



Comparing Intervention Fidelity Measures

Kristin Duppong Hurley & Matthew Lambert, UNL

Mark Van Ryzin, OSLC

May 16, 2013

Seattle Implementation Research Conference



Funding sources

- This project was funded by the **NIMH** grant #R34MH080941
- With support from **IES**, Department of Education grant R324B110001
- Dr. Duppong Hurley is fellow with the **Implementation Research Institute (IRI)**, at the George Warren Brown School of Social Work, Washington University in St. Louis; through an award from the National Institute of Mental Health (R25 MH080916-01A2) and the Department of Veterans Affairs, Health Services Research & Development Service, Quality Enhancement Research Initiative (QUERI).



Goal: Comparison of Fidelity Measures

Primary Focus

- Psychometrics
- Comparability
- Prediction

Application

- Research
- Sustainability in Practice



Where did I study these simultaneously?

Boys Town
Family-style homes
Teaching Family Model

60+ homes
6-8 youth/home

Disruptive Behavior Diagnosis
10-18 years of age



Study Enrollment

- Family Teaching (FT) Homes
 - 64 (81%) participating
- Supervisors of FT's
 - 23 (96%) participating
- Youth
 - 145 (68%) participating



Six different fidelity assessments



Goal: Usefulness and inter-changability

Psychometrics * Comparability * Prediction



Basic Psychometrics

- Good distribution of use of response options
- Internal consistency scores on conceptual scales are acceptable:

Supervisor ratings (TC=.97, MS=.93, SG=.92, RBFS=.96)

FT Self-ratings (TC=.87, MS=.83, SG=.89, RBFS=.91)

Youth ratings (TC=.86, MS=.84, SG=.78, RBFS=.95)

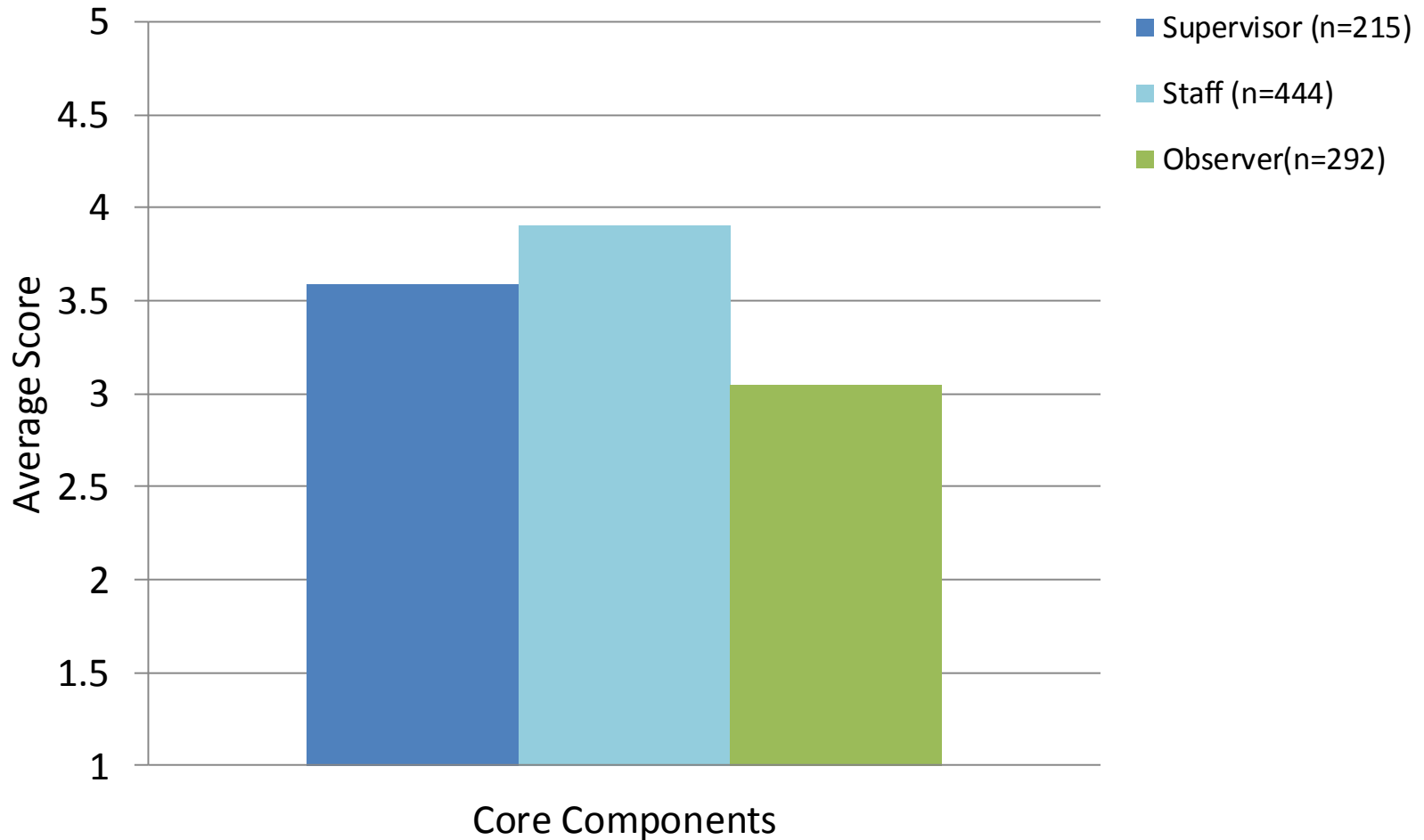
External Observer Obs 6 (TC=.89, MS=.64, SG=.72, RB=.78)

