The Wisdom of Crowds
Collaboratively Developing and Establishing the Validity & Reliability of a Student Teaching Evaluation Form

Valid and Reliable Instruments for Educator Preparation Programs (VARI-EPP)
Candidate Preservice Assessment of Student Teaching (CPAST): Spring 2017
Rubric and assignments may not be shared without permission

- Pedagogy Evaluation
- Dispositions Evaluation
- Goals

Pedagogy	Alignment	Dispositions	Alignment
Planning for Instruction and Assessment		**Professional Commitment and Behaviors**	
A.	InTASC 7a	A.	
B.	InTASC 7b	B.	InTASC 10d
C.	InTASC 6b	C.	InTASC 9o
D.	InTASC 2c	D.	InTASC 9o
		E.	InTASC 3d
Instructional Delivery		**Professional Relationships**	
E.	InTASC 7c	F.	InTASC 10b
F.	InTASC 5d	G.	InTASC 10j
G.	InTASC 8b	**Critical Thinking and Reflective Practice**	
H.	CAEP 1.5		InTASC 9n
I.	InTASC 3d		
Assessment			
J.	CAEP 2.3		
K.	InTASC 6d		
L.	InTASC 7d		
Analysis of Teaching			
M.	CAEP 1.2		

Erica Brownstein, Carolyn Kaplan, and Xiangquan (James) Yao, (The Ohio State University)

Kristall J. Day (Ohio Dominican University)

Agenda
Candidate Preservice Assessment of Student Teaching (CPAST)

- CPAST Form
  - Who designed
  - How designed
  - Instrument description
    - Training
    - Resources

- Evidence of Validity & Reliability
  - Data collection
  - Demographic data
  - Analyses conducted
  - Results

Pre-Assessment

- Does your EPP use a unit-wide instrument to assess student teaching?

- Has the instrument has been analyzed for Validity and Reliability?

- Do you have comparison scores (with other EPPs)?

Discuss with your neighbor

- What are five (5) essential components in a formative and summative student teaching assessment form?

It takes an army to build a form...

University of Toledo
V. Stewart

Wilmington College
M. Hendricks

The Ohio State University
F. Beickelman
M. Bendixen-Noe
P. Bode
E. Brownstein
K. Day
M. Fresch
C. Kaplan
C. Warner
M. Whittington

University of Akron
W. Jewell

Cleveland State University
A. Price
A. Crell

Ohio University
C. Patterson

University of Dayton
C. Bowman

Wright State University
T. Kahrig

Kent State University
J. Arhar
S. Turner

Bowling Green State University
D. Gallagher

Wittenberg University
S. Brannan
T. Whitlock

How was the CPAST Form developed?

Year 3

- Faculty/staff from 10 EPPs revised form

- Implemented with 23 EPPs and 1203 candidates

- Validity & Reliability analyses conducted

Year 2

- Faculty/staff from 8 EPPs revised form

- Piloted with candidates from 8 EPPs and 322 candidates

- Validity & reliability analyses conducted

Year 1

- Initial form developed and implemented at one EPP

* Based on (Ball & Forzani, 2010; Council of Chief State School Officers, 2011; Danielson, 2011; Gargani & Strong, 2014; Marzano, Pickering, & Pollock, 2001; SCALE, n.d.a, among others)

What is the CPAST Form?

Candidate Preservice Assessment of Student Teaching (CPAST) Form

- A *formative* and *summative assessment* during *student teaching*, aligned to *CAEP and InTASC Standards.*

- The rubric has *two subscales:*
  - (1) Pedagogy and
  - (2) Dispositions

- Developed over three years

- Collaboration with 26 institutions

Resources

- Required CPAST Form Training for Supervisors
  - Required 80% score on supervisor training quizzes
  - Initial training ~90 minutes (self-paced)
  - Refresher training ~30 minutes (self-paced)

- Optional Cooperating Teacher/Student Teacher Training
  - ~20 minutes (self-paced)

- "Look Fors" Document
  - Elaborates on qualities and behaviors for a given level of performance, describing
    - Sources of Evidence
    - Possible Evidence

What data were collected?

Data
- Midterm & Final CPAST Consensus Score
    - Combined score from Supervisor, Candidate, & Student Teacher
- edTPA scores
- State test scores
- Demographics

All IHEs

IRR IHEs
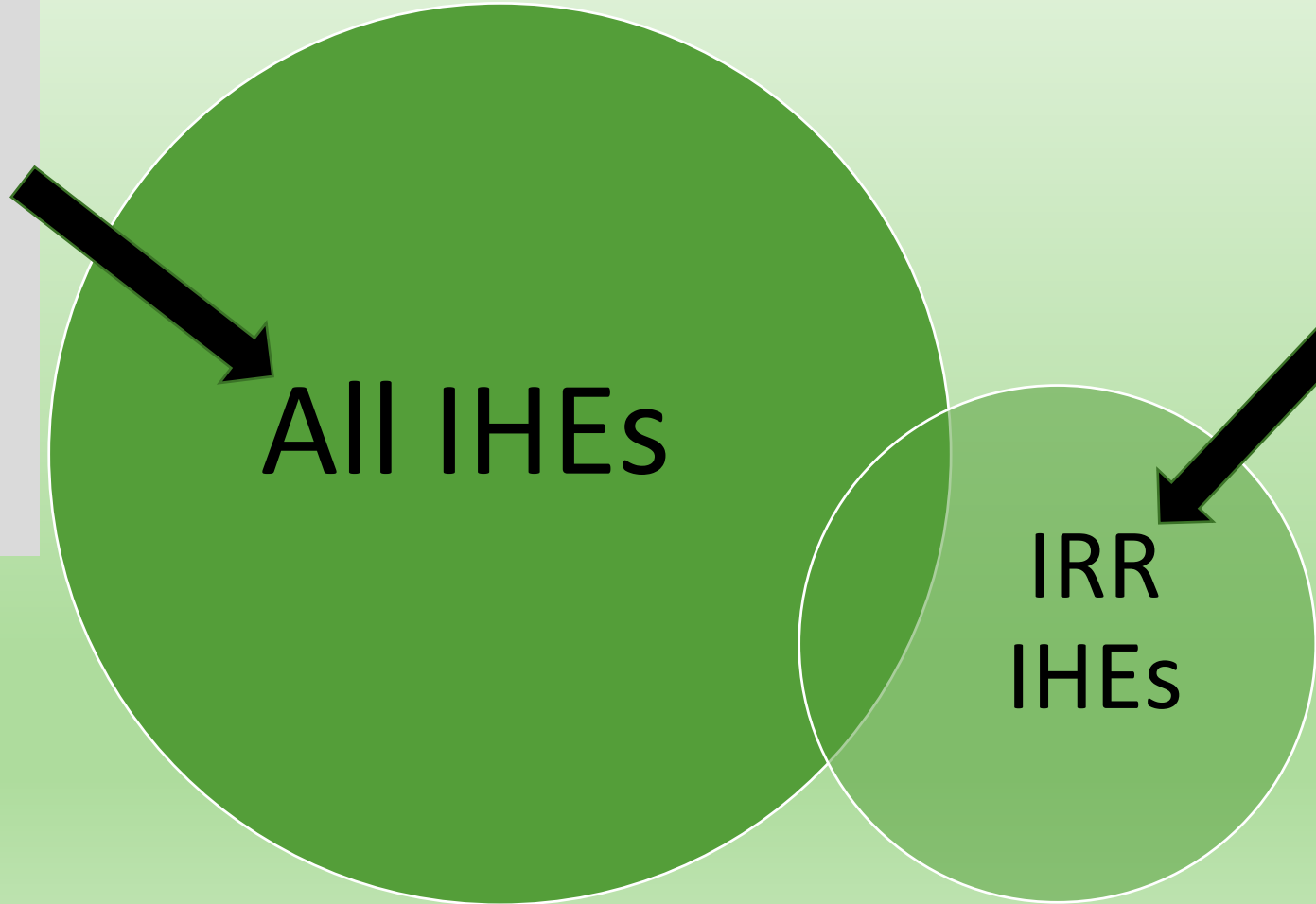
Data plus
- three observations
- IRR Supervisor final evaluation

Demographic Data of Participating Institutions

Institution Type	# of Institutions	# of Student Teachers
Urban - Public	6	481
Suburban - Public	2	359
Rural - Public	0	0
Urban - Private	5	208
Suburban - Private	6	85
Rural - Private	4	70
Total	23	1203

*Note - Sample includes representation from a variety of licensure areas, placement types, race, grade level, and gender.

Detailed demographic data available upon request.

Types of Validity and Reliability Analysis Conducted

Validity

- Content Validity

- Construct Validity

- Concurrent Validity

Reliability

- Test-retest Reliability

- Internal Consistency Reliability

- Interrater Reliability

- Addressed areas in the CAEP Evaluation Framework for Assessments

Content Validity

- Method: Content Validity Ratio

- Participants: Three Experts
  - K-12 Teacher
  - EPP Faculty
  - Psychometrician

- Measurement & Results
  - Clarity = average CVR of 0.94
  - Importance = all items reached a value of 1
  - Representativeness = average CVR of 0.94

- The results indicate CPAST has *a high content validity* (Wilson, Pan & Schumsky, 2012)

Construct Validity: CFA Results

- **Model fit indexes** RMSEA (0.048), CFI (0.980) and TLI (0.978) indicated that the **hypothesized two-factor model fit the data reasonably well**

- Factor loadings range from 0.676 to 0.841, all at .001 significance level, indicating that **all the items are moderately or strongly associated with their corresponding latent factors**.

- **Pedagogy and Dispositions scales were correlated** (r=0.873), which is supported by the literature (Kuzborska, 2011)

Concurrent Validity

- Concurrent Validity examined by correlating CPAST scores with
    - state standardized professional knowledge and skills test scores and
    - edTPA scores

- *No significant* correlation between
    - CPAST Form and state professional knowledge and skills test
    - CPAST Form Disposition Rows and edTPA

- *Significant* correlation between
    - CPAST Form Pedagogy Rows and edTPA

Internal Consistency Reliability

- Cronbach Alpha coefficient

- Results
  - 0.907 for the Pedagogy subscale
  - 0.831 for the Dispositions subscale
  - 0.929 for the total scale

- Subscales and the total scale display *good internal consistency*

*Note: A commonly accepted rule of thumb for describing internal consistency is as follows : α ≥ 0.9 Excellent, 0.9 > α ≥ 0.8 Good, 0.8 > α ≥ 0.7 Acceptable, 0.7 > α ≥ 0.6 Questionable, 0.6 > α ≥ 0.5 Poor, 0.5 > α Unacceptable

Test-Retest

- Test-retest reliability examined by correlating Pedagogy and Disposition scores in the
    - midterm and
    - final term under the CFA model.

- Pedagogy & Disposition *significantly correlated* at midterm and final term
    - (Pedagogy =.752; Disposition =.845)

Inter-rater Reliability

- Adjacent agreement and Kappa-n statistics
  - Same methods used by SCALE for edTPA analyses

- Adjacent agreement
  - Same or within one

- Kappa-n
  - Accounts for chance agreements

- Exceeds CAEP requirements
  - Adjacent agreement = 98%
  - Average Kappa-n = 0.97

*Note: CAEP requires assessment to have a interrater reliability greater than 0.8.

Item	Agreement Rate	Kappa-N
	100%	1.00
	100%	1.00
	100%	1.00
	100%	1.00
	100%	1.00
	100%	1.00
	100%	1.00
	100%	1.00
	96.9%	0.96
	100%	1.00
	100%	1.00
	100%	1.00
	100%	1.00
	87.5%	0.83
	87.5%	0.85
	90.6%	0.86
	100%	1.00
	96.9%	0.96
	96.9%	0.96
	96.9%	0.96
	96.9%	0.96

Would YOU like to join our team?

What questions do you have?

Thank you for your time!

- If you have any questions AT ANY TIME, feel free to contact:

Name	Email	Phone	Topic
Erica Brownstein	Brownstein.2@osu.edu	(614) 292-1414	"Big Picture" project questions, rubric questions
Carolyn Kaplan	Kaplan.169@osu.edu	(614) 292-5044	Training, Timeline, IRR Participants
Xiangquan (James) Yao	Yao.298@osu.edu	(614) 292-5044	Data Collection

Also, see us at Spring 2017 CAEPCon March 24 at 11 am
And Fall 2017 CAEPCon September TBD.