Testing Measurement Invariance of Quality Rating Causal Models in Tutorial-Based Assessment

Sukolrat Ingchatcharoen, Kamonwan Tangdhanakanond and Shotiga Pasiphol
Chulalongkorn University, Thailand
The first author is grateful for financial support throughout the 90th Anniversary of Chulalongkorn University Fund (The Ratchadaphiseksomphot Endowment Fund).

Tutorial-based assessment in problem-based learning (PBL) by raters, namely, tutor-, peer- and self-raters, were considered to have significance in recommending guidelines for judging student ability, including course efficiency and educational arrangements. However, inconsistencies across the ratings from tutors, peers and students were usually encountered in practice. Therefore, this study focused on testing the measurement invariance of a quality rating causal model in PBL across tutor-, peer-, and self-raters to compare similarity and differences in research model variables or parameters according to rater groups.

Methods

Participants
- Sample units were counted as PBL groups from five public Thai universities.
- Totally 120 PBL groups, each of which was composed of one instructor and two students (simple random sampling was conducted from a list of approximately 8 to 12 students in each group in order to obtain two students per group in order to have one student perform self-ratings and another student perform peer-ratings).

Measures
- The first set of questionnaires for raters contained 39 items, which measured seven variables consisting of rater goals, ability for rating, conscientiousness (modified from Goldberg’s (2001) IPIP Scales for the construct of conscientiousness), rater’s motivation, accountability, perception of PBL standards and the comparison process.
- The second set included the tutorial-based assessment form. This assessment was used to measure the rater error variable and contained 19 items with a 4-level response scale of “Needs Improvement”, “Moderate”, “Good”, and “Very Good”.

Data Analysis
The rater error or severity-leniency analysis was performed from the tutorial-based assessment scores of the tutor, peer and student rater groups using Facets version 3.71.4. All data was then analyzed to determine the measurement invariance of the quality rating causal model in the tutorial-based assessment using LISREL version 8.52.

Results
In results, an unrestricted baseline model was acceptable ($\chi^2 = 34.804$, df = 54, p = .980, RMSEA=0.000, and CFI = 1.00). The chi-square different test between Models 1 and Models 2 was significant (p < .01), indicating a significant different across tutor-, peer- and self-raters. For this comparison, the CFI also indicated that a substantial change in fit had occurred (1.00 vs. 0.98). This results indicated that the factor loadings were not invariance across tutor-, peer- and self-raters.

Conclusion
In conclusion, to produce quality ratings in tutorial-based assessment in PBL by tutors, peers and students, rater context factors consisting of rater’s motivation, accountability, conscientiousness, rater goals and ability for rating should be developed. Future research could use multilevel analysis techniques to more systematically understand factors involving quality ratings such as rating instrument, assessment techniques, and assessment criteria in studies within the model.

References