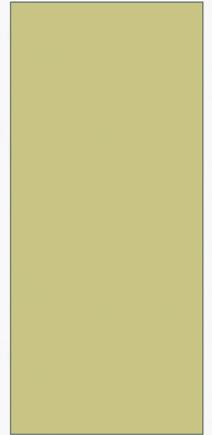


# Rater Nonindependence: A Methodological Problem in Universal Risk Assessments

MARISSA SMITH-MILLMAN, MA; PAUL D FLASPOHLER, PHD;  
MELISSA MARAS, PHD; JONI SPLETT, PHD; KRISTY WARMBOLD,  
MA



# OUTLINE

- 1.) Universal risk assessments: academic versus behavioral
- 2.) Possible problems with teacher-rated risk assessments
- 3.) Between-teacher rating differences
- 4.) What do we do about this?
  - a.) Explanation of the intraclass correlation (ICC)
  - b.) Example of this method
- 5.) Implications for research and practice

# UNIVERSAL SCREENING

**Tier III**

- Academic screening
  - DIBELS
  - Woodcock Reading
  - Curriculum-based
  - *All completed*

**Tier II**

- Behavioral screening
  - Systematic
  - Behavior
  - Strengthening
  - DFB
  - C

**Tier I**

(BESS)

# POSSIBLE PROBLEMS WITH TEACHER RATED RISK ASSESSMENTS

## **Possible problems**

- Teachers may interpret questions differently
- Some teachers may be high raters
- Others may be low raters
- Could be due to a non-random distribution of difficult students to certain classrooms or other classroom level influences

## **So what?**

- This may lead to the misidentification of some students and the non-identification of other students who are in need

# BETWEEN-TEACHER RATING DIFFERENCES

- Office discipline referrals (ODRs):
  - 2/3 of all ODRs in one middle school came from 25% of teachers (Skiba et al., 1997)
  - In an elementary school, 70% of school staff gave 1-5 ODRs, while 6.2% of the staff gave over 25 (Putnam et al., 2003)
- Mashburn and colleagues (2006) found that 15-33% of the variance in preschool teachers' ratings of their students' competence was due to teacher differences

# WHAT DO WE DO ABOUT THIS?: RESEARCH PERSPECTIVE

- Use multilevel modeling (MLM) techniques
- This enables us to do two things:
  - Determine the *degree* of rater difference in scores
  - Look at other variables while *accounting for* rater differences

# INTRACLASS-CORRELATION (ICC)

- How do you calculate this?
  - Run a multi-level model with no predictors to get 2 pieces of information: between-subjects variance and within-subjects variance
  - ICC = between-subjects variance / between-subjects variance + within-subjects variance

$$\rho = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

- What does it mean?
  - This gives you an estimate of the *degree* of rater-differences in a given set of measures

# EXAMPLE: DESSA-MINI DATA

## Covariance Parameters

*Within-subjects  
variance*

Estimates of Covariance Parameters<sup>a</sup>

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Residual	128.391653	7.261783	17.680	.000	114.919349	143.443352
Intercept [subject = Rater]	33.192809	9.748351	3.405	.001	18.666195	59.024487

a. Dependent Variable: DESSA1 score.

*Between-subjects  
variance*

$$\rho = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

$$33.192809 / (33.192809 + 128.391653) \approx$$

.205

# EXAMPLE: DESSA-MINI DATA CON'T

- In two samples (one from OH and one from MO), we examined the ICC in DESSA-Mini ratings
  - In the Ohio sample,  $ICC = .205$
  - In the Missouri sample,  $ICC = .184$
- What does this mean?
  - In the Ohio sample, we know that 20.5% of variance in DESSA-Mini scores is due to either teacher- or classroom-level differences, NOT differences between students
  - In the Missouri sample, we know that 18.4% of variance in DESSA-Mini scores is due to either teacher- or classroom-level differences, NOT differences between students

# THE MULTI-LEVEL MODEL

- MLM can be used to answer other questions, while controlling for rater-differences
- By entering the rater into the model at level 1, you can control for rater-differences to ask questions about universal risk assessment data such as:
  - Does risk assessment data predict academic achievement?
  - Is risk assessment data related to discipline referrals?
  - And many more
- Controlling for rater-differences will provide more accurate, less biased results

## IMPLICATIONS FOR RESEARCH

When using universal screening data completed by teachers, use *MLM* to control for possible teacher or classroom effects

# IMPLICATIONS FOR PRACTICE

- Possible that some at-risk kids are not being identified and not getting necessary treatment
- Possible that kids who are not really at-risk are being misidentified and given treatment they do not need
- So, what should you do?
  - Multi-rater assessments of students identified as “at-risk” by universal screening
  - Training for teachers on how to complete the universal screener
  - Use a variety of indicators
  - Present information to school staff to identify classrooms in need of more support

THANK YOU

QUESTIONS?