Algebra is not like trivia: Self-assessment in an online math tutor

Rachel A. Jansen[1], Anna N. Rafferty[2] & Thomas L. Griffiths[1]
University of California, Berkeley[1] & Carleton College[2]

Introduction

- In order to be self-directed, online learners need to know what information they lack.
- Personalized online educational tools can help these learners gain new knowledge.

→ Understanding more about learners' perceived ability may help improve the personalization of online educational tools.

- In the cognitive science literature, there is a general finding that people are miscalibrated in their performance judgments.[1,2]
- Much of the research on self-assessment is in domains not taught in school.
- There is mixed evidence about students' ability to self-assess in school-taught domains.[3]

Research Questions

- **Do online participants perform similarly to participants from previous cognitive science studies?**
- **How well are people able to self-assess on an algebra task?**

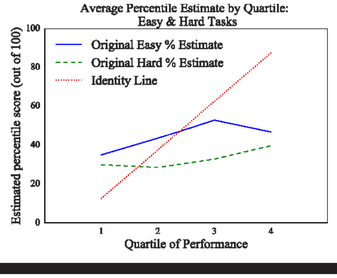
Because most people have experienced a great deal of feedback on their algebra performance, it is likely that individuals will be better calibrated to their algebraic equation solving ability than to their performance answering trivia questions.

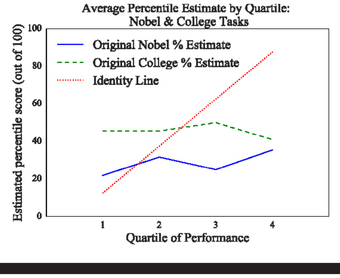
Experiment 1: Trivia

40 participants were recruited from Amazon's Mechanical Turk (19 female, mean age = 30.9 years).

Materials from Burson et al. (2006)[1] were used from two domains: college acceptance rates and years when Nobel prizes were obtained. For both, there was an easy and a difficult set of 10 questions. Participants estimated their percentile estimate after each set (e.g., "Think about these particular 10 estimates. Compared to other participants in this study, how good are you at determining the year of the Nobel Prize within 5 years of the correct year?").

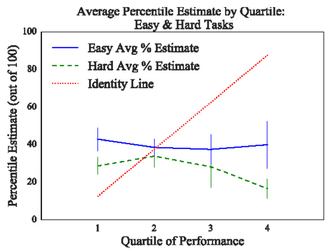
Original results broken out by difficulty (left) and domain (right) indicated overall poor calibration:

