USCOTS 2013
Breakout session 3C

Preparing to Teach K-12 Statistics: Assessing Teachers' Readiness

May 18, 2013

Our Presentation Team

Supported in part by CAUSE (under NSF DUE #0618790)

Leigh M. Harrell-Williams
Georgia State University

M. Alejandra Sorto
Texas State University

Rebecca L. Pierce
Ball State University

Lawrence M. Lesser
The University of Texas at El Paso

Randall Groth
Salisbury University

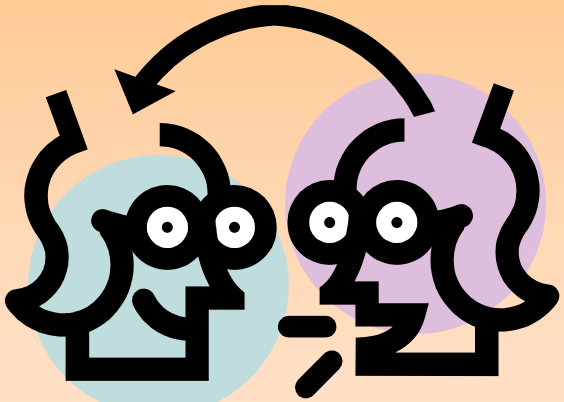
Teri J. Murphy
Northern Kentucky University

Skit Reflection Questions

- What does it mean to “be ready to teach”?
- What does a teacher really need to know to “be ready”?
  - Content knowledge
  - Pedagogical knowledge
  - Mix of both
  - Other knowledge
- How can we measure teacher readiness?

RULES OF THIS BREAKOUT SESSION

Please Speak Up



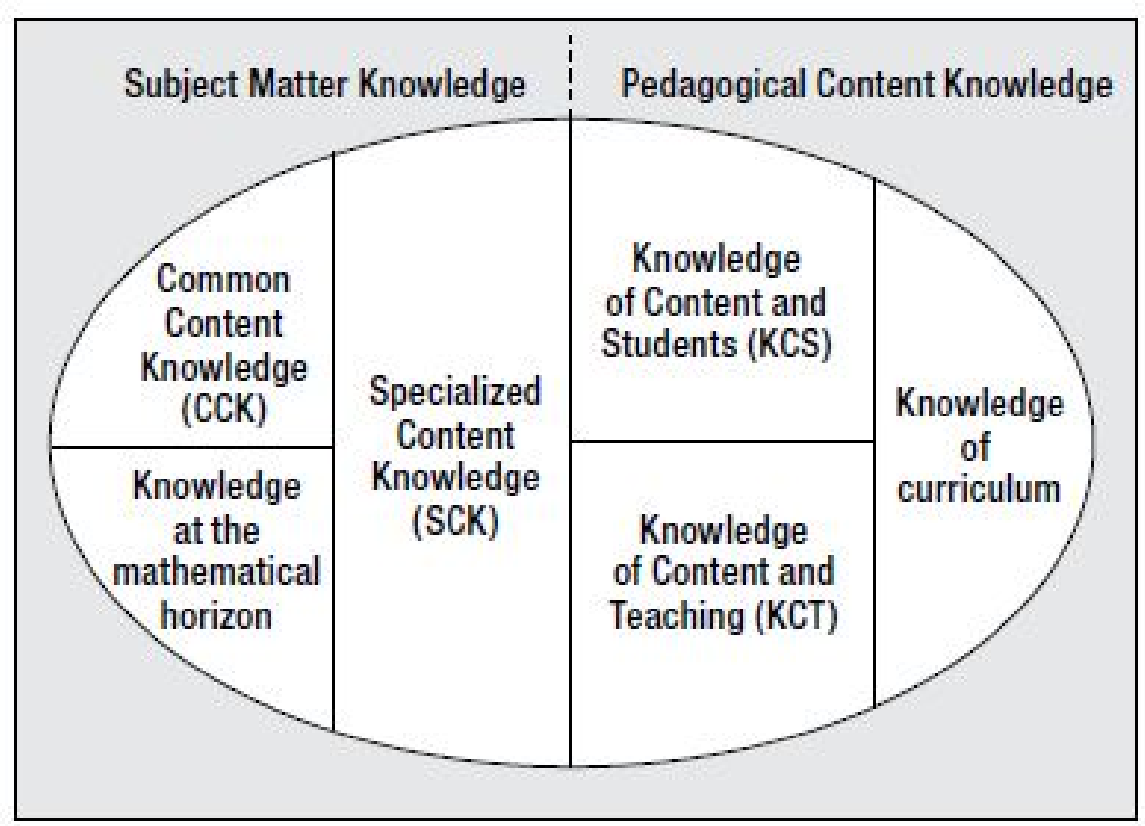
Share with Us All!