Int. Observer (TC=.88, MS=.79, SG=.81, RBFS=.86)

- EFA and CFA suggest a single implementation factor for each measure

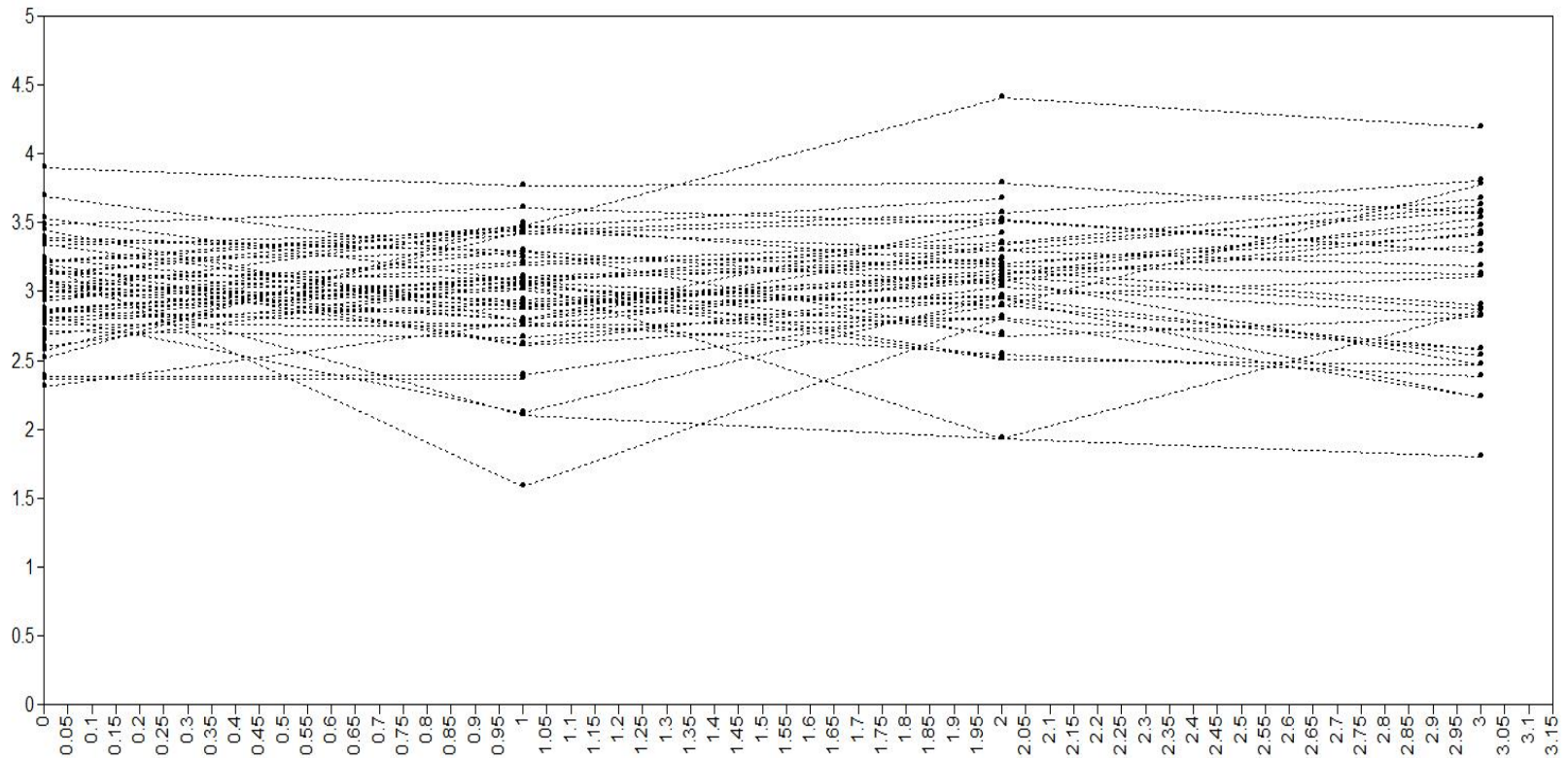


Ratings of Home by Assessment



Observer Ratings of Home - Longitudinal

Core (N = 53)



