

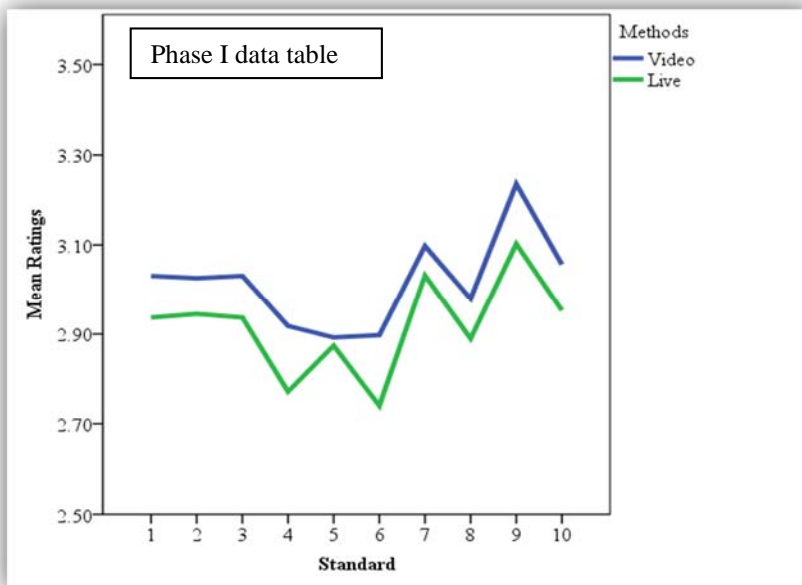
OVERVIEW

Beginning fall of 2013, many EPPs across the state of Georgia began using the Teacher Assessment on Performance Standards (TAPS) as a summative assessment at the end of the clinical practice. As a result, each EPP was encouraged by the Georgia Department of Education (GaDOE) to have at least one faculty or staff member complete the Teacher Keys Effectiveness System (TKES) credentialing. Further, to make a clear distinction between the state’s valid and reliable evaluation system, including the TAPS, and the EPPs’ use of the standards and rubrics, the pre-service instrument was named the Intern Keys.

National and state accreditation of educator preparation programs expects EPPs to use valid and reliable instruments. The adaptation of the TAPS for the Intern Keys meant that validity and reliability would need to be established to support the accreditation process for EPPs across the state. The Georgia Network for Transforming Educator Preparation (GaNTEP), in collaboration with the Council of Chief State School Officers (CCSSO), awarded a grant to the University of Georgia to conduct this project. The project was conducted in two phases with Phase II being a replication of the first phase. This brief provides an update on the project with results from Phase II of the project.

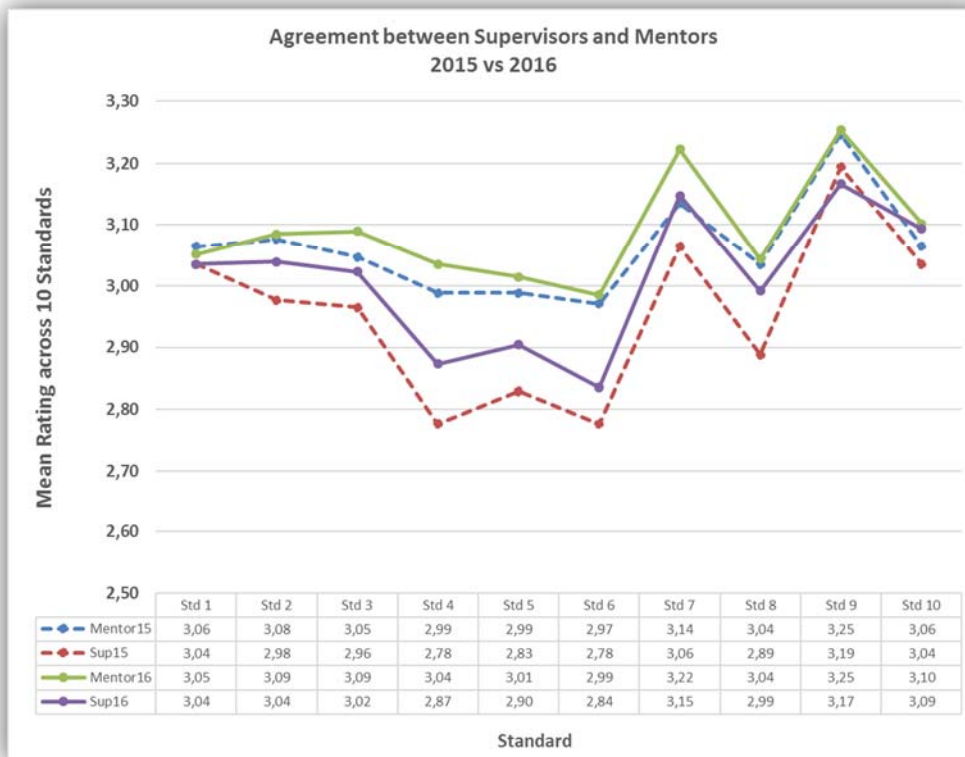
WHAT WE LEARNED

- Overall, the Intern Keys instrument has a high reliability. The results are consistent across different methods. Over time, the level of agreement between two raters, supervisors and mentor teachers, has increased.
- Among the 10 standards, Assessment Uses and Differentiated Instruction are given lower scores by raters, indicating that these are more difficult for candidate teachers to achieve. In contrast, professionalism is the easiest standard to meet.



- The result shows raters who only watched video tended to rate significantly higher than those who are trained with live method(s) on those standards.
- The results indicate high reliability of internal consistency among all 10 standards. Also, the internal consistency of mentors and supervisors is equal, indicating that professionals in these two roles apply the instrument in a similar way.
- The generalizability coefficient indicates the Intern Keys instrument is reliable for generalization.

COMPARISON OF RATINGS FOR PHASE I AND II



- There is a general increase in ratings, but more importantly, a decrease in mentor and supervisor differences.
- Standards 9 level of agreement saw the biggest reduction in the gap between raters.
- Standard 6, Assessment, continues to be the most challenging standard for candidates however, average ratings have improved as well as level of agreement between raters for this standard.

WHERE WE GO FROM HERE

The Phase II report for this project provides some suggestions for future studies related to the validity and reliability of the Intern Keys.

**Suggestions for future studies**

- Use edTPA data as a solid criterion measure
- Seek access to TEM data to test hypotheses that candidate variables – gender, ethnicity, content specialty, level of degree, geographic locality, etc. – may predict some of the variance in Intern Keys scores.
- Examine the demographic and professional characteristics of the evaluators to discover any correlates and reveal any biases.
- Consider predictive validity of the Intern Keys when compared to the Teacher Keys TAPS score.
- Conduct a qualitative analysis of the evidence cited with each rating to further explore levels of agreement.

To view the full report, visit <http://epr.coe.uga.edu/evaluation-systems/intern-keys-validation-project/>.