Development of a General Undergraduate Estimation Skills Survey (GUESS)

Andrew J. Macdonald[1] (ajmacd@physics.ubc.ca), Sarah A. Burke[1,2] (saburke@phas.ubc.ca), Cynthia E. Heiner[1,3,*] (cynthia.heinerfu-berlin.de)

[1]*Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4*

[2]*Department of Chemistry, University of British Columbia, Vancouver, BC, Canada V6T 1Z1*

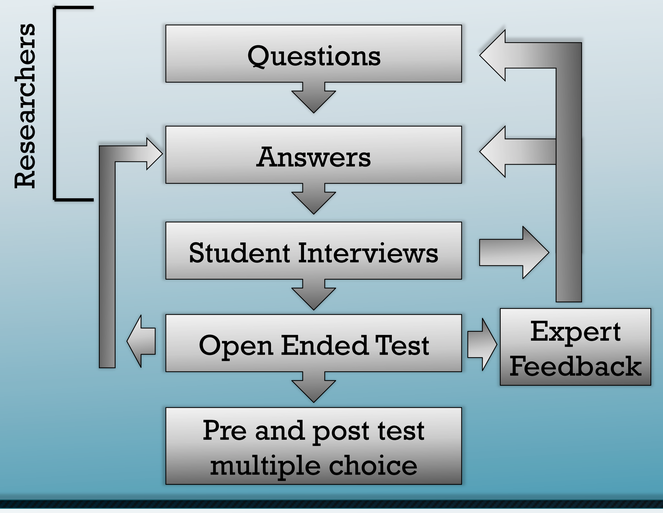
[3]*Carl Wieman Science Education Initiative, University of British Columbia, Vancouver, BC, Canada V6T 1Z1*

[*]*Current Address: Free University Berlin, 14195, Berlin, Germany*

Motivation

- Making good estimations is something that many people, from scientists to public policy makers, are required to do everyday.[1] The ability to make an accurate estimation of a quantity despite incomplete information is a crucial real world skill.
- We developed a short 10 question formative assessment called the General Undergraduate Estimation Skills Survey (GUESS) to give meaningful results about the estimation skills of students.

Method

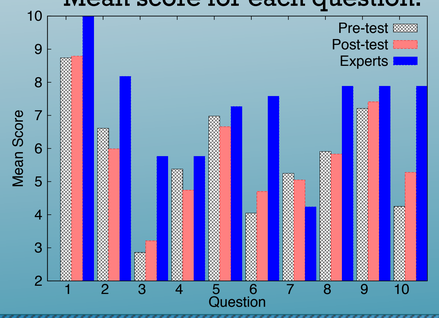


Scores

GUESS scores by sample population.

Sample	Size	Mean	S. E.
Pre-test	301	5.72	0.10
Post-test	521	5.77	0.08
Matched Post-test	170	5.81	0.14
Grad Students	24	7.08	0.39
Post-docs	5	7.40	0.40
Faculty	4	8.00	0.91
Experts	33	7.24	0.30

Mean score for each question.



Sample Questions

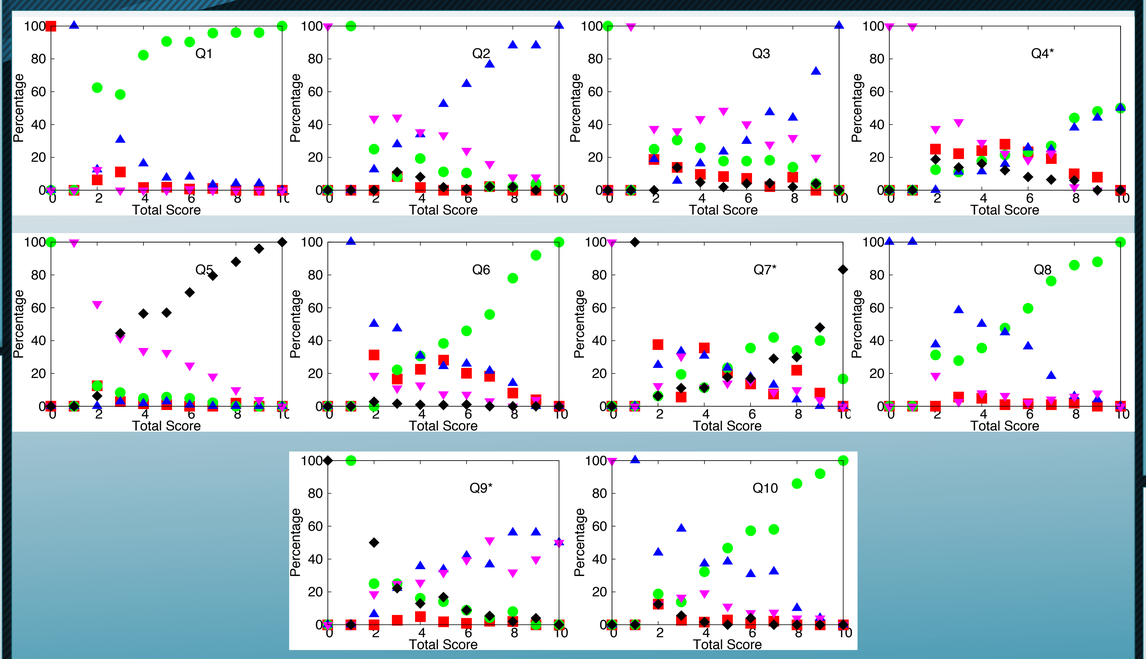
Below are examples of questions that appeared on the GUESS.