Goal: Usefulness and inter-changability

Psychometrics * Comparability * Prediction



Do Ratings Differ by Supervision Unit?

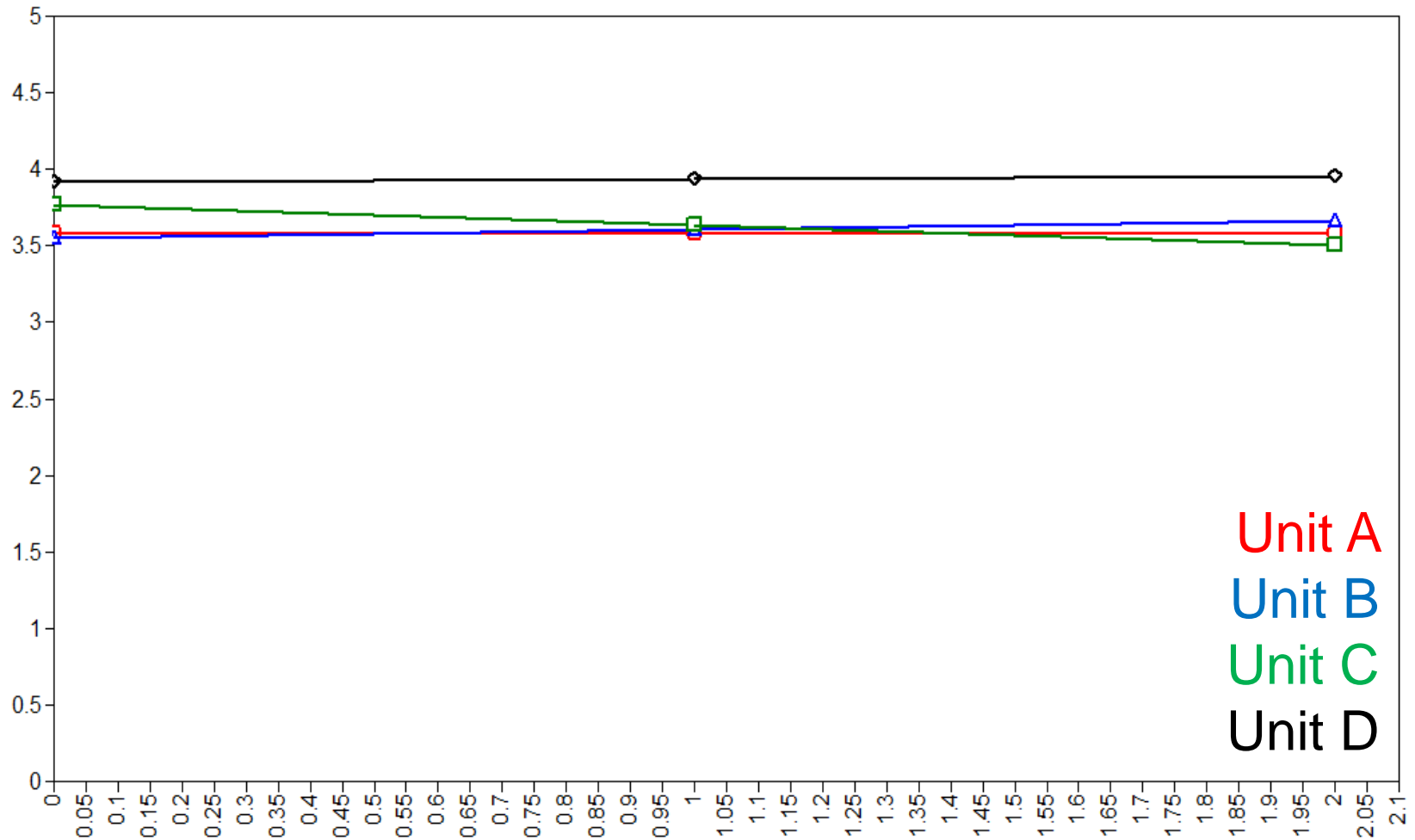
Curious if supervisory units had different patterns of fidelity levels...



