



Validity Evidence Using Generalizability Theory for an Objective Structured Clinical Examination (OSCE): a multiple occasions rationale

Michael J. Peeters, PharmD, MEd, FCCP, BCPS*; M. Kenneth Cor, PhD, MEd, BScEng#;
Bryan M. Bishop, PharmD, BCPS*; Sarah E. Petite, PharmD, BCPS*; Michelle N. Schroeder, PharmD, BCACP, CDE*

*University of Toledo College of Pharmacy & Pharmaceutical Sciences, Toledo OH, USA

#University of Alberta, Faculty of Pharmacy & Pharmaceutical Sciences, Edmonton AB, Canada



Purpose

- To evaluate how our reliability of OSCE grading changed depending on number of stations over multiple weeks/occasions

Key Findings [Implications]

- This OSCE was successfully modeled with multivariate G-theory
- Students accounted for only some variation in OSCE scores
 - Variation in scores from other influences as well (station, rater, week)
- Our reliability improved by increasing the number of stations each week and/or number of weeks

Why did we do this study? [Background]

- OSCEs are gold-standard for performance assessment (skills)¹
 - Can be used to help evaluate PharmD students' practice-readiness
- Pharmacy colleges/schools should be generating validation evidence for assessments used to make important decisions; including OSCEs²
- Kane's Framework for validation³
 - Scoring > Generalization > Extrapolation > Implications
- Generalizability Theory (G-theory) is one method for reliability⁴
- Prior studies in pharmacy education rarely report the use of G-theory to produce reliability evidence as part of the validation process⁵

What did we do? [Methods]

- IRB approved
- Assessment of OSCE scores from 3rd-year PharmD students
 - In Spring Semester before their final-year rotations
 - Assessed skills with: counselling & communication, case presentation, prescription verifying, medication reconciliation, non-adherence, drug information, pharmacy calculations and general drug knowledge
- OSCE included 14 stations over 3 weeks with 4 or 5 stations per week
 - 1 or 2 stations per week were scored by faculty-raters
 - 3 stations per week required students' written responses
- mGENOVA software was used for Multivariate G-theory
 - $p \times w^o$ model used (Students crossed with stations nested in 3 weeks/occasions)

What did we find? [Results]

- Ninety-seven students completed this OSCE
- Stations were scored independently
- G-study: Our estimated g-coefficient (reliability) for the total score based on three weeks and approximately five stations per week was estimated at **0.74**

Our Variation Sources (%)

	OSCE Week1	OSCE Week2	OSCE Week3
student	17%	5%	28%
station	12%	14%	16%
p x w (and other error)	71%	81%	56%

Further analysis (decision studies) revealed how our reliability changed depending on combinations of stations per week

