



Analyzing Observational Ratings of Classroom Interactions: An Application of Generalizability Theory to inCLASS Scores

Abry, T. S., Booren, L. M., Meyer, J. P., & Downer, J.
University of Virginia



INTRODUCTION

- Observational measures can reduce the bias of self-report measures, and can be used to capture complex interactions in naturalistic settings unlike many direct assessments (Neisworth & Bagnato, 2004 ; Volpe, DiPerna, Hintze, & Shapiro, 2005).
- Information gathered through observational measures is limited to the time and setting of its administration. Multiple observations are often needed to ensure the data is representative of the subject(s) being observed.
- Generalizability theory (Cronbach, Gleser, Nanda, & Rajaratnam, 1972) can simultaneously account for multiple sources of error in observed and hypothetical data collection designs, as well as estimate the variance associated with each design facet and the interaction between them.

THE INCLASS

- The Individualized Classroom Assessment Scoring System (inCLASS; Downer, et al., 2010) is an observational measure assessing preschool children's competencies in relation to school readiness.
- Scored on a scale of one to seven with higher scores indicating higher levels of competency.
- Assesses preschool children's classroom interactions in ten dimensions categorized into four domains.
 - Teacher (Positive Engagement with Teacher & Teacher Communication)
 - Peer (Peer Sociability, Peer Assertiveness, & Peer Communication)
 - Task (Engagement, Self-reliance, & Behavior Control)
 - Conflict (Teacher Conflict & Peer Conflict)
- inCLASS data can be analyzed at child or classroom level because children are randomly selected from the classroom.

RESEARCH QUESTIONS

- RQ 1: For each of the four inCLASS domains, what amount of variance is attributable to the classroom (the object of measurement) and each study facet, including students nested in classroom, observation occasion, and inCLASS dimension (item) as well as the interaction between facets?
- RQ 2: For each of the four inCLASS domains, what is the optimal number of observation occasion and children observed per classroom needed to result in reliable estimates of preschool children's interactions in the four domains of teacher, peer, task, and conflict?

METHODS

- Participants
 - 99 classrooms
 - 198 children (two randomly selected from each classroom)
 - 50 % male
 - Average age 50.11 months (SD = 6.38)
 - 65% Hispanic/Latino, 21% Caucasian
- Children observed for three alternating 15 minute observation cycles and rated on the ten inCLASS dimensions at the end of each observation
- Analytic approach
 - In the first phase a generalizability study was conducted for each inCLASS domain (item crossed with observation crossed with children nested within classrooms; ios:c) where classroom is the object of measurement.
 - In the second phase a decision study was conducted for each inCLASS domain with increasing numbers of children observed per classroom (up to 10) and observation occasions (up to 12).

RQ1 RESULTS: G-STUDIES

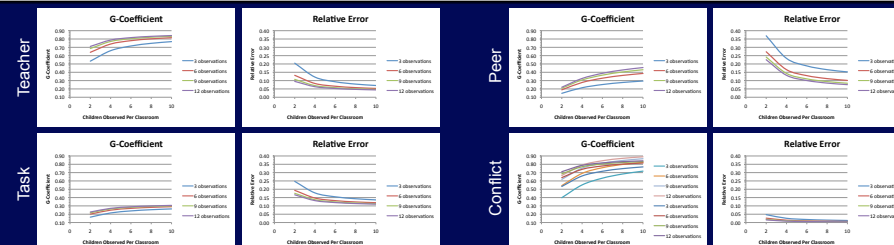
- Across domains, the highest proportions of variance were attributable to observation occasion crossed with students nested in classrooms (os:c) and item (i). Variance attributable to the object of measurement, classroom (c) was lower than anticipated. The residual error term (ios:c) also accounted for a substantial proportion of variance across the four domains.

Generalizability Study Results for the inCLASS Domains: i x o x s:c Design

Facet	Teacher		Peer		Task		Conflict	
	Variance Component	Percent of Total Variance	Variance Component	Percent of Total Variance	Variance Component	Percent of Total Variance	Variance Component	Percent of Total Variance
Classroom: c	0.2343	13%	0.0637	3%	0.0487	1%	0.0313	7%
Students nested in classrooms: s:c	0.0303	2%	0.3103	14%	0.0665	2%	0.0000	0%
Observation Occasion: o	0.0440	2%	0.0002	0%	0.0006	0%	0.0000	0%
Item: i	0.2878	16%	0.5011	23%	1.7892	54%	0.0069	2%
Classroom x Occasion: co	0.0000	0%	0.2228	10%	0.0448	1%	0.0000	0%
Classroom x Item: ci	0.0590	3%	0.0553	3%	0.2782	8%	0.0000	0%
Occasion x Student Nested in Classroom: os:c	0.6195	34%	0.5505	26%	0.3252	10%	0.1162	27%
Item x Student Nested in Classroom : is:c	0.0564	3%	0.0224	1%	0.0822	2%	0.0210	5%
Occasion x Item: oi	0.0037	0%	0.0000	0%	0.0105	0%	0.0000	0%
Classroom x Occasion x Item: coi	0.0451	3%	0.0471	2%	0.0047	0%	0.0214	5%
Item x Occasion x Student Nested in Classroom: ios:c	0.4226	23%	0.3715	17%	0.6838	21%	0.2290	54%
Total Variance	1.8027	100%	2.1450	100%	3.3344	100%	0.4258	100%

RQ 2 RESULTS: D-STUDIES

The generalizability coefficient ranged from 0.15 to 0.54 across inCLASS domains for the observed design (two students per class, three observations each). G-coefficients for a design in which 10 students per classroom are observed 12 times ranged from 0.31 to 0.89. Relative measurement error for the observed design ranged from 0.05 to 0.37 across domains and ranged from less than 0.01 to 0.11 in the most extreme hypothetical design.



DISCUSSION

- In the teacher and conflict domains, large g-coefficients would result from substantial increases in children observed per classroom or observation occasions, or from moderate increases to both.
- In the peer and task domains, g-coefficients were lower than anticipated, even in the most stringent of hypothetical designs where 10 children would be observed on 12 occasions.
- It is recommended that inCLASS data be reported at the dimension level for the peer and task domains. Due to high proportions of variance attributable to item.

- Variance attributable to the interaction between observation occasion and student nested in classroom may be diminished by limiting the types of activity settings observed, thereby decreasing task-sampling variability, or by stratifying observations according to the specific "tasks" or classroom activities observed.
- Future studies using the inCLASS will increase the number of children observed per classroom as well as the number of occasions on which they are observed. Additionally, children will be systematically observed across a variety of activity settings. This will allow for analyses in which activity setting is included as a facet.

ACKNOWLEDGEMENTS

This study is part of the Interagency Consortium for School Readiness Outcome Measures (ICSROM) and co-funded by the U.S. Departments of Health and Human Services and Education. The opinions expressed are those of the authors and do not represent views of the U.S. Departments of Health and Human Services or Education.

Poster presented at the National Council for Measurement in Education Annual Conference: April 29 – May 3, 2010 (Denver, CO).

Correspondence concerning this poster should be addressed to Tashia Abry, email tsa2@virginia.edu.