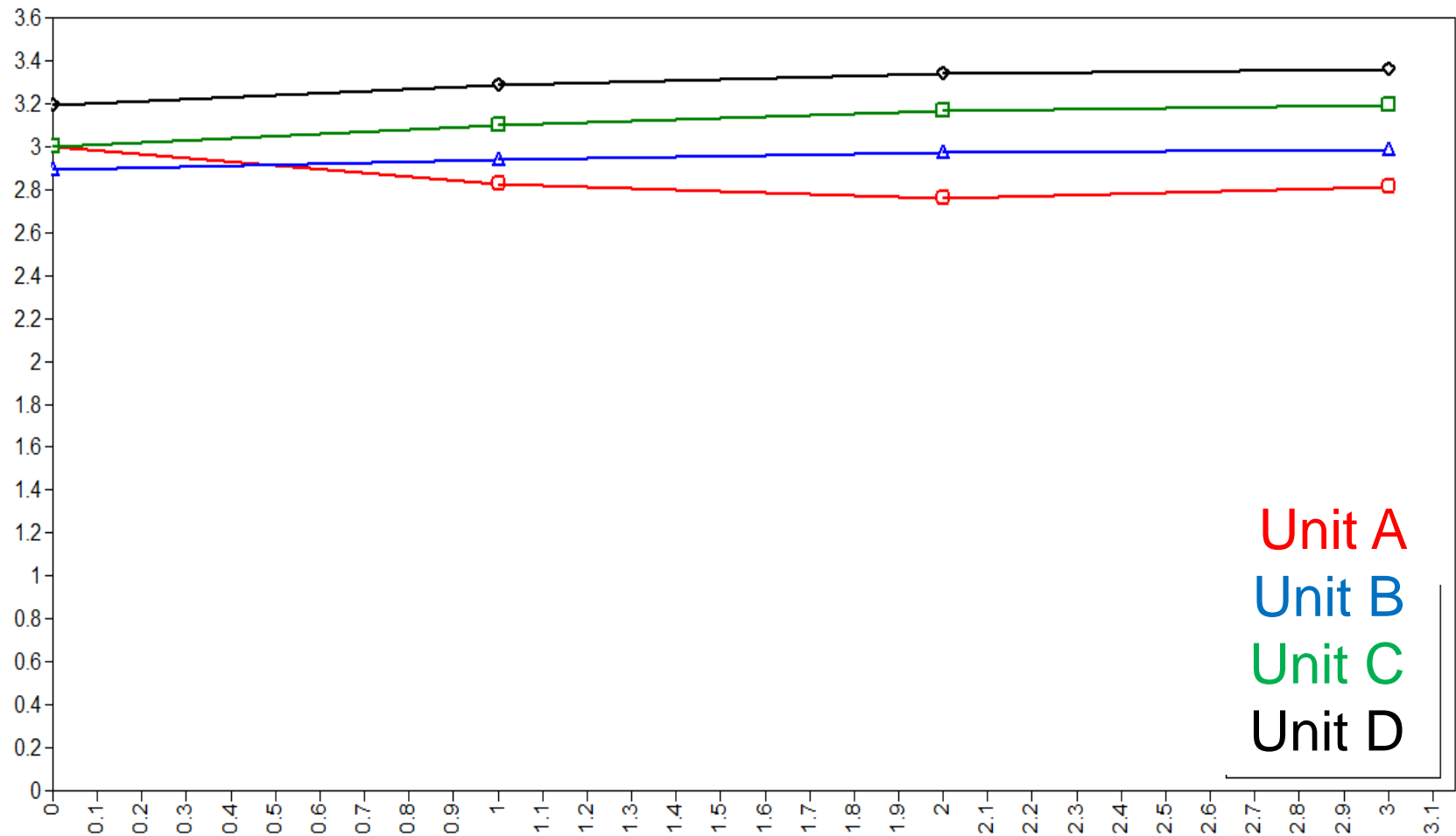
...do any units need additional training?



