Author: Jennifer Maria Luoto
TITLE: Disentangling tensions between the comparative education field and classroom observation systems

Your full name: Jennifer Maria Luoto

Affiliated authors with institutions: -

Affiliation: Department of Teacher Education and School Research, University of Oslo

Current position: Postdoctor

Title of your paper: Disentangling tensions between the comparative education field and classroom observation systems used comparatively

Abstract

Classroom observation research often use standardized observation system for the purpose of capturing and comparing the quality of teaching across different national and international contexts. This study investigates how three inherent concerns within the comparative education field are addressed in three observation systems designed for comparative use, and in comparative studies that apply these systems. The concerns are *conceptualization of teaching quality*, *attention to context*, and *implications of results,* that encompass the theoretical framework of this study. Key documents describing the identified observation systems as well as studies applying them comparatively (*N*=15) were analyzed using qualitative content analysis. Preliminary findings indicate a similar *conceptualization of teaching quality* at domain-level across the three systems, yet differences in operationalizing the domains into observable practices and divergent assumptions about the relationship between teaching and learning outcomes. For *attention to context*, the identified studies include different aspects of context (classroom, academic, structural, cultural) yet in only one study the context is included in the discussion of patterns of teaching quality, and then only the cultural context. Regarding *implication of results*, these studies together are delivering both policy, practice and research implications. They all to a different degree indicate that we can learn from the contexts that have the highest scores as defined in the observation systems (in some studies, this is specified as in practices correlated with student outcomes). However, there is no discussion of *how* we can learn from high achieving teachers and *why* we see differences in patterns of teaching quality, which leaves much of the potential of comparative research to improve or reflect on teaching untapped. This study concludes with discussing how studies using observation systems comparatively can benefit from perspectives by the comparative education field.

Extended summary

Introduction

Comparative classroom observation studies often use *standardized observation systems* to conceptualize and operationalize teaching quality into measurable practices that can be systematically compared (Praetorius et al., 2019). In contrast, the *comparative education field* is mostly concerned with *understanding* how societal and cultural aspects shape teaching (Schriewer, 2021; Schweisfurth et al, 2020). At the same time, comparative scholars argue that teaching is insufficiently studied comparatively (Alexander, 2009; Suter, 2019), recognizing that to understand *how* context shapes teaching and *why*, we need to know *what*

teaching looks like. This is where observation systems potentially serve a purpose for the comparative education field.

This study aims to contribute to bridging the gap by examining how concerns intrinsic to comparative education are addressed in three observations systems and studies that apply them comparatively. The research questions are:

1. *How is teaching quality conceptualized in international observation systems and how is this conceptualization legitimized for comparative research?*
2. *What aspects of contexts are considered and how and what kind of implications of results are suggested in in cross-national comparative classroom studies using international observation systems?*

Theoretical framework

Three recurrent concerns in comparative education literature on classroom research frame the analysis of this study; *conceptualizations of teaching quality*, *attention to context*, and *implications of results*.

Teaching quality is often defined as teaching and learning opportunities related to some student outcome. This conceptualization tend to build on different theoretical traditions and empirical research mainly from Anglo-Saxon and Central European contexts (Praetorius & Charalambous, 2018). *Conceptualizations of teaching quality* thus refer to the origins of the conceptualization, how it is justified for comparative application, and what assumptions it holds about the relationship between teaching and learning.

Comparative education scholars unanimously argue that comparative classroom research should be situated within a context of different levels, and that meaningful comparisons of teaching explore the relationship between these interacting levels (Alexander, 2009; Schweisfurth, 2019). *Attention to context* therefore include concerns about what different levels of context are reported (i.e., classroom, academic, structural and cultural context) and how context is used to understand findings.

Scholars within this field also warn against how results of comparative educational studies may be used for political purposes as well as for uncritically promoting universal views of concepts like teaching quality (Phillips & Schweisfurth, 2014; Reynolds, 2006). *Implications of results* thus encompass suggested implications of results for policy, pedagogical practice, and research.

