Deni Basaraba, Ph.D., Yetunde Zannou, Ph.D., Dawn Woods, Paul Yovanoff, Ph.D., & Leanne Ketterlin-Geller, Ph.D.!

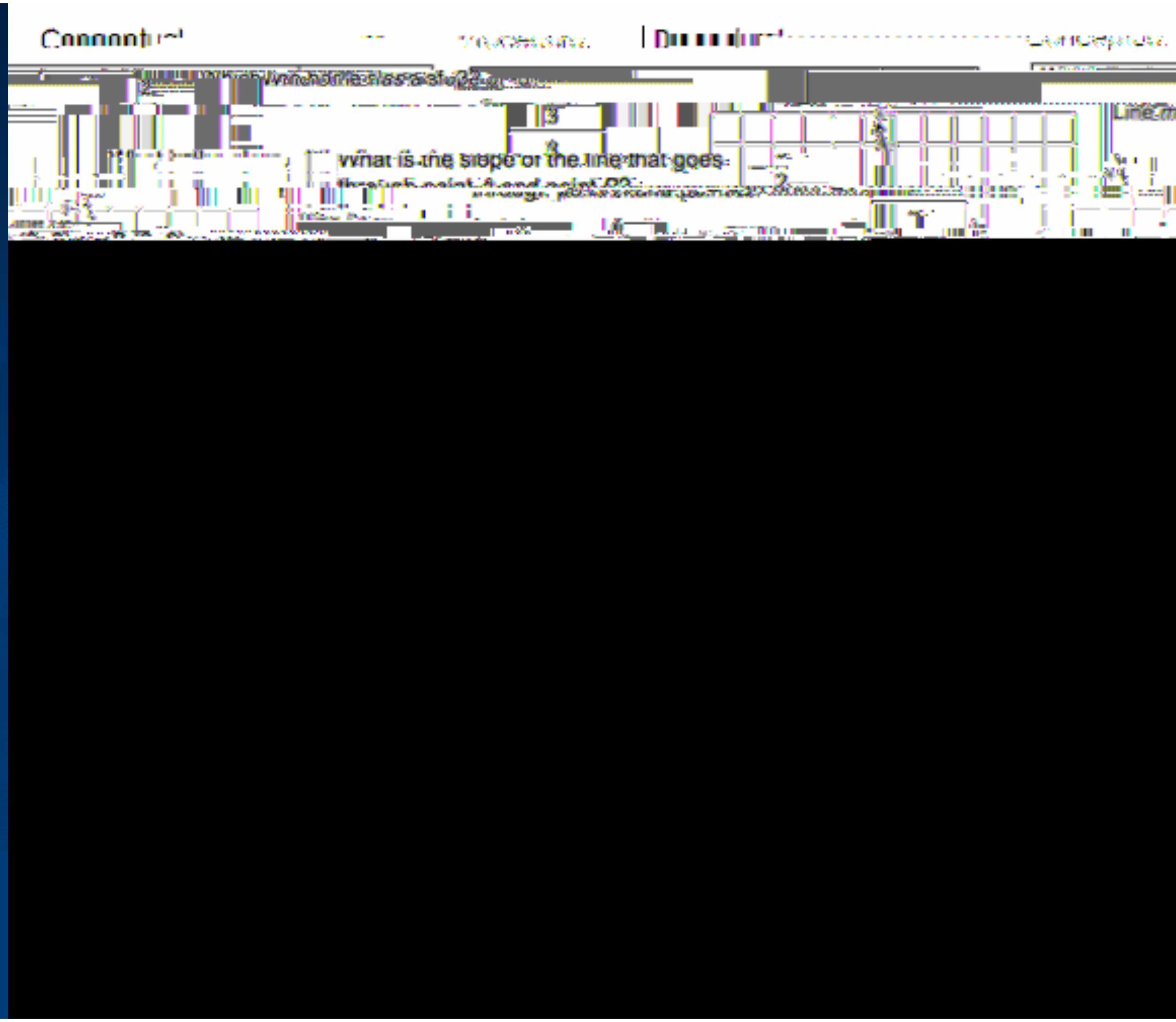
- Describe the assessment context for this study!
- Provide theoretical rationale for verbal protocols (aka “think-alouds”)!
- Explore the relation between metacognition and performance on mathematics items!
- Describe methods for coding students# responses collected during the verbal protocols!
- Present results of analyses!
- Explore directions for future research!