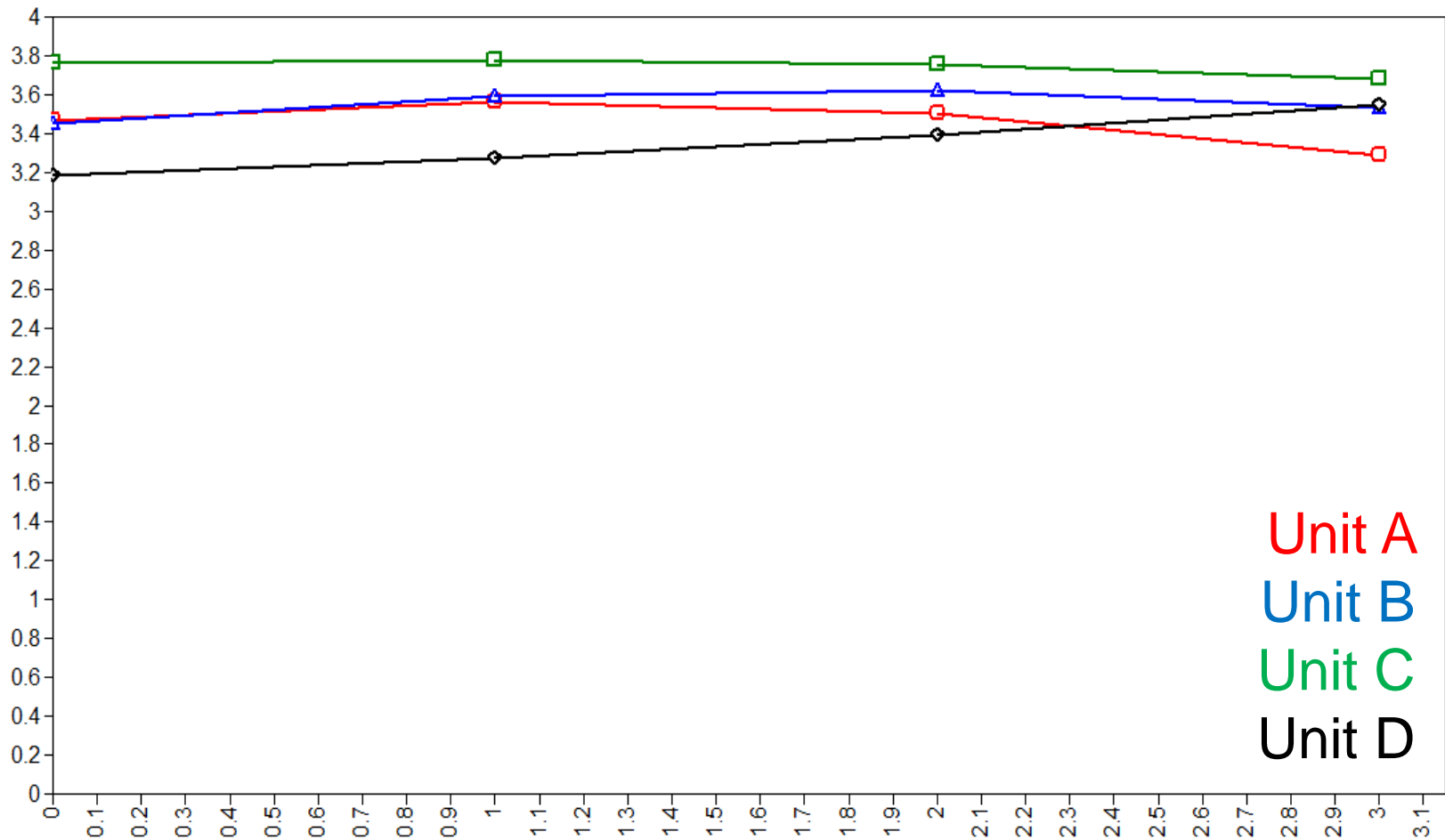
Ratings by Unit – Internal Observation



Ratings by Unit – External Observation



Ratings by Unit - Supervisor



Do Ratings Differ by Unit?

It seems that some raters are more harsh/lenient than others

Caution for common supervisor-type evaluations

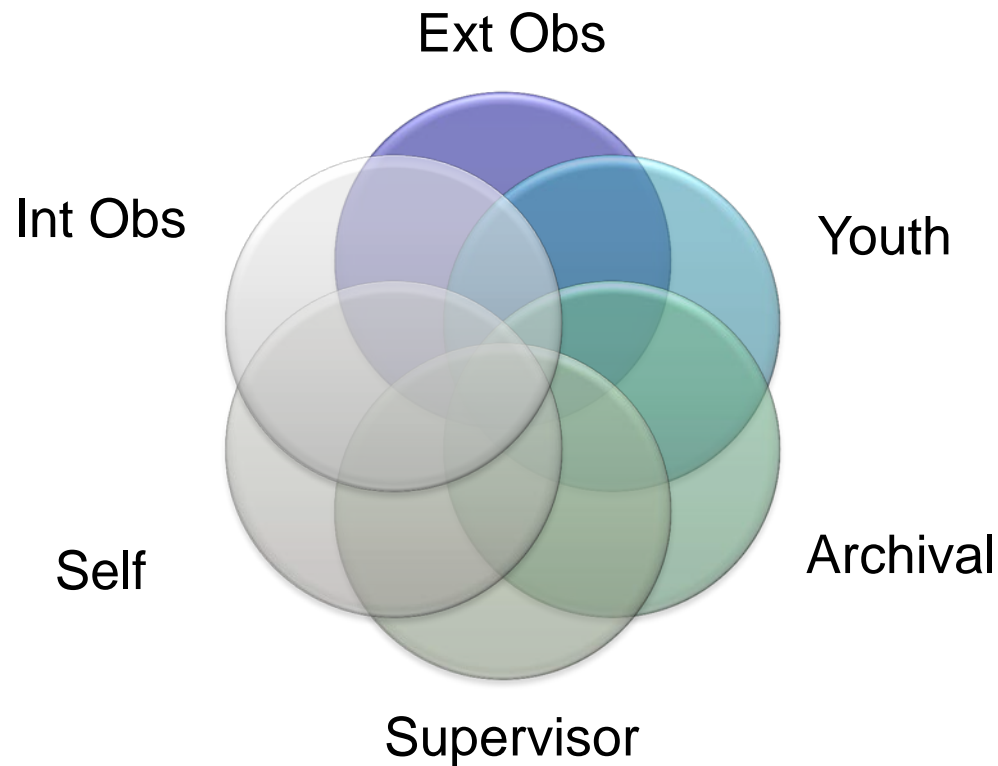


Goal: Usefulness and inter-changability

Psychometrics * Comparability * Prediction



Correlations Among the Measures



Fidelity Measure Correlations

	Self	Superv.	P Card	Internal	Youth	External
Staff Self Ratings	-					
Supervisor Ratings	.15					
Point Cards % positive	-.08	-.17	-			
Internal Observation	.31*	.56*	-.03	-		
Youth Ratings	.27*	.12	-.16	-.09	-	
External Observation	.07	.19*	-.07	.27*	-.06	-

At one point in time (similar for other time points, with some fluctuations)