Some Assessments/Instruments

- Content knowledge exams
  - CAOS, SCI, LMT
- Classroom assessments
  - Japanese Lesson Study
  - Writing prompts
- Assessing attitude/beliefs/efficacy
  - Efficacy: SETS, SELS, CSSE
  - Attitudes: SATS, ATS
  - Beliefs & Practices: STI
  - Anxiety: SAS, STARS, SAM, SAI

TO MAKE IT BIG IN STATISTICS EDUCATION, YOU MUST USE ACRONYMS!!! THERE WILL BE A TEST ON THESE AT THE CLOSING SESSION.

What do we mean by "Content Knowledge"?



Hill, H.C., Ball, D.L., & Schilling, S.G. (2008). p.377

Content Knowledge Exams: Statistics Concept Inventory (SCI)

- Teri Reed-Rhoads, Robert A. Terry, Andrea D. Stone, Kirk Allen, Teri J. Murphy
- Items drafted 2002
- Designed for topics in an engineering statistics course
- 4 Areas: Descriptive, Probability, Inferential, Graphical

- http://cihub.org/resources/statistics

Content Knowledge Exams: Statistics Concept Inventory (SCI)

1. The following are temperatures for a week in August: 94, 93, 98, 101, 98, 96, and 93. By how much could the highest temperature increase without changing the median?

   a) Increase by 8°
   b) Increase by 2°
   c) It can increase by any amount (*correct*)
   d) It cannot increase without changing the median

2. A researcher performs a t-test to test the following hypotheses:

   $H_0 : \mu \leq \mu_0$

   $H_1 : \mu > \mu_0$

   He rejects the null hypothesis and reports a p-value of 0.10. Which of the following must be correct?

   a) The test statistic fell within the rejection region at the significance level
   b) The power of the test statistic used was 90%
   c) Assuming the null is true, there is a 10% possibility that the observed value is due to chance (*correct*)
   d) The probability that the null hypothesis is not true is 0.10

Comprehensive Assessment of Outcomes in Statistics (CAOS)

- Bob delMas, Joan Garfield, Ann Ooms, Beth Chance (2007)
- Learning outcomes associated with an algebra-based introductory statistics course
- Cronbach's alpha = .82
- All items on the CAOS test were written to require students to think and reason, not just compute, use formulas, or recall definitions

Comprehensive Assessment of Outcomes in Statistics (CAOS): Items from Tests of Significance topic

7. A newspaper article claims that the average age for people who receive food stamps is 40 years. You believe that the average age is less than that. You take a random sample of 100 people who receive food stamps, and find their average age to be 39.2 years. You find that this is significantly lower than the age of 40 stated in the article ($p < .05$). What would be an appropriate interpretation of this result?

   a) The statistically significant result indicates that the majority of people who receive food stamps is younger than 40.
   b) Although the result is statistically significant, the difference in age is not of practical importance.
   c) An error must have been made. This difference is too small to be statistically significant.

Comprehensive Assessment of Outcomes in Statistics (CAOS): Items from Tests of Significance topic

9. A researcher conducts an experiment on human memory and recruits 15 people to participate in her study. She performs the experiment and analyzes the results. She obtains a *p*-value of .17. Which of the following is a reasonable interpretation of her results?

   a) This proves that her experimental treatment has no effect on memory.
   b) There could be a treatment effect, but the sample size was too small to detect it.
   c) She should reject the null hypothesis.
   d) There is evidence of a small effect on memory by her experimental treatment.

