Statistical problem-solving cycles while solving simulation tasks during guided interviews

Department of Educational Psychology

Jonathan Brown, Robert delMas, & Andrew Zieffler
Quantitative Methods in Education, University of Minnesota

CEHD COLLEGE OF EDUCATION + HUMAN DEVELOPMENT

Background

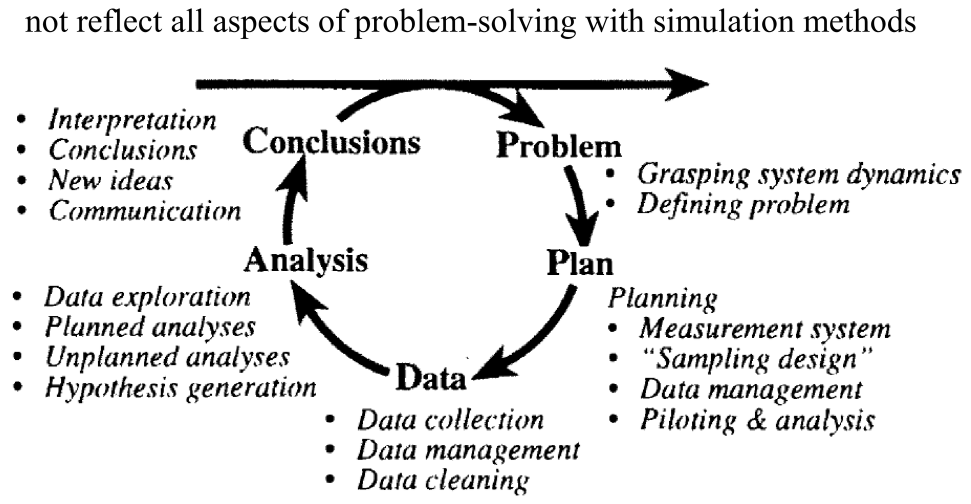
- Simulation methods (e.g., randomization test, bootstrapping) offer an alternative method of inference to, or augment, traditional parametric methods typically taught in introductory coursework (Cobb, 2007; Rossman & Chance, 2014)
- Simulation methods often employ dynamic software tools such as Fathom® (Finzer, 2002) or TinkerPlots™ (Konold & Miller, 2015)
- Little empirical evidence exists about students' reasoning while solving statistical tasks when employing simulation methods

Research Questions

- What is the nature of student statistical problem-solving using simulation following completion of an introductory course?
- How does the nature of student statistical problem-solving using simulation fit into a larger problem-solving framework?

Framework for Interpreting Results

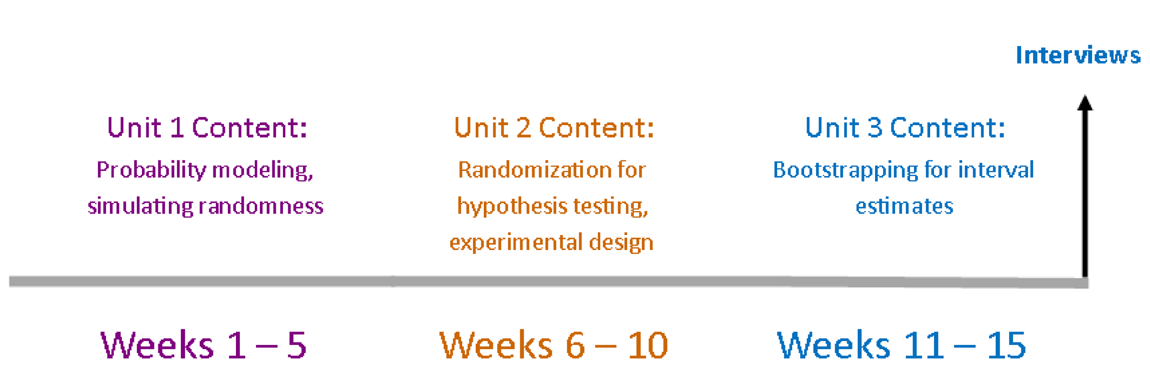
- Steps for statistical problem-solving with simulation methods have been proposed (Maxara & Biehler, 2006; Noll et al., 2016), though not empirically placed in a larger problem-solving framework
- The Problem, Plan, Data, Analysis, Conclusion (PPDAC) cycle is one model for statistical investigation (Wild & Pfannkuch, 1999), but may not reflect all aspects of problem-solving with simulation methods



Methodology

Participants, Course, and Data Collection

- Eight student volunteers from multiple sections of an undergraduate introductory simulation-based statistics course
- Change Agents for Teaching and Learning Statistics (CATALST) course
  - Curriculum from an NSF-funded three-year teaching experiment (Garfield et al., 2012)
  - Focus on student construction of probability models and employing simulation to explore statistical phenomena and conduct statistical inference and estimation
  - Use of TinkerPlots™ software throughout course
  - Emphasis on in-class activities completed in groups
  - Three course units, each five weeks in duration
- Data were collected from participant interviews at the end of the course (15 weeks), as shown:



