Developing an assessment for concepts in introductory statistics and data science

P Burckhardt, P W Elliott, C Evans, S Hyun, K Lin, A Luby, C P Makris, M Meyer, J Orellana, R Yurko, G Weinberg, J Wieczorek, R Nugent & A Reinhart
Dept. of Statistics & Data Science, Center for the Neural Basis of Cognition, Carnegie Mellon University, and Dept. of Mathematics and Statistics, Colby College

Overview
- We are developing an assessment of introductory statistics concepts
- Think-aloud interviews with students helped us uncover new misconceptions and improve assessment questions
- We administered the assessment to = 200 students at two different institutions
- Goal: Provide ways to measure student learning, so we can conduct more pedagogical research

Assessing Learning in Intro Stats
- To improve teaching, need to assess what the students are learning
- Must avoid ambiguous questions and confirm that questions actually measure student learning, not just test-taking skills
- Used think-alouds: have students think aloud while answering draft assessment questions (see Adams and Wieman 2011, Burckhardt et al. 2017)
- Think-alouds elicit misconceptions and misreadings, and help us revise and write new questions

Think-Aloud Results
- Conducted 36 interviews, each = 1 hour long, in rounds timed to topics introduced in 36-200
- Interviews tested roughly 50 draft questions
- Used student feedback to revise questions, then re-tested in later think-aloud rounds

Data Collection
- After think-alouds, built revised assessment
- Students also asked to rate their confidence in each answer
- Used ISLE to administer the assessment to 95 students in 36-202 in Fall 2018; results presented here
- Administered paper version to 117 introductory students at Colby College, in 5 course sections taught by 3 instructors

Spotting False Confidence
Misconceptions can be identified by finding questions students get wrong while confident (in red):

Next Steps
- Student beliefs are surprising: to build a good assessment, think-alouds are invaluable
- We recommend this process for anyone writing assessments in any field
- Will collect pre/post data in 36-200 in Spring 2019 to assess student learning
- Will survey instructors to get their input on the assessment topics and questions
- Results will guide new pedagogical experiments
- The validated test can be used to assess learning and aid redesign for new Dietrich General Education curriculum

References
Burckhardt et al. (2017), Teaching & Learning Summit
Hyun et al. (2018), eCOTS

Acknowledgments
Thanks to the Eberly Center for advice, to the Department of Statistics & Data Science for financial support, and to all the students who participated in interviews and assessments.

Assessing Learning in Intro Stats
- To improve teaching, need to assess what the students are learning
- Must avoid ambiguous questions and confirm that questions actually measure student learning, not just test-taking skills
- Used think-alouds: have students think aloud while answering draft assessment questions (see Adams and Wieman 2011, Burckhardt et al. 2017)
- Think-alouds elicit misconceptions and misreadings, and help us revise and write new questions

Think-Aloud Results
- Conducted 36 interviews, each = 1 hour long, in rounds timed to topics introduced in 36-200
- Interviews tested roughly 50 draft questions
- Used student feedback to revise questions, then re-tested in later think-aloud rounds

Data Collection
- After think-alouds, built revised assessment
- Students also asked to rate their confidence in each answer
- Used ISLE to administer the assessment to 95 students in 36-202 in Fall 2018; results presented here
- Administered paper version to 117 introductory students at Colby College, in 5 course sections taught by 3 instructors

Spotting False Confidence
Misconceptions can be identified by finding questions students get wrong while confident (in red):

Next Steps
- Student beliefs are surprising: to build a good assessment, think-alouds are invaluable
- We recommend this process for anyone writing assessments in any field
- Will collect pre/post data in 36-200 in Spring 2019 to assess student learning
- Will survey instructors to get their input on the assessment topics and questions
- Results will guide new pedagogical experiments
- The validated test can be used to assess learning and aid redesign for new Dietrich General Education curriculum

References
Burckhardt et al. (2017), Teaching & Learning Summit
Hyun et al. (2018), eCOTS

Acknowledgments
Thanks to the Eberly Center for advice, to the Department of Statistics & Data Science for financial support, and to all the students who participated in interviews and assessments.