Comprehensive Assessment of Outcomes in Statistics (CAOS)

- Each of the two previously presented was answered correctly by 60% of the students (N = 4036).
- The sample consists of students who were undergraduate college students enrolled in a non-Calculus based introductory statistics course at a 2-year College, 4-year College, or University in the United States of America between fall 2005 and fall 2012.

Content Knowledge Exams: Learning Mathematics for Teaching (LMT)

- Developed by Learning Mathematics for Teaching Project at Univ. of Michigan
  - Heather Hill, Stephen Schilling, Deborah Ball
- Data, Probability, and Statistics module (grades 4-8): piloted 2008
  - 31 items: 18 probability items, 7 data items, and 6 statistics items.

Content Knowledge Exams: Learning Mathematics for Teaching (LMT)

12. Mrs. Jackson is getting ready for the state assessment, and is planning mini-lessons for students focused on particular difficulties that they are having with adding columns of numbers. To target her instruction more effectively, she wants to work with groups of students who are making the same kind of error, so she looks at a recent quiz to see what they tend to do. She sees the following three student mistakes:

I)
$$\begin{array}{r} {\scriptstyle 1} \\ 38 \\ 49 \\ +\,65 \\ \hline 142 \end{array}$$

II)
$$\begin{array}{r} {\scriptstyle 1} \\ 45 \\ 37 \\ +\,29 \\ \hline 101 \end{array}$$

III)
$$\begin{array}{r} {\scriptstyle 1} \\ 32 \\ 14 \\ +\,19 \\ \hline 64 \end{array}$$

Which have the same kind of error? (Mark ONE answer.)