- Developed as one component of a comprehensive, response-to-intervention initiative designed to increase the preparedness of students to meet standards and pass assessments in algebra in Grades 2-8 (Texas Algebra Ready)!
- Components of the initiative include:!
  - Online professional development academies that focus on core and supplemental mathematics instruction!
  - Assessments (Universal Screener and Diagnostic)!
  - Sample intervention lessons!

- Learning progressions represent hypotheses about the development of students' understanding about a target construct (e.g., algebra readiness) (Duschl, Schweingruber, & Shouse, 2007)
  - Include descriptions of successively more sophisticated ways of thinking about the target construct students engage in as they learn over time (Corcoran, Mosher, & Rogat, 2007)

- For the ESTAR Universal Screener, this development is represented via two components:!
  - Knowledge representation (foundational, bridging, and target content knowledge as defined by state content standards)!
  - Levels of cognitive engagement (levels of cognitive processing with which students are expected to engage with the content)!
- Although the content and levels of cognitive engagement increase in their sophistication and complexity, items were also written to three different levels of relative difficulty (e.g., an item could assess foundational content knowledge, target strategic competence, and be considered a relatively “easy” item)!

- According to the National Research Council (2001) and others, mathematical proficiency requires the following:!
  - Conceptual understanding (CU): Understanding of mathematical concepts and operations, the relations them, and/or a procedure works !
  - Procedural fluency (PF): Ability to follow a sequence of certain, defined actions flexibly, efficiently, and accurately!
  - Strategic competence (SC): Ability to formulate, represent, and solve


What is the slope of the line that goes

- Process of having students "think-aloud" while completing a task!
- Students are asked to say what they are looking at, thinking, and doing (including strategies they are using) while completing a task!

all

- Goal: To see first-hand the of task completion, rather than just the final product!

- Can be useful during the test development process because having students "think-aloud" while solving problems can provide information about!
  - Cognitive processes students were engaged in while solving the problem (Ericsson & Simon, 1993)!
  - Students# understanding (or misunderstanding) of the content and constructs being assessed (Almond et al., 2009)!
  - Whether items are of comparable difficulty or understood similarly for

- Researchers (Deseote, 2009; Deseote, Roeyers, & Buysse, 2001; Jacobse & Harskamp, 2012) have used verbal protocols to explore the relationship between procedural, predictive, and planning metacognition and students# performance on mathematics problems!
  - Predictive: Congruence between students# prediction of their selecting the correct response and whether they selected the correct response!
  - Planning: Ability to articulate what you would need to do to solve the problem (identification and application of problem-solving steps)!
  - Procedural: Accurate understanding of the procedures and strategies needed to solve the problem using the information given!
- Findings include (a) moderate correlations between metacognition and mathematics performance and (b) that indicators of metacognition can explain some of the observed variability in student math performance!

Participants: 10 4th grade students with varying mathematics ability!

Procedures: !

- Asked students to solve 10 multiple-choice mathematics items!
- Had students respond to 10 retrospective think-aloud questions after solving each problem (coded 5)!
- Developed a rubric for each item outlining the expected components of responses of students demonstrating
  and understanding of the assessed content!
- Had 2 independent reviewers code students# responses targeting predictive, planning, and procedural metacognition for each item!

- Procedural Metacognition!
  - Defined as one#s knowledge of the methods or strategies needed to achieve one#s goals (i.e., solving the problem), understanding those strategies work, and how they can be applied to solve the problem (Deseote et al., 2001; Montague, 1992)!
  - During the retrospective think-aloud, we asked students:!
    - Q5: What strategies and steps did you take to solve the problem?!
    - Q6: Does your answer for this problem make sense? Why?!

- Although not directly related to metacognition, there is reason to believe that students# understanding of the content might mediate his/her ability to identify or develop a plan to solve the problem!
    - During the retrospective think aloud, we asked students:!
        - Q2: What do you know about [content assessed by the problem]?!
    - We identified the critical information needed to solve the problem and examined the extent to which students# responses to these questions included that information!

