Writing practical pre- and post-tests for concepts in introductory courses


Overview
- We are developing an assessment of introductory statistics concepts
- Think-aloud interviews with students helped us uncover new misconceptions and improve assessment questions
- Pre- and post-tests can help you improve your courses by measuring student learning
- Interviews and assessments revealed new misconceptions about correlation and causation

Assessing Learning in Intro Stats
- To improve teaching, need to assess what the students are learning
- Used think-alouds: students think aloud while answering draft assessment questions (Adams and Wieman 2011, Burckhardt et al. 2017)
- Think-alouds elicit misconceptions and misreadings, and help us revise and write new questions
- Think-alouds allow us to make effective pre- and post-tests of complex misconceptions, guiding our teaching

Think-Aloud Results
- Conducted 42 interviews, each 1 hour long, in rounds timed to topics introduced in 36-200 (from Spring 2018 through Summer 2019)
- Tested roughly 50 draft questions in the interviews
- Used student feedback to revise questions, then re-tested in later think-aloud rounds

Data Collection
- After think-alouds, built revised assessment and gave it to students in 36-200 "Reasoning with Data" as a pre- and post-test
- Expanded the assessment to include students in "Introduction to Statistics and Data Science" at Colby College, starting Spring 2019
- Used ISLE, an online statistics learning platform, to administer the assessment to 379 students and extract the question-level answer data

Combining Results
In post-test, most students correctly answered when correlation doesn’t imply causation, but not when correlation does imply causation:

<table>
<thead>
<tr>
<th></th>
<th>books wrong</th>
<th>books right</th>
</tr>
</thead>
<tbody>
<tr>
<td>vitamin-c wrong</td>
<td>5</td>
<td>51</td>
</tr>
<tr>
<td>vitamin-c right</td>
<td>5</td>
<td>28</td>
</tr>
</tbody>
</table>

Possible reasons to explore:
- Belief that correlation does not imply causation even when causal conclusions can be drawn
- Know the phrase "correlation does not imply causation," but can’t recognize causal language that doesn’t use keywords like “causation”
- Incomplete understanding of why randomization is useful (or of distinction between random sampling and assignment)

Causal Diagrams of the Questions

- **books**
  - Confounder
  - Number of books read
  - Nearsightedness

- **vitamin-c**
  - Confounder
  - Recovery time for cold

Causal diagrams for the books and vitamin-c questions. The diagram for books is consistent with how students think about the question, but for vitamin-c, the students do not realize the arrow from any confounder to "meditation" has been deleted by randomization. While students don’t learn causal diagrams in 36-200, their understanding of causality from think-aloud interviews appears consistent with a causal diagram that includes arrows for confounding variables, regardless of the scenario.

Next Steps
- Currently writing and testing new questions on correlation and causation
- Exploring other frequently misunderstood concepts in introductory statistics
- Expanding interviews to more undergraduate institutions
- Hope to experimentally test new ways to teach these concepts
- These same methods could work for your course too! Understanding student learning could lead you to new teaching strategies

Overheard in Think-Aloud Interviews
During the think-aloud interviews, some students gave reasoning suggesting that they are unwilling to ever infer causation:
- “Correlation does not imply causation is a universal rule” (books)
- “When can we ever say something causes something else?” (candy-test)
- “I think the word ‘causes’ is too strong... my friend who’s a stats major always tells me you can’t say this causes that—there’s always other factors” (vitamin-c)
- “Usually [you] can’t assume causation” (vitamin-c)

References
Burckhardt et al. (2017), Teaching & Learning Summit.
ISLE: http://www.stat.cmu.edu/isle/

Acknowledgments
Thanks to S. Hyun, K. Lin, and R. Yurko for conducting think-aloud interviews, and G. Weinberg for advice and support. Thanks also to the Eberly Center for advice, to the Department of Statistics & Data Science for financial support, and to all students who participated in interviews and assessments.