

A Comparative Study of the CLASS and the Stallings Instruments on Teaching Evaluation in Developing Countries--With Chile As Case Study

Zhao Ying

Central China Normal University, Wuhan Hubei 430000, China

ABSTRACT. *Across the world, there is growing interest in observing teachers' classroom practice, for teacher performance evaluation and incentives, program impact evaluation, formative feedback to teachers, and research on the determinants of student learning. For observations to generate meaningful data, they must be carried out using valid instruments, standardized protocols, and trained observers. A number of different observation instruments have been used in the US, but only two have gained any traction to date in developing countries-the Classroom Assessment Scoring System, CLASS, and the Stallings classroom observation instrument.*

KEYWORDS: *Classroom observation instruments; Measures of effective teaching in developing countries; Class; The stallings classroom observation instrument*

1. Introduction

The CLASS, established by Robert Pianta, is an instrument which rates teaching efficiency and teacher-student dynamics in classroom. It generates qualitative ratings of teaching on a scale from 1-7 through 3 broad domains: "emotional support, classroom organization, and instructional support" (Pianta, Hamre, and Mintz, 2012). The Stallings instrument, developed by Jane Stallings, produces quantitative data on teacher-student interaction in classrooms across 4 main variables: "Teachers' use of instructional time, Teachers' use of materials, Core pedagogical practices, and Teachers' ability to keep students engaged" (World Bank Group, 2015, p2).

The paper exploits the consistency between Stallings method and CLASS in terms of teacher's classroom management and scores in their own domains. My focus is to explore the significance and limits of the research, as well as its methodology.

2. Significance and Limits of This Research

According to the authors, the paper is the first comparative research between CLASS and Stallings instruments in a developing country setting, which provides developing countries with more references in terms of teaching monitoring and is enlightening for further research in this regard. The CLASS is a more complicated system and "requires operating personnel receiving training for certification and annual reexamination" (WANG, 2014, p44). It entails necessary localization in terms of language and curriculum. The exploration of the more economical Stallings method presents a pragmatic alternative for developing countries, because Stallings instrument is language and curriculum-neutral and observers are not required of much training. "This instrument's relative simplicity makes it suitable for large-scale samples in developing country settings" (Jukes, Vagh, and Kim, 2006).

Some limits are unavoidably spotted. The lack of the 12th dimension of CLASS by the previous Chilean research team results in a failure in comparing student engagement. The Stallings instrument research in Chile was conducted with recorded videos. Some classrooms are crowded so that camera angles could not guarantee a 360-degree cover of all students. Many details such as the students' absent-mindedness, non-engagement, and misbehavior that might be meaningful for the research could be missed out. The 51 Chilean teachers picked up are from Santiago Metropolitan region and two adjacent regions which is the capital city and developed region in the country. The 51 teachers are ranked in the middle (basic/competent) by the Docentemas (Chie's teaching evaluation system). Therefore, its representativeness is in question. Even so, the authors drew such conclusion as

“Chile may be ready to tackle the more complex teaching skills because efficient classroom management is largely achieved” (Bruns, Gregorio, and Taut, 2016, p29).

3. A Critical Review of Research Methodology

The paper employs many research methods, such as literature research, comparative research, and case study.

Through reviewing CLASS researches conducted in the US and six Latin American and Caribbean countries, the authors figured out that teacher’s individual effectiveness differs widely, has critical significance in student’s outcomes, is not necessarily indicated by teacher’s formal qualification, and is correlated with teacher’s different practice in classroom.

By comparing the US MET and ELA teachers and Chilean math teachers observed with CLASS, the authors found the general profiles are basically similar. Worth mentioning is that only math class was observed in Chile, while math and language classes were observed in the US. What’s more, 698 teachers are observed in the US while only 51 for Chile. The difference in sample number poses a problem. Again, the authors concluded that observed Chilean math teachers boast a higher proficiency in ICT than LAC counterparts. However, the authors did not pinpoint the specific locations in which similar observations were conducted in LAC countries. If they took place in remote and underdeveloped regions in LAC countries, the significance of such comparison might be jeopardized.

The paper attaches great importance to case study. For comparing the consistency of CLASS and Stallings instruments in developing country context, the authors conducted a Stallings instrument based research using videos recorded in Chilean math classroom. The objectives of such case study is to figure out the consistency of Stallings and CLASS instruments in measuring teacher performance. However, the research followed CLASS protocol of not coding immediately when the video started. This choice yields the result where Chilean math teachers registered in this sample a very high 85% of time on instruction, which equals to good practice benchmark developed by Stallings throughout the US. Such a result can not be regarded as well-based.

4. Conclusion

The paper compared CLASS and Stallings instruments on Chilean sample class, which explores the consistency of the two teaching observation instruments in their own domains. The research found teachers’ scores on the CLASS classroom organization domain were modestly correlated with the most important Stallings measures such as teachers’ use of time for instruction, classroom management, and teachers’ ability to keep students engaged. However, Stallings measures were not correlated with teachers’ scores on the other two CLASS domains of instructional support and emotional support.

The paper points out the advantages and disadvantages of the two instruments. Therefore, it provides a reference for developing countries in effort to effectively observe classroom dynamics and teacher-student interaction, as well as evaluate teachers’ performance. Although the research exposed several defects, it still is enlightening for further exploration and research in this regard.

References

- [1] Barbara B, Soledad G, Sandy T (2016). Measures of Effective Teaching in Developing Countries, RISE Working Paper, no.9, pp.1-29.
- [2] Jennings J L, T A DiPrete (2010). Teacher Effects on Social and Behavioral Skills in Early Elementary School. *Sociology of Education*, no.4, pp. 135-159.
- [3] Jackson C Kirabo, J Rockoff, D O Staiger (2014). Teacher Effects and Teacher-Related Policies. *Annual Review of Economics*, no.6, pp. 1-34.
- [4] Jukes M, S B Vagh, Y S Kim (2006). Development of Assessments of Reading Ability and Classroom Practice, Unpublished manuscript, World Bank, Washington, D.C.
- [5] Pianta R C Hamre, B K, Mintz S (2012). Classroom assessment scoring system: Secondary manual, Teachstone.
- [6] Lei W (2014). Implementation of Classroom Assessment Scoring System in Quality Evaluation of American Early Childhood Settings and its Implications, Nanjing: *Studies in Foreign Education*, vol.11, no.41, pp.44-45.
- [7] World Bank Group (2015). Conducting Classroom Observations--User Guide, no.2, pp. 18-19.