Module 5: Wrap-up of G Theory

Main topics

- Wrapping up the key concepts in G theory
- Going through some examples of the G study designs

Basic ideas of G theory
- multifaceted measurement errors

- G theory can identify
  - Measurement errors from various sources compared to CTT
  - Different types of measurement errors depending on the relative or absolute interpretations
  - Different types of measurement errors depending on whether the facets are fixed or random

Basic ideas of G theory
- universe and universe score

- Universe
  - Is defined as the collections of possible measurement collections that will be generalized, e.g., the infinite universe vs. restricted universe
  - Is represented by all the possible levels at the facets in which the researchers have interest
- Universe score
  - Is the estimate score for the universe
  - Can be considered as true score in the context of CTT

Basic ideas of G theory
- rel. and abs. interpretations

- Rel. and abs. interpretation
  - Measurement scores/results can be either used to index the absolute levels of performance or just rank/compare the levels
  - Different types of interpretations are associated with different ways to estimate the measurement errors and G coefficients

Basic ideas of G theory
- procedures to conduct G/D studies

- Obtain the variance component estimates from EMS in the ANOVA outputs
- Obtain the VC estimates from running statistical programs (e.g., Genova or SPSS)
  - When to use Genova, Urgenova, or SPSS?

Other issues (1): hidden facet

- Facet that is involved in the G study design, may or may not affect the variance components
  - Occasion facet in a p x i design, which could be consider as p x (i:o)
  - Item order facet
- We cannot obtain estimates for the VCs related to the hidden facet.

Other issues (2): nested and crossed facets

- A facet can be nested or crossed with other facets.
- For a give crossed design (one facet or multifaceted model), you can run a crossed model and different nested models. Often the crossed model yields the highest rel. G coefficient.
- So what about the person variance?

Other issues (3): random and fixed facets

- In the multifaceted model (and even one facet model), a facet can be either random of fixed.
- When a facet is fixed, we need to apply different rules to estimate the VCs, depending if we need to sample all or just some levels of the facet.

Other issues (4): future directions of G theory

Theoretical issues

- Integration of IRT and G theory in the estimation of variance
- Use of Bayesian network approach to estimate the VCs
- Exploration of matrix sampling (both of G and D contexts) in estimating VCs
- Estimation of co-variance structure

Applications of G theory (i)

- Description of the study: Each student (p) at a piano recital plays a musical selection (m) of his or her own choosing, with all students evaluated by the same 2 judges (j).
- Questions for discussion
  - What's the variance that researchers were interested in?
  - What is the design of the G study?

11

Applications of G theory (ii)

- Description of the study
  - The reliability of UW course evaluation system
  - 42 teachers (t) taught 2 different courses (c)
  - 14 students (p) from each class completed the evaluation form
  - The form included 10 items (i)
- Questions for discussion
  - What's the variance that researchers were interested in?
  - What is the design of the G study?

12

Applications of G theory (iii)

- Students participate in the engineering competition by designing the ROVs and preparing the reports.
- Their multiple projects are scored, for example, the technical report, the presentation, or the poster.
- The scoring rubrics were created to capture multiple aspects. Some are more relevant to engineering design and others are related to communication.
- We are interested in running the project x rater x aspect (for the scores relevant to the engineering design) to identify how the measurement procedure can be optimized.

Applications of G theory (iv)

- The assessment was initially created by the Center of Educational Leadership.
- Individuals are asked first to watch a 10-minute video and then respond to three open-ended questions regarding what they notice and wonder, what feedback they plan to provide to the instructor, and what pd they would like to plan.
- The current scoring rubric evaluates individuals' responses in four domains with multiple sub-rubrics.

Applications of G theory (iv)

In this project, we then:

- refined the assessment by asking each subject to watch two short clips of lessons
- compared the responses between two groups of leaders, principals and content specialists/coaches
- ran the g/d studies
  - using the p x video x rater
  - using the (p : group) x video x rater

Sharing: now your examples

- Describe the design of your study
- Share 2 or 3 main findings