a) I and II

b) I and III

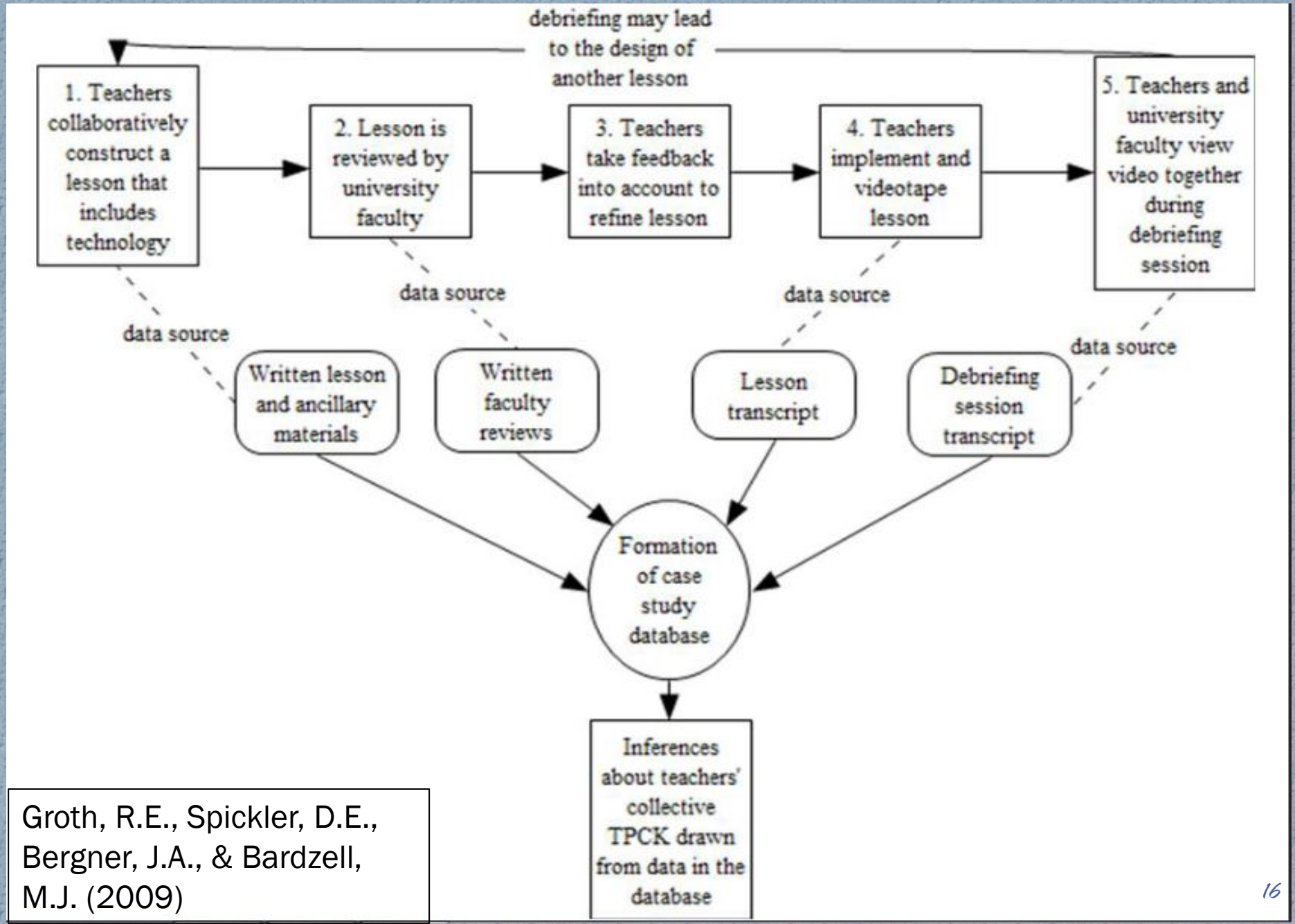
c) II and III

d) I, II, and III

Classroom Assessments: Japanese Lesson Study

- A Japanese professional development activity that has its roots in the early 1900s and has widespread use in the Japanese elementary and middle schools

- A cyclic process with four stages: study, plan, teach, and reflect



Groth, R.E., Spickler, D.E., Bergner, J.A., & Bardzell, M.J. (2009)

Classroom Assessments: Writing Prompts

SKT: Statistical Knowledge for Teaching

Sample writing prompt	Article	Reading comprehension question types	Relevant SKT components
In your own words, explain how snap cubes can be used to determine the arithmetic mean of a set of quantitative data.	Franklin and Mewborn (2008)	Literal comprehension	Knowledge of content and teaching
How are hat plots similar to box-and-whisker plots? How are they different?	Watson, Fitzallen, Wilson, and Creed (2008)	Reorganization	Specialized content knowledge
On p. 438, the authors commented in regard to item 3, "This type of item assesses students' conceptual understanding of mean." Explain what the authors may mean by "conceptual understanding." How is it different from other types of understanding?	Zawojewski and Shaughnessy (2000)	Inference	Curriculum knowledge

Groth, R.E. (2012)

Sample writing prompt	Article	Reading comprehension question types	Relevant SKT components
Why did Eric, Paloma, and Kenji each have different estimates for the number of fish in the population of Lake Amanda? How much variability in student estimates do you think you would have if you had 25 students in your class? Why?	Morita (1999)	Reorganization, prediction	Common content knowledge
On p. 417, the author claimed, "They (the students) have formulated on their own this fundamental idea in statistical inference: larger samples tend to yield less sampling variability and therefore more accuracy." Do you agree with this claim? Why or why not? What evidence is provided in the article to support the claim?	Morita (1999)	Evaluation	Horizon knowledge, knowledge of content and students
In your own words, describe the different methods students used in combining their individual samples during the capture-recapture activity. Which method do you find the most appealing? Why?	Morita (1999)	Personal response	Knowledge of content and students

Groth, R.E. (2012)

Measuring *Efficacy*, Attitudes & Beliefs

- Current Statistics Self-Efficacy (CSSE) - 2003
- Self-Efficacy to Learn Statistics (SELS) – 2003

- Self-Efficacy to Teach Statistics (SETS)
  – 2008, 2012

Measuring Efficacy: Current Statistics Self-Efficacy (CSSE) Self-Efficacy to Learn Statistics (SELS)

- Sara Finney & Gregory Schraw (2003)
- 6 point scale
- 14 items
- Focus: topics from AP/College-level Intro Stat

Measuring Efficacy: SELS & CSSE

Please rate your confidence in **learning** the skills necessary while you're in this class to successfully complete the following tasks. The item scale has 6 possible responses: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence. For each task, please mark the one response that represents your confidence in **learning** the skills necessary in this course to successfully complete the task.