Goal: Usefulness and inter-changability

Psychometrics * Comparability * Prediction



Study Outcome Measures

- Staff & Youth Rated Behavior Measures
 - **Child Behavior Checklist (CBCL: Ext, Int) Staff ONLY**
 - Symptom Functioning & Severity Scale (SFSS: Ext, Int)
 - Behavior Emotional Rating Scale (BERS: SI)
- Archival/Youth Records
 - Youth behavior incidents (2 months intervals)
 - Number of psychotropic medications

All assessments at intake, 6, and
12 months, but incidents



Fidelity Measures Prediction of CBCL Mental Health Outcomes (N=112)

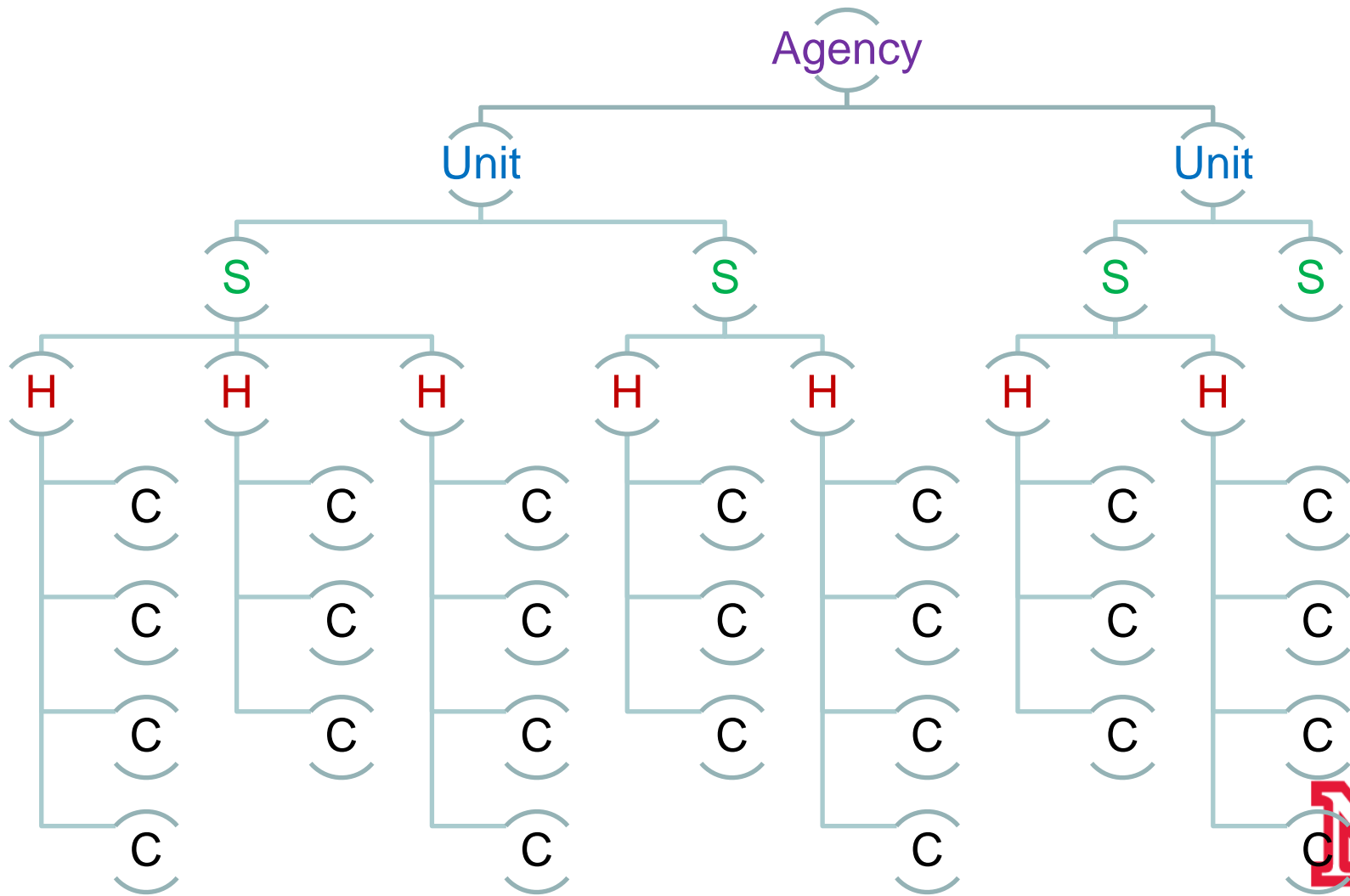
Fidelity Measure	Family Teacher Rated CBCL Outcomes	
	6 months	12 months
Youth-rating	CBCL (-) Int, Ext	CBCL (-) Ext
Point Card--% positive interaction	CBCL (-) Int, Ext	
Staff Self-rating		
Supervisor	CBCL Ext (+)	CBCL Ext (+)
External Observation	CBCL Int ,Ext (+)	CBCL Int (+)
Internal Observation		

Green= higher fidelity and improved outcomes

Red=higher fidelity and worse outcomes



Nesting Issues With Assessing Fidelity

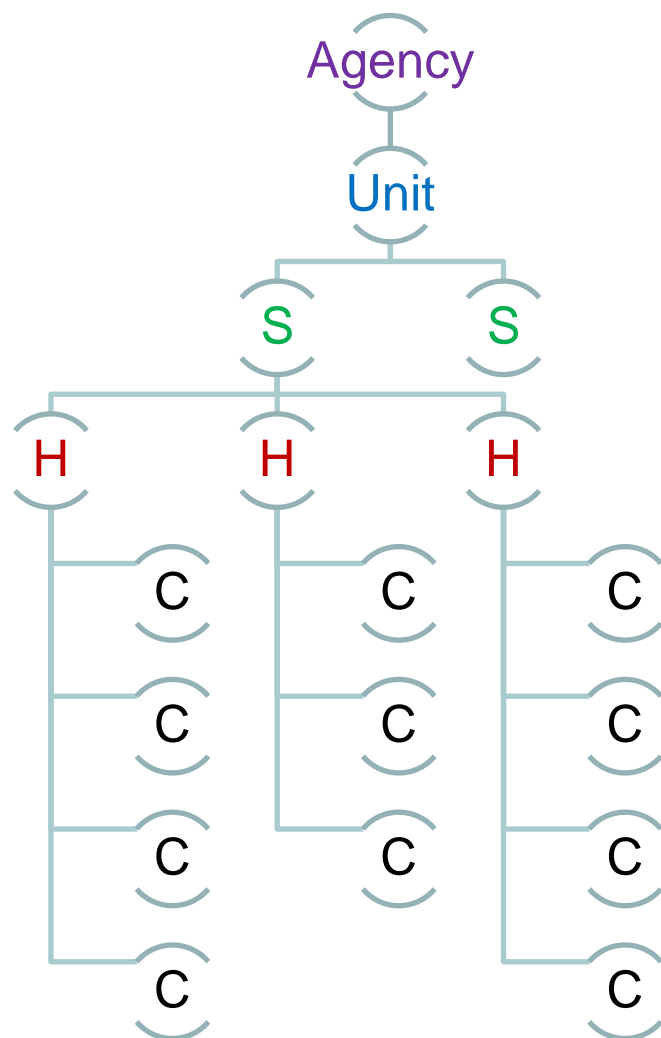


HLM Analysis of Fidelity

- Found group-level of rating of fidelity (external observations) did not predict CBCL scores
- But individual-level ratings of fidelity (youth ratings) correlated with CBCL scores
- Corresponds with Zvoch, 2012



Consider issue of level with fidelity



- Supervisors common source, may have bias
- Provider is typical focus
- How often? Who?
- Is it feasible to collect at client level? How often? Method?



Implications

- Fidelity measures can vary substantially
- Try to include multiple assessments
 - Different respondents
 - Different “levels” (group & individual)
- Challenges for sustainability
 - Supervisor most likely respondent during care
 - Costly to assess fidelity at individual-level



Many Thanks to...

- The youth and staff at Boys Town that made this study possible
- The graduate and undergraduate students that have worked on the project
- All of those that helped consult on the study, especially Michael Epstein, Betsy Farmer and John Landsverk





© 2007, THE BOARD OF REGENTS OF THE
UNIVERSITY OF NEBRASKA. ALL RIGHTS RESERVED.
