

# Tips and Pitfalls of Converting to Simulation-Based Inference for Large Lectures

Daisy Philtron and Patricia Buchanan, The Pennsylvania State University

## Project

A team at Penn State is converting a large-enrollment introductory statistics course from a traditional curriculum to a simulation-based inference (SBI) curriculum using the Lock5 textbook. In Spring 2017 the new course was piloted with a class of 240 students. A control course using a traditional curriculum was taught at a similar timeslot with the same exam schedule.

**STAT 200:**  
4-credit course  
with  $\approx$  2000  
students per  
semester

**Weekly  
format:**  
2 lectures  
(240-700)  
2 labs ( $\approx$  80)

## Teaching team:

3-5 instructors  
8-12 graduate teaching assistants  
20 undergraduate learning assistants  
Undergraduate guided study group  
leaders

## Next Steps

- Full roll-out of SBI curriculum in Fall 2017.
- Development of lab materials by team of stat 200 instructors.
- Development of deep question banks by graduate teaching assistants.

**Reference:** Lock, R.H., Lock, P.F., Lock Morgan, K., Lock, E.F., & Lock, D.F. (2017). *Statistics: Unlocking the power of data, 2<sup>nd</sup> Edition*. Hoboken, NJ: Wiley.

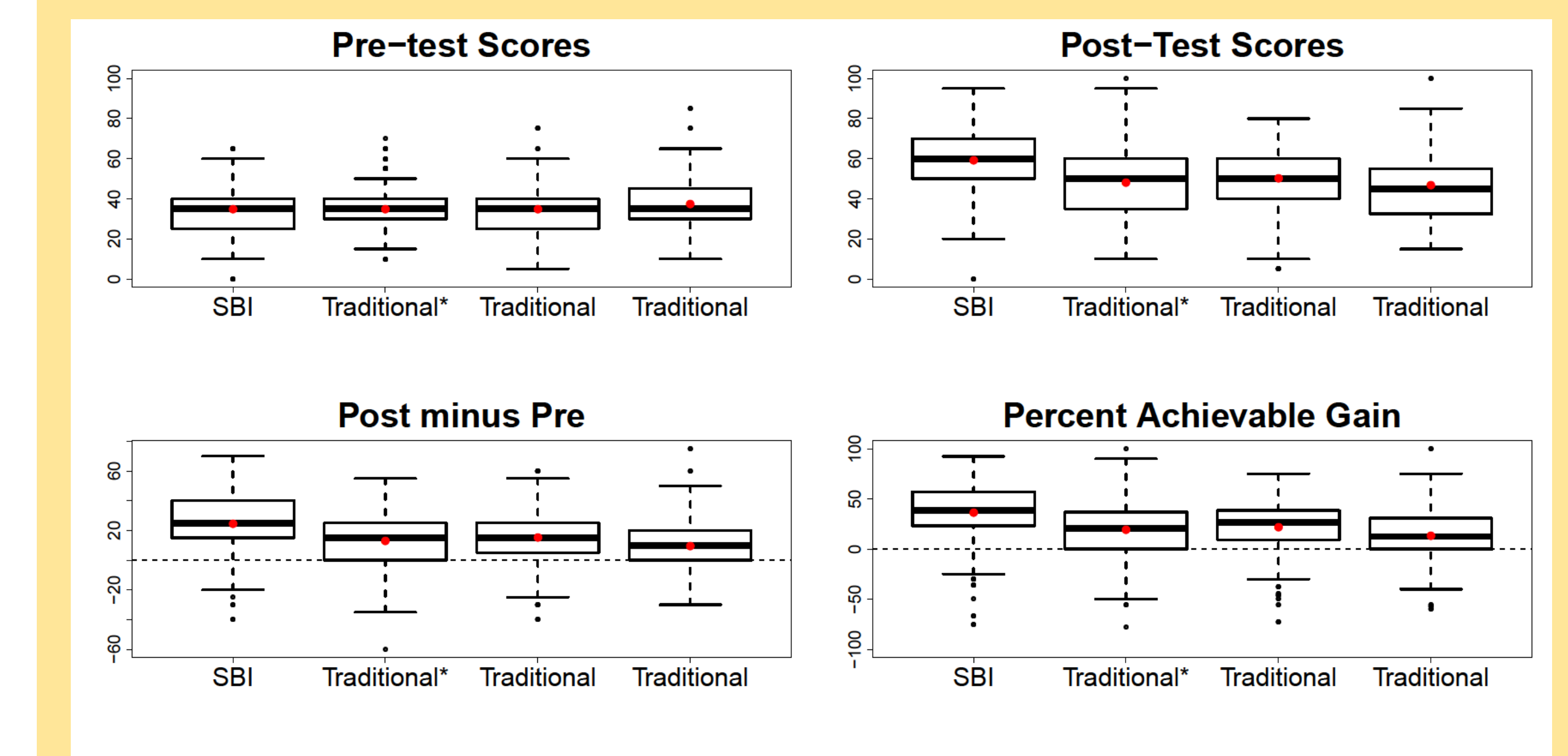
## Tips

- Keep every member of the teaching team up to date on course content.
- Make space for students to interact with SBI software. We used Statkey extensively in active learning labs.
- Ask questions about the software's output and interface to ensure students understand the 'what and why.'
- To maintain interest, source data in real-time using Google Forms.
- Use clickers and a tool such as Doceri for interactive lecture slides
- Reinforce learning with high-frequency, low-stakes exams.

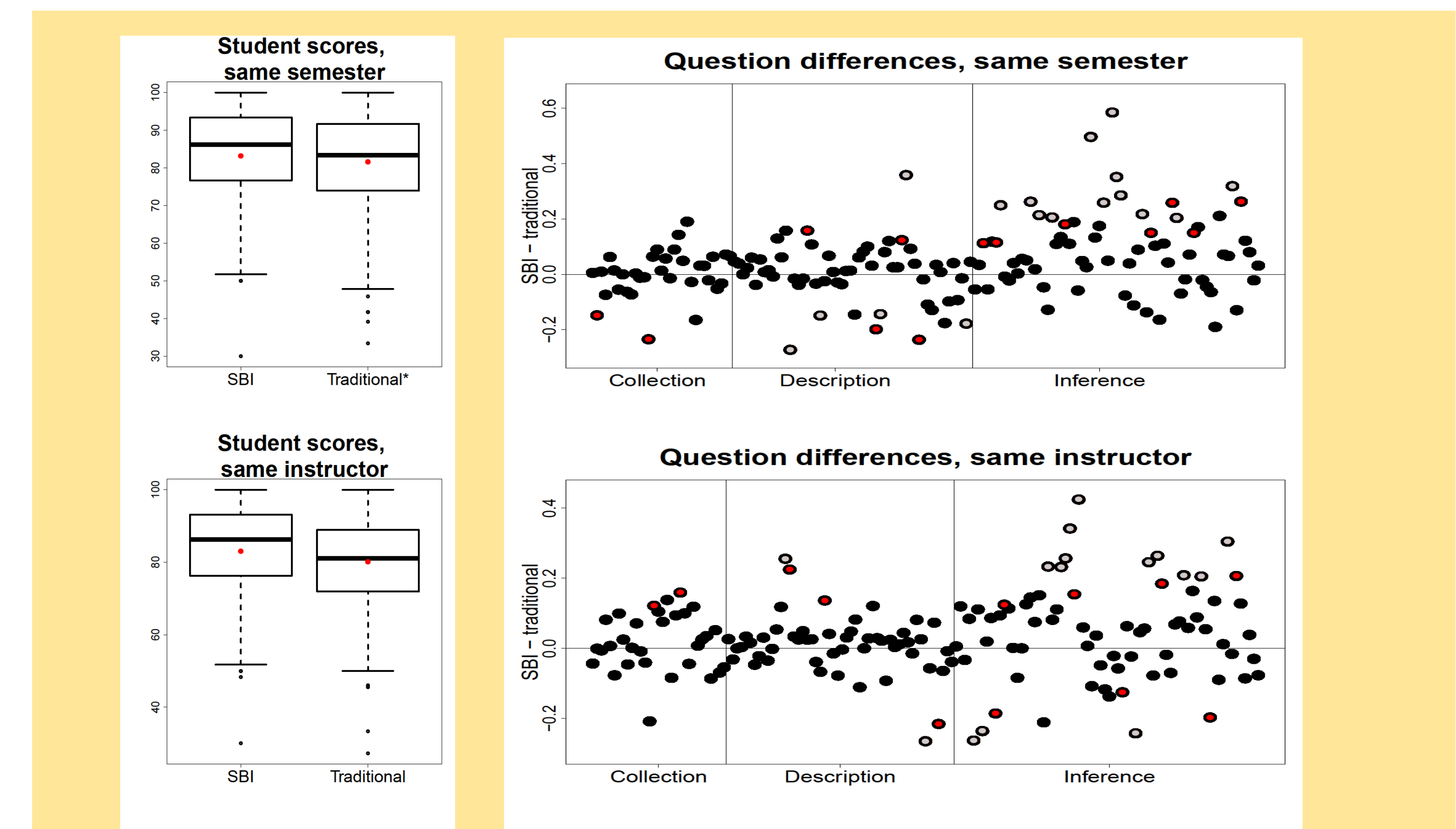
## Pitfalls

- Lack of engagement in large lectures
- Student sabotage of crowd-sourced data (don't use Google Sheets).
- Creation of meaningful automated assessments for lab activities
- Preparation of high-frequency exams

## Results



We compare pre and post scores from the GOALS instrument tests between the SBI pilot course (SBI), the control course (Traditional\*), and other traditionally taught sections offered at different timeslots. Pre-test scores are comparable across all sections, but post-test scores are significantly higher ( $p$ -value $<.001$ ) for the pilot course.



Scores on common cumulative exams were compared between the pilot course and the control course (top row), and between the pilot course and a traditional course taught by the same instructor one year ago. A t-test for the difference in means is not significant for the first comparison ( $p$ -value=0.11), but is significant for the second ( $p$ -value=0.013).