Interviews and Tasks (problem type, one solution approach)

- Each interview consisted of the interviewer prompting and guiding participants through multiple statistical tasks, most of which were solved with a simulation method using TinkerPlots™ software
- *Task 1 (Single proportion hypothesis test, randomization test)*
  - Hypothetical computer game, randomly assigning four colors to a 5x5 grid, where a user attempts 50 times to correctly guess where the color blue will appear to earn points
  - Story about disagreement between the game's creator and their friend, regarding a supposedly imporbably high score
  - Task prompt: "What could you do to decide if the student was correct to not believe that his friend hit blue 27 out of 50 turns?"
- *Task 2 (Single proportion estimation, bootstrap confidence interval)*
  - Participants presented with results 2011 Gallup poll where 43% of those surveyed chose Barak Obama over George W. Bush as having been a better president
  - Task prompt: "What would you estimate for the percent of all U.S. adults who would have said that Barack Obama has been a better president than George W. Bush during that same period of time (from Sept. 15 and 18, 2011)?"
- *Task 3 (Two-sample difference in proportions, randomization test)*
  - Another 2011 Gallup poll, but this scenario focused on extent of public awareness of Occupy Wallstreet movement, depending on geographic location of respondent
  - Results shared with participants included percentage of those in the East/West (62) vs. South/Midwest (49) who indicated they " Very or Somewhat Closely" follow the movement
  - Task prompt: "Based on this data, how could you decide if the implication by Gallup (U.S. adults on the East and West coasts were more likely to pay attention to news about Occupy Wall Street than U.S. adults in the South or Midwest) is a legitimate claim?"
- *Task 4 (Two-sample proportion difference estimation, bootstrap confidence interval)*
  - Same scenario and data as Task 3, with a new prompt and task type
  - Task prompt: "What would you estimate for the true difference between the percent of all U.S. adults on the East or West coast and the percent of all adults in the South or Midwest who would have said they were following news about Occupy Wall Street very or somewhat closely during the same time period (Nov. 19 and 20, 2011)?"

Interview Coding and Analysis

- Each participant statement or set of related statements with the interviewer categorized into a "phase"
- Phase determination was informed by the particular behaviors exhibited by the participant relevant to the problem-solving process
- A participant was considered to be in a specific phase until one of the following occurred:
  - the interviewer asked a question that pushed the interview to a new phase
  - the participant talked their way from one phase to another, with limited interviewer involvement

Results

Observed Problem-solving Phases with Example Quotes

- Six phases resulted from the analysis, defined as follows:

P **Plan:** attempting to describe how to solve a given task prior to using TinkerPlots™

M **Model:** describing TinkerPlots™ work to set up the solution, and where models were being constructed but not yet simulated

S **Simulate:** the act of running the model, as opposed to building it; more of a marker in the problem-solving process, as participant statements during this phase type was limited

E **Evaluate:** discussion / work following simulation, including deciding to run more simulations and making sense of results

C **Conclusions:** directly answering the research question; most cases were the complete contextual extension of a final Evaluate phase

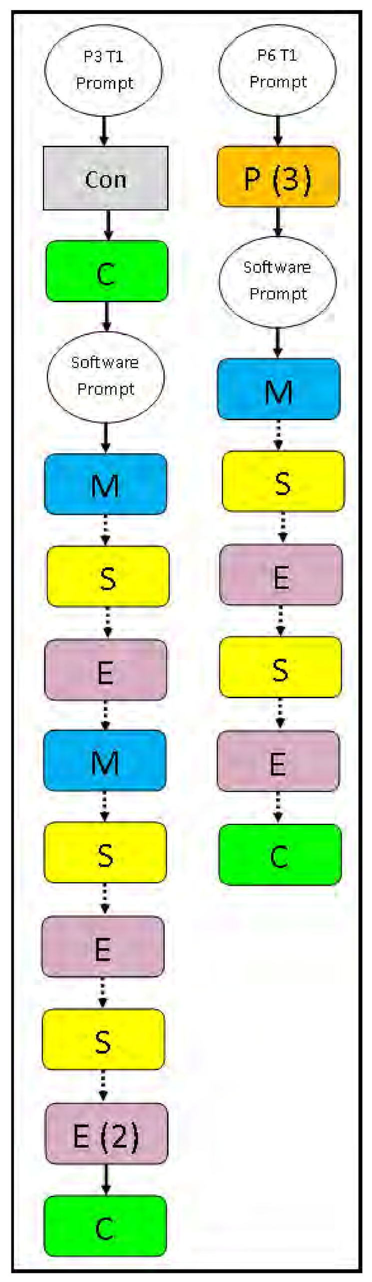
Con **Context:** questioning or attempting to understand a task's context; limited instances of this

Example Quotes

- Task 1, Participant 5, Plan phase:
  "…you could create some sort of test where you use his program over and over many times, and see the average number of how many blue squares at a time. You could see … if 27 was like way out…"
- Task 2, Participant 2, Model phase:
  "…I have a sampler, and I would change it to a spinner…. I would show the percent on there, and better would be 43%. …I would change the draw to one, and the repeat to 1004."