SELS

	No confidence at all					Complete Confidence
1. Identify the scale of measurement for a variable.	1	2	3	4	5	6
2. Interpret the probability value (p-value) from a statistical procedure.	1	2	3	4	5	6

CSSE

Please rate your confidence in your **current ability** to successfully complete the following tasks. The item scale has 6 possible responses: (1) no confidence at all, (2) a little confidence, (3) a fair amount of confidence, (4) much confidence, (5) very much confidence, (6) complete confidence. For each task, please mark the one response that represents your confidence in your **current ability** to successfully complete the task.

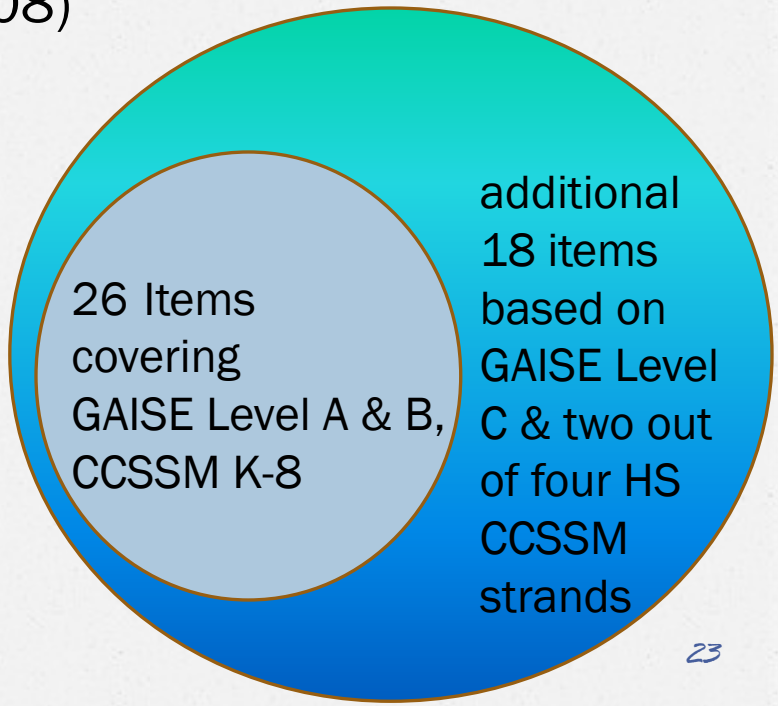
	No confidence at all					Complete Confidence
1. Identify the scale of measurement for a variable.	1	2	3	4	5	6
2. Interpret the probability value (p-value) from a statistical procedure.	1	2	3	4	5	6

Measuring Teacher Efficacy: Self-Efficacy to Teach Statistics (SETS)

- Authors: Harrell-Williams, Sorto, Pierce, Murphy, Lesser (2008)
- CAUSE Webinar (August 2012): http://www.causeweb.org/webinar/teaching/2012-08/
- Harrell-Williams, L.M., Sorto, M.A., Pierce, R.L., Murphy, T.J., & Lesser, L.M. (in press). Validation of scores from a measure of teachers' self-efficacy to teach middle grades statistics. Journal of Psychoeducational Assessment.