VP Question	Exemplary	Proficient	Developing	Emerging

- Due to limited number of cases, 3 single, linear regression models were conducted with the total number of items correct as the outcome and a type of metacognition as a predictor!

Predictor	*F*	*p*	*b*	*t*	*p*	Adjusted R^2

- Our indicators of planning metacognition (Q3 & Q4) had the strongest relation with students# performance on the multiple-choice mathematics items!
- Contrary to previous research (Deseote, 2009; Jacobse & Harskamp, 2012) our indicator of p.2(o) -0.0.0.0.0.0p0.2 (f) T 19 0 0 19 0 -15.2 T(290.2 (f) T 2 (a) h T

Verbal Protocol Data!

- Small sample, only one grade level!
- Lower levels of inter-rater agreement for coding some of the questions!
- Likely a relation between students# content knowledge and their ability to plan to solve a problem and follow procedures to solve that problem that weren#t accounted for in our analyses!
- Unable to fully examine relation between predictive metacognition and student performance!
  - Explore relation between prediction, evaluation, and/or indicators of persistence!

Verbal Protocol Data!

- Examine rationales for G4 ratings to see if there are specific reasons why the inter-rater agreement was low for specific items; revise rubric if necessary!
- Analyze G2 and G3 data!
- Consider revising indicator of predictive metacognition to include evaluation (confidence in selection of correct response after solving the problem), which is consistent with prior research (Deseote, 2009; Jacobse & Harskamp, 2012) and students# rating of their perceived difficulty of the item!
- Consider logistic regression as alternative analysis to examine the relation of these predictors to students# performance on each item!

- Participants: 2,548 students who responded to between 22-25 multiple-choice items written for the ESTAR US item bank (206 Grade 4 items)!
- Procedures:!
  - IRT item difficulties estimated with a bifactor model in Testfact !
    - Bifactor model includes a general, underlying latent factor (algebra-readiness) and four domain-specific factors (CU, PF, SC, AR) – 1 model for each knowledge

- Chi-square difference tests comparing the model fit of the unidimensional, single-factor and bi-factor model revealed that the bifactor model fit the data better for all three knowledge representations:!

Knowledge Representation	Bifactor ^{2}	Single Factor ^{2}	^{2} Difference	Difference df	*p*

Knowledge Representations	Percent of Variance Explained					Additional variance explained by specific factors

- Kruskal-Wallis tests were conducted to examine the equality of population medians (and means) among groups of items!
  - Items are rank-ordered with respect to item difficulties (low to high)!
  - Item types (CU, PF, SC, and AR) are compared with respect to rank!
  - If item types have different median or mean values, then item types are different with respect to item difficulties!
- One-way ANOVAs were conducted to test for statistically significant differences across the means of item difficulty types!
- Results indicated no significant differences across the item types!

- Comparisons of the single factor and bifactor model revealed that the bifactor fit the data better and that a not insignificant amount of

Model comparisons!

- Examine the percent of variance within each item type (CU, PF, SC, and

Almond, et al. (2009).

Dover, NH: Measured Progress.!

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. 189-225.!

Corcoran, T., Mosher, F., & Rogat, A. (2010). (CPRE Research Report #RR-63). Consortium for Policy Research in Education!

Deseote, A. (2009). Metacognitive prediction and evaluation skills and mathematical learning in third-grade students. 435-446.!

Deseote, A., Roeyers, H., & Buysse, A. (2001). Metacognition and mathematical problem solving in grade 3. 435-448.!

Ercikan et al. (2010). Application of think aloud protocols for examining and confirming sources of differential item functioning identified by expert reviews. (2), 24-35.!

Ericsson, K. A., & Simon, H. A. (1993). Cambridge, MA: MIT Press.!

Jacobse, A. E., & Harskamp, E. G. (2012). Toward efficient measurement of metacognition in mathematical problem solving. 133-149.!

Montague, M. (1992). The effects of cognitive and metacognitive strategy instruction on mathematical problem solving of middle school students with learning disabilities.