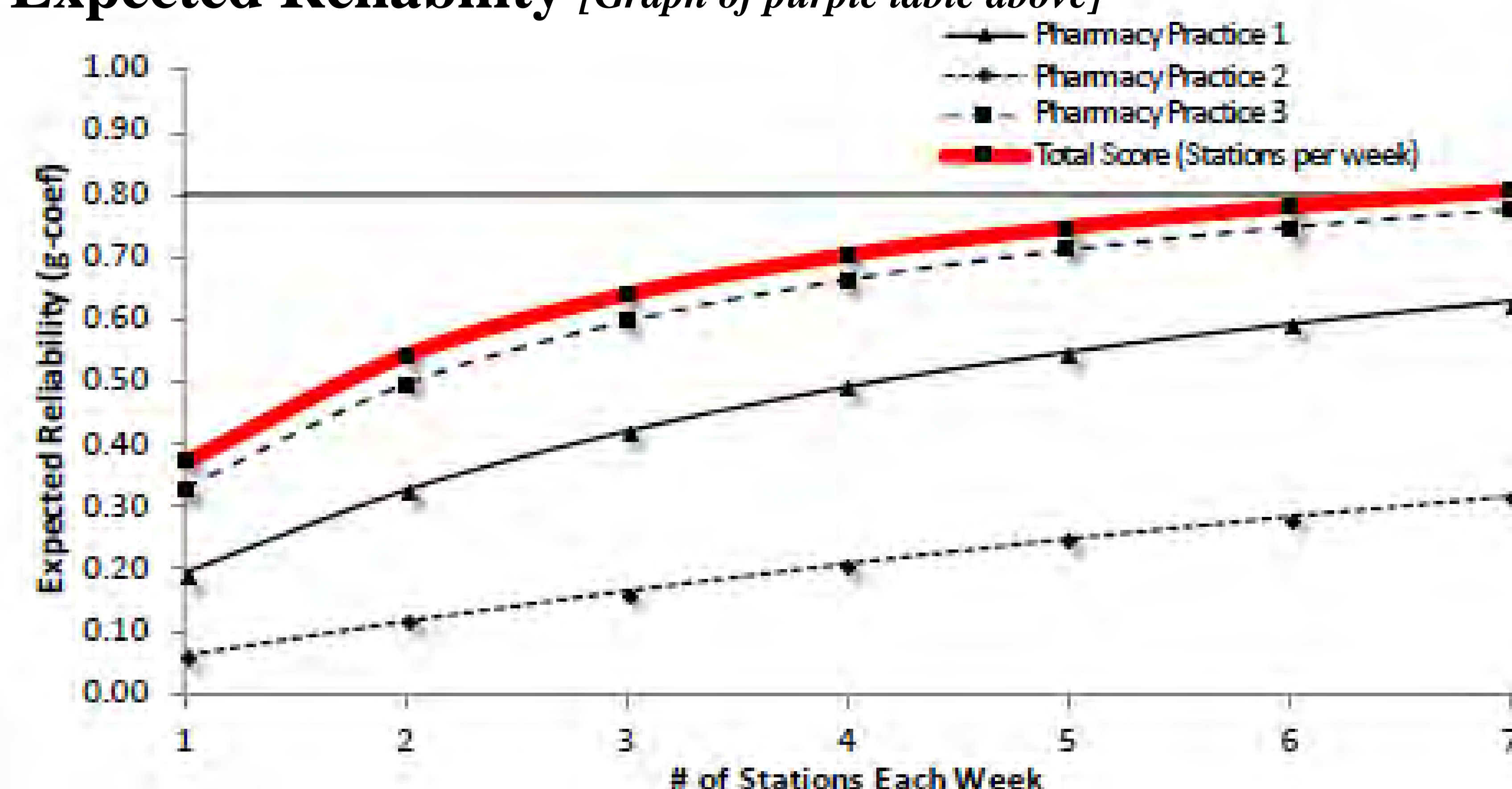
Decision studies (G-Coefficients Estimates)

Week#	Stations/week						
	1	2	3	4	5	6	7
Week 1	0.19	0.33	0.42	0.49	0.55	0.59	0.63
Week 2*	0.06	0.12	0.17	0.21	0.25	0.28	0.32
Week 3	0.33	0.50	0.60	0.66	0.71	0.75	0.78
Total-Score	0.37	0.54	0.64	0.70	0.75	0.78	0.81**

*Improving stations especially in Week #2 should also help improve our total-score reliability

**Threshold=0.80

Expected Reliability [Graph of purple table above]



References

- Bursicot KAM, Roberts TE, Burdick WP. Structured assessments of clinical competence. In: Swanwick T, ed.. *Understanding Medical Education: Evidence, Theory and Practice*. 2nd ed. Chichester, UK: Association for the Study of Medical Education; 2014: 293-304.
- Accreditation Council for Pharmacy Education. Accreditation Standards and Key Elements for the Professional Program in Pharmacy Leading to the Doctor of Pharmacy Degree ("Standards 2016"). Published February 2015. Available at: <https://www.acpe-accredit.org/pdf/Standards2016FINAL.pdf>.
- Peeters MJ, Martin BA. Validation of learning assessments: a primer. *Curr Pharm Teach Learn*. 2017; 9(5):925-933.
- Cor MK, Peeters MJ. Using generalizability theory for reliable learning assessments in pharmacy education. *Curr Pharm Teach Learn*. 2015; 7(3):332-341.
- Hoover MJ, Jung R, Jacobs DM, Peeters MJ. Educational testing validity and reliability in pharmacy and medical education literature. *Am J Pharm Educ*. 2013; 77(10):article 213.