Measuring Teacher Efficacy: Self-Efficacy to Teach Statistics (SETS)

- Two versions:
  - Middle Grades SETS (2008)
  - High school SETS (2012)



Measuring Teacher Efficacy: Self-Efficacy to Teach Statistics (SETS)

Using a scale of {1, 2, 3, 4, 5, 6} where 1 = not at all confident, 2 = only a little confident, 3 = somewhat confident, 4 = confident, 5 = very confident, 6 = completely confident, please rate your confidence in ***teaching middle school students*** the skills necessary to complete the following tasks successfully:

	Not at all confident					Completely confident
1. Collect data to answer a posed statistical question in contexts of interest to middle school students.	1	2	3	4	5	6
2. Recognize that there will be natural variability between observations for individuals.	1	2	3	4	5	6

Measuring Efficacy, *Attitudes & Beliefs*

Attitudes

- Statistics Attitude Survey (SAS*) - 1980
- Attitude Towards Statistics (ATS) – 1985
- Students' Attitudes toward Statistics (STATS) - 1992
- Survey of Attitudes Towards Statistics (SATS) - 1992, 2003

Beliefs & Practices

- Statistics Teaching Inventory (STI) – Piloted 2009

Measuring Attitude: Student Attitude Towards Statistics (SATS)

- Candace Schau, Ph.D., CS Consultants, LLC
- SATS-28 (1992): Affect, Cognitive competence, Value, Difficulty
- SATS-36 (2003): Interest & Effort items added
- 7-point Likert response scale

Measuring Attitude: Student Attitude Towards Statistics (SATS)

	Strongly disagree			Neither disagree nor agree			Strongly agree
1. I plan to complete all of my statistics assignments.	1	2	3	4	5	6	7
2. I plan to work hard in my statistics course.	1	2	3	4	5	6	7
3. I will like statistics.	1	2	3	4	5	6	7
4. I will feel insecure when I have to do statistics problems.	1	2	3	4	5	6	7
5. I will have trouble understanding statistics because of how I think.	1	2	3	4	5	6	7

Measuring Teacher Practices & Beliefs: Statistics Teaching Inventory (STI)

- Joan Garfield, Bob delMas, Andy Zieffler
- Forced choice questions regarding 4 areas: teaching practice (11), assessment practice (10), teaching beliefs (10), assessment beliefs (6)
- Supplemental questions on teacher, course and institutional characteristics
- CAUSE Webinar (Nov. 2010): http://www.causeweb.org/webinar/teaching/2010-11/
- Paper in March 2012 *J. of Statistics Education*

Measuring Teacher Practices & Beliefs: Statistics Teaching Inventory (STI)

Part 1: Teaching Practice

Please rate the extent that each of the following are used by you (and/or your assistants) to teach this particular introductory statistics course. These statements apply only to in-class settings such as lectures, labs, and discussion.

	Never	Seldom	Some of the Time	Most of the Time	All of the Time
1. Teacher presentations (e.g., lectures, demonstrations, etc.) are used to help students learn statistics.					
2. Activities are used to help students learn statistics.					

Measuring Teacher Practices & Beliefs: Statistics Teaching Inventory (STI)

Part 3: Assessment Practice

Please consider all assessments that are used in this particular course to evaluate student learning. Indicate your agreement or disagreement with each of the following statements as they reflect your assessment of student learning for this particular course.

	Disagree	Agree
16. My assessments include a variety of assessment types (e.g., homework, quizzes, projects, minute papers, etc.).		
17. One use of my assessments is to reveal whether students are using statistical language properly.		

Measuring Teacher Practices & Beliefs: Statistics Teaching Inventory (STI)

Part 4: Teaching Beliefs

Please rate the extent to which you agree with each of the following statements as they reflect your beliefs (but not necessarily your actual teaching) regarding the teaching and learning of introductory statistics

	Strongly Disagree	Disagree	Agree	Strongly Agree	Undecided
26. Rules of probability should be included in an introductory statistics course.					

34. Indicate the type of data that you believe helps students learn statistics best.
- ❑ All constructed data
- ❑ Mostly constructed data
- ❑ Equal amounts of constructed data and real data
- ❑ Mostly real data
- ❑ All real data

Measuring Teacher Practices & Beliefs: Statistics Teaching Inventory (STI)

Part 5: Assessment Beliefs

Please rate the extent to which you agree with each the following statements as they reflect your beliefs about assessment (e.g., homework, quizzes, exams, projects, etc.) in a non-calculus based introductory course.