Q3. The pressure inside a typical latex balloon is approximately:

A) 0.7 atm
B) 1.0 atm
C) 1.1 atm
D) 3.0 atm
E) 10 atm

Q10 How much time does it take to drive along the highway from Las Vegas, NV to Salt Lake City, UT? (Assume NO stopping on the way)

A) 3 hours
B) 7 hours
C) 10 hours
D) 15 hours
E) 20 hours



Item response curves for each question of the GUESS. Questions with a (*) have two correct acceptable answers. A(■), B(●), C(▲), D(▼), and E(◆).

Item Response Curves

- To test the validity of the GUESS we built Item Response Curves for each question.[2] Item Response Curves plot the percentage of students that chose each answer against student ability. We chose to use total score as a proxy for student ability and used the curves to evaluate the effectiveness of our distractors.
- Analyzing the curves above we see:
  - -Questions 2,3,6, and 10 are effective multiple choice questions
  - -Questions: 4,7, and 9 are moderately effective
  - -Questions 1,5, and 8 are ineffective and should be changed for future versions

Confidence

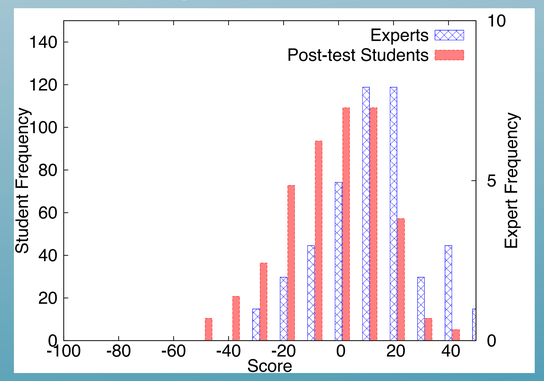
- Each of the GUESS questions also asked students to categorize how confident they were in their answers. In order to correlate confidence with correctness we developed a new marking scheme, similar to those applied in other multiple-choice diagnostics.[3]

GUESS confidence grading scheme.

	Correct	Incorrect
Very Confident	5	-10
Somewhat Confident	4	-8
Not so Confident	3	-3
Not Confident	2	-1
Complete Guess	1	0

- Plotting the expert and student post-test distribution of scores we see that expert confidence correlates more strongly with correctness than the students.

Confidence graded distribution of scores.



Looking forward

- Our use of multiple research procedures in the creation of the GUESS allows us reasonable confidence in the test's construct validity. Further iterations could benefit from our analysis, and need to be deployed to assess reliability.
- Once a fully validated and reliable survey has been developed it can be employed to assess the effectiveness of instructional strategies designed to target estimation skills.

Acknowledgements & References

Thanks to James Day, Natasha Holmes, and all the instructors and teaching assistants of PHYS 153.

[1] E. L. Munnich *et al.*, in *Proc. of the 26th Ann. Conf. of the Cognitive Science Society*, Mahwah, NJ, USA, 2004, pp. 426-432

[2] G. A. Morris *et al.*, *American Journal of Physics* **74**, (5) (2006)

[3] R. B. Frary, Applied Measurement in Education **2**, (1) 79-86 (1989)