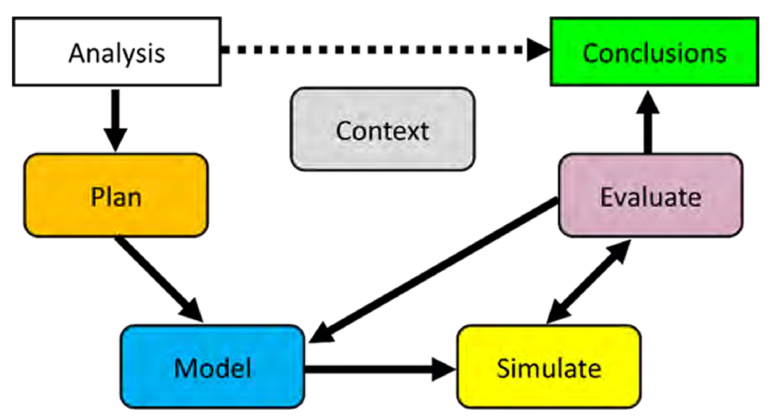
Example Trajectories of Phases

- Task 1 (T1) trajectories for participants 3 (P3) and 6 (P6)
  - Phase change from interviewer (solid arrow) or participant (dotted)
  - Numbers in parentheses indicate how many times the interviewer prompted the participant for additional discussion within the phase



Proposed Simulation Problem-Solving Sub-Cycle

- A typical trajectory emerged, as shown, amid numerous variations
- No particular trajectory was associated with unanimous success, and in multiple cases, revisiting previous phases was needed for solving the task
- Trajectories indicated the need for a "simulation sub-cycle" linking the "Analysis" and "Conclusions" stages in the PPDAC framework from Wild & Pfannkuch (1999):



Discussion

- *Study artifacts potentially influencing trajectory/outcomes*
  - Mostly high performing students in the sample
  - Highly structured nature of guided interviews (less open-ended)
  - Interviewer effects, such as subtle prompting variations
  - Usage of TinkerPlots™ software
  - Similarity of coursework and behavior to study tasks (less novelty)
- *Role of the "Plan" phase*
  - Style of prompt mostly necessitated a Plan phase of some kind, thus difficult to assess if students would self-initiate a Plan phase or how a Plan may or may not contribute to task success
  - Evidence that planning prior to engaging in a simulation task may be worthwhile (Biehler & Prömmel, 2010)
- *Pedagogical approach*
  - Due to no clear trajectory for success or failure, success in a statistics class using simulation may occur in multiple ways, depending on several factors
- *Future Research*
  - Evaluate the sensitivity of the observed trajectories to shifts in study characteristics, including further disentangling study tasks from the course being taken and the software employed

References

- Biehler, R., & Prömmel, A. (2010). Developing students' computer supported simulation and modelling competencies by means of carefully designed working environments. In C. Reading (Ed.), Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia. Voorburg, The Netherlands: International Statistical Institute.
- Chance, B., & Rossman, A. (2006). Using simulation to teach and learn statistics. In A. Rossman & B. Chance (Eds.), Proceedings of the Seventh International Conference on the Teaching of Statistics (pp. 1-6).
- Cobb, G. W. (2007). The Introductory Statistics Course: A Ptolemaic Curriculum? Technology Innovations in Statistics Education, 1(1). ISSN 1933-4214.
- Finzer, W. (2002). Fathom Dynamic Data Software (Version 2.1). [Computer Software]. Emeryville, CA: Key Curriculum Press.
- Garfield, J., delMas, R., & Zieffler, A. (2012). Developing statistical modelers and thinkers in an introductory, tertiary-level statistics course. ZDM-The International Journal on Mathematics Education, 44(7), 883-889.
- Konold, C. & Miller, C. (2015). TinkerPlots® Version 2.3 [computer software]. Emeryville, CA: Key Curriculum Press.
- Maxara, C., & Biehler, R. (2006). Students' probabilistic simulation and modeling competence after a computer-intensive elementary course in statistics and probability. In A. Rossman & B. Chance (Eds.), Proceedings of the Seventh International Conference on the Teaching of Statistics. Salvador, Brazil.
- Noll, J., Gebresenbet, M., & Glover, E. D. (2016). A modeling and simulation approach to informal inference: Successes and challenges. In D. Ben-Zvi & K. Makar (Eds.), The teaching and learning of statistics: International perspectives (pp.139-150). New York: Springer.
- Rossman, A. J., & Chance, B. L. (2014). Using simulation-based inference for learning introductory statistics. Wiley Interdisciplinary Reviews: Computational Statistics, 6(4), 211-221.
- Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. International Statistical Review, 67(3), 223-248.