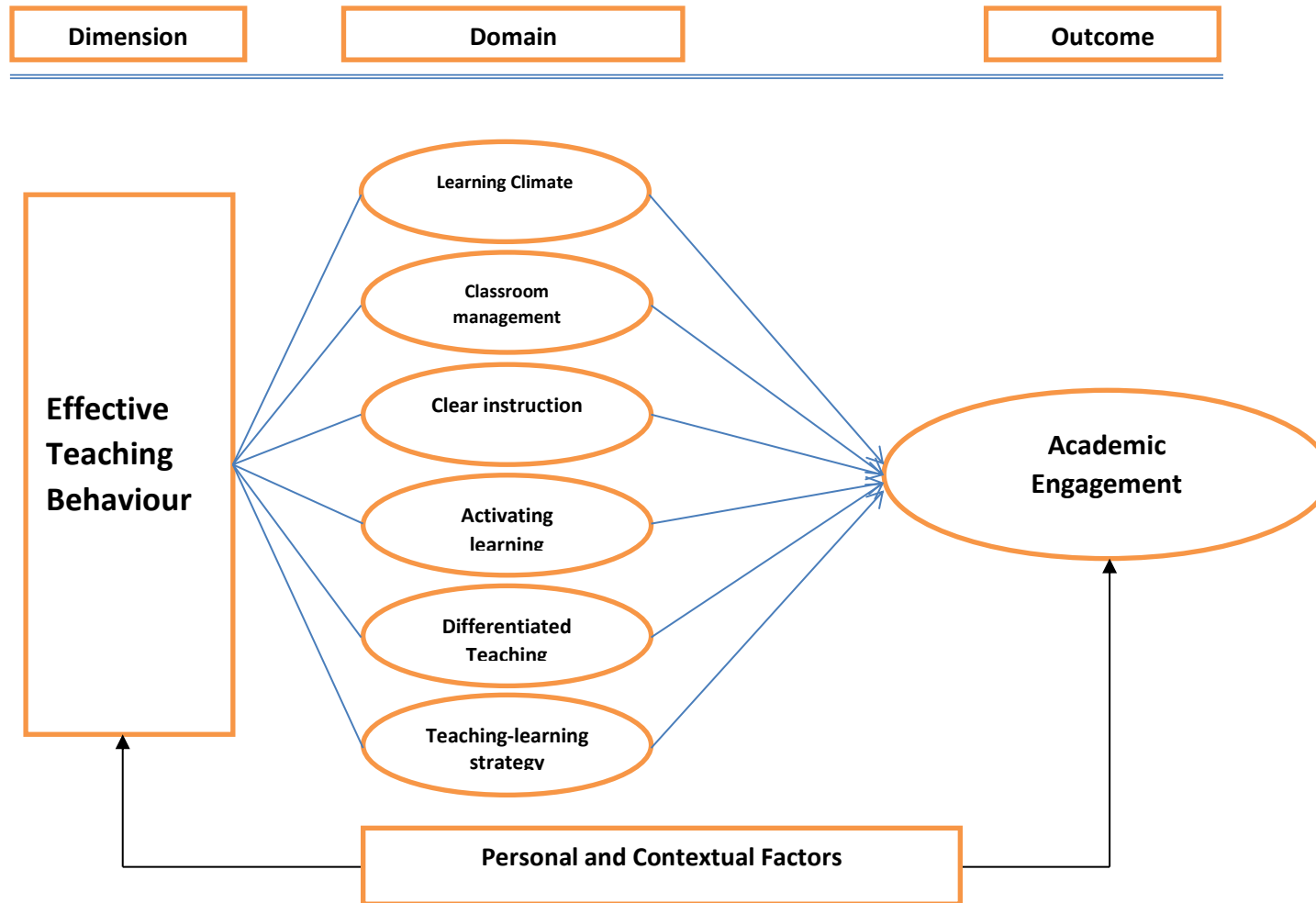


Figure 1. The hypothesised model effective teaching behaviour