Methods

Qualitative content analysis (Elo & Kyngäs, 2008) was applied to study three international observation systems and their application in comparative studies; International Comparative Analyses of Learning and Teaching (ICALT; van de Grift, 2007), International System for

Teacher Observation (ISTOF; Teddlie et al, 2006), and Global Teaching Insights Observation System (GTI; OECD, 2019b). The preliminary sample is 15 documents including published descriptions of the systems as well as comparative classroom observation studies, which sometimes overlap (Table 1). The descriptions of the systems constitute one unit of analysis, related to RQ1, while the analyzed studies relate to RQ2.

Table 1. Sample

Obs. system	**Texts**	Key text	Comparative study
ICALT	***T1.*** van de Grift (2007)	*x*	*x*
	T2. van de Grift (2014)	*x*	*x*
	T3. van de Grift et al., (2017)	*x*	*x*
	T4. Maulana et al (2019)		*x*
	T5. Maulana et al. (2020)		*x*
ISTOF	***T6.*** Teddlie et al., (2006); Kyriakides et al., (2010); Muijs et al 2018	*x*	
	T8. Miao et al., (2015)		*x*
GTI	***T9.*** OECD (2020)		*x*
	T10. OECD, 2019; Opfer, 2020; Bell et al., 2020; Bell, 2020; Castellano & Bell, 2021	*x*	

The analytical framework encompasses the three concerns by the *academic comparative education field* (Cowen, 2014), operationalized as questions (Table 2).

Table 2. Analytical framework

Conceptualization
How is the observation systems conceptualizing teaching quality? What assumptions about teaching and learning?
Whose conceptualization of quality teaching?
What are the limitations of a certain conception of teaching quality?
Context
How is the studies attending to context and using context?
What is the justification for the comparison?
Implication
What implications are drawn?

Findings

The observation systems have similar overarching conceptualizations of teaching quality, while they differ in assumptions about the relationship between teaching and its outcomes. ICALT and ISTOF focus on teaching practices empirically related to student achievement gains, building on teaching effectiveness research and consensus among different countries. GTI's conceptualization emphasizes both opportunities teachers provide and how students use opportunities, and recognizing cognitive as well as non-cognitive

outcomes, building on different countries' conceptualizations, TALIS and PISA frameworks, and teaching effectiveness literature.

Different aspects of context is highlighted across studies. Some studies discussed results in relation to structural and cultural contexts, yet disregarded classroom and academic contexts. Others mentioned context only as sample background variables, or ignored context as important in claiming that the effect of (their conceptualization of) teaching quality is universal. In GTI, academic context is important as it focuses only on quadratic equations. Yet while cultural and structural context are considered imperative for understanding teaching, there is no information on these context levels, and classroom context (e.g., previous achievement) is included in the analysis only to control for moderation within the relationship between teaching and learning.

The ICALT studies, focusing mostly on measurement properties, draw *research implications* that ICALT can be used reliably and validly in comparative classroom studies. Some studies also suggest broader *policy and pedagogical implications* of results such as supporting international benchmarking efforts and setting criteria for school evaluations, as ICALT and ISTOF scores are considered to partly explain student achievement differences across countries. While GTI asserts that it is not ranking teaching quality nor promoting copy-pasting of practices, it suggests policy and practice implications in stating that identifying 'what works' is the goal with correlating observations with student outcomes.

Significance & relevance

The comparative education provide valuable perspectives to increase the usefulness, relevance and credibility of comparative classroom studies, for example regarding how to treat conceptualizations of teaching quality, context and implications in sensitive ways. Such approach is also argued important in using comparative research as a means to improve teaching and not only measure it.

Reference

Alexander, R. (2009). Towards a comparative pedagogy. In R. Cowen & A. M. Kazamias (Eds.), *International handbook of comparative education* (pp. 923-942). Dordrecht: Springer.

Bell, C. (2020). The development of the study observation coding system. In OECD (Ed.), *Global teaching insights: Technical reports*. Retrieved from https://www.oecd.org/education/school/GTI-TechReport-Chapter4.pdf

Bell, C., Klieme, E., & Praetorius, A.-K. (2020). Conceptualizing teaching quality into six domains for the study. In OECD (Ed.), *Global teaching insights: Technical report*. Retrieved from https://www.oecd.org/education/school/GTI-TechReport-Chapter2.pdf

Bell, C. A., Dobbelaer, M. J., Klette, K., & Visscher, A. (2019). Qualities of classroom observation systems. *School Effectiveness and School Improvement*, 1-27. doi:10.1080/09243453.2018.1539014

Castellano, K. E., & Bell, C. A. (2021). *Video component score characteristics*. Retrieved from https://www.oecd.org/education/school/GTI-TechReport-Chapter19.pdf

Cowen, R. (2014). Comparative education: stones, silences, and siren songs. *Comparative Education, 50*(1), 3-14. doi:10.1080/03050068.2013.871834

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *J Adv Nurs, 62*(1), 107-115. doi:10.1111/j.1365-2648.2007.04569.

Kyriakides, L., Creemers, B. P. M., Teddlie, C., & Muijs, D. (2010). The International System for Teacher Observation and Feedback: A Theoretical Framework for Developing International Instruments. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education (Third Edition)* (pp. 726-734). Oxford: Elsevier.

Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., . . . Fadhilah, N. (2020). Observed teaching behaviour in secondary education across six countries: measurement invariance and indication of cross-national variations. *School Effectiveness and School Improvement*, 1-32. doi:10.1080/09243453.2020.1777170

Maulana, R., Smale-Jacobse, A., Helms-Lorenz, M., Chun, S., & Lee, O. (2019). Measuring differentiated instruction in The Netherlands and South Korea: factor structure equivalence, correlates, and complexity level. *European Journal of Psychology of Education*. doi:10.1007/s10212-019-00446-4

Miao, Z., Reynolds, D., Harris, A., & Jones, M. (2015). Comparing performance: a cross-national investigation into the teaching of mathematics in primary classrooms in England and China. *Asia Pacific Journal of Education, 35*(3), 392-403. doi:10.1080/02188791.2015.1056593

OECD. (2020). *Global Teaching InSights: A video study of teaching*. Retrieved from Paris, France: https://www.keepeek.com//Digital-Asset-Management/oecd/education/global-teaching-insights_20d6f36b-en#page3

OECD. (2019). *The TALIS Video Study Observation System*. Retrieved from http://www.oecd.org/education/school/TALIS_Video_Study_Observation_System.pdf

Opfer, V. D. (2020). *An overview of the study*. Retrieved from OECD: https://www.oecd.org/education/school/GTI-TechReport-Chapter1.pdf

Phillips, D., & Schweisfurth, M. (2014). *Comparative and international education: An introduction to theory, method, and practice* (2nd ed. ed.). London: Bloomsbury.

Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: looking back and looking forward. *ZDM : The International Journal on Mathematics Education, 50*(3), 535-553. doi:10.1007/s11858-018-0946-0

Praetorius, A.-K., Rogh, W., Bell, C., & Klieme, E. (2019). Methodological challenges in conducting international research on teaching quality using standardized observations. In L. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 269-288). London, UK: SAGE.

Schriewer, J. (2021). Comparison and explanation: A long saga. *Comparative Education, 57*(4), 445-451. doi:10.1080/03050068.2021.1982555

Schweisfurth, M., Thomas, M. A. M., & Smail, A. (2020). Revisiting comparative pedagogy: Methodologies, themes and research communities since 2000. *Compare*, 1-21. doi:10.1080/03057925.2020.1797475

Schweisfurth, M. (2019). Qualitative comparative education research: Perennial issues, new approaches and good practice. In L. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE handbook of comparative studies in education* (pp. 258-268). Los Angeles, USA: SAGE reference.

Suter, L. E. (2019). The status of comparative education research in the 21st century: An empiricist's views. In L. E. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE Handbook of comparative studies in education* (pp. 3-24). London, UK.: SAGE.

Teddlie, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006). The international system for Teacher Observation and Feedback: Evolution of an international study of teacher effectiveness constructs. *Educational research and evaluation, 12*(6), 561-582. doi:10.1080/13803610600874067

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*(3), 295-311. doi:10.1080/09243453.2013.794845

van de Grift, W. J. C. M., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School effectiveness and school improvement*, *28*(3), 337-349. doi:10.1080/09243453.2016.1263215

Xu, L., & Clarke, D. (2019). Validity and comparability in cross-cultural video studies of classrooms. In L. Xu, G. Aranda, W. Widjaja, & D. Clarke (Eds.), *Video-based research in education: Cross-disciplinary perspectives* (pp. 19-33). New York: Routledge.