	Strongly Disagree	Disagree	Agree	Strongly Agree	Undecided
36. Traditional assessments (e.g., exams, quizzes) should be used to evaluate student learning.					

41. Indicate your belief in the purpose of student assessment.
- ❑ Only to measure student achievement
- ❑ Mostly to measure student achievement
- ❑ Equal emphasis on measuring achievement and providing feedback to improve student learning
- ❑ Mostly to provide feedback to improve student learning
- ❑ Only to provide feedback to improve student learning

Measuring Efficacy, Attitudes & Beliefs (**Focus on Anxiety**)

- Statistics Anxiety Rating Scale (STARS) - 1980
- Statistics Anxiety Inventory (SAI) - 1991
- Statistics Anxiety Scale (SAS*) - 1992
- Statistics Anxiety Measure (SAM) – 2007
- Statistical Anxiety Scale (SAS*) – 2008

Suggested Discussion Questions

- Have you used any of these assessments/instruments to measure readiness to teach statistics? If so, how?
- Could you see yourself using one that you haven't before? If not, what would keep you from doing so?

Thank you for attending our breakout session!

Contact information:

- Using the SETS instruments: Rebecca Pierce, rpierce@bsu.edu

- Lesson study and writing prompts: Randall Groth, regroth@salisbury.edu

References

- Curcio, Frances R. *A User's Guide to Japanese Lesson Study: Ideas for Improving Mathematics Teaching*. Reston, VA: National Council of Teachers of Mathematics, 2002.

- delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58.

- Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28(2)*, 161–186.

- Groth, R.E., Spickler, D.E., Bergner, J.A., & Bardzell, M.J. (2009). A qualitative approach to assessing technological pedagogical content knowledge. *Contemporary Issues in Technology and Teacher Education*, *9* (4). http://www.citejournal.org/vol9/iss4/mathematics/article1.cfm

References

- Groth, R.E. (2011). Improving teaching through lesson study debriefing. *Mathematics Teacher, 104,* 446-451.

- Groth, R.E. (2012). The role of writing prompts in a statistical knowledge for teaching course. *Mathematics Teacher Educator, 1*, 23-40.

- Harrell-Williams, L., Sorto, M.A., Pierce, R. L., Lesser, L., & Murphy, T.J. (August 2012). Measuring confidence to teach statistics to middle and high school grades: The development and validation of the SETS instruments. Webinar for Consortium for the Advancement of Undergraduate Statistics Education (CAUSE). http://www.causeweb.org/webinar/teaching/2012-08/

- Harrell-Williams, L.M., Sorto, M.A., Pierce, R.L., Lesser, L.M., & Murphy, T.J. (2013). Validation of scores from a measure of teachers' self-efficacy to teach middle grades statistics. *Journal of Psychoeducational Assessment.*

References

- Hill, H.C., Ball, D.L., & Schilling, S.G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*, 372-200.

- Kaplan, J.J., Cervello, K., & Corcoran, E. (2009). Lesson study as a tool for professional development: A case of undergraduate calculus. Proceedings of the 2009 Conference on Research in Undergraduate Mathematics Education. http://mathed.asu.edu/crume2009/Kaplan_LONG.pdf

- Learning Mathematics for Teaching (2008). Mathematical Knowledge for Teaching (MKT) Measures. Probability, Data, and Statistics. Ann Arbor, MI: University of Michigan.

References

- Otten, S., Park, J., Mosier, A. & Kaplan, J.J. (2009) Lesson study as a tool for research: A case of undergraduate calculus. Proceedings of the 2009 Conference on Research in Undergraduate Mathematics Education. Available online: http://mathed.asu.edu/crume2009/Otten_LONG.pdf

- Schau, C., Stevens, J., Dauphinee, T., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes toward Statistics. *Educational and Psychological Measurement, 55*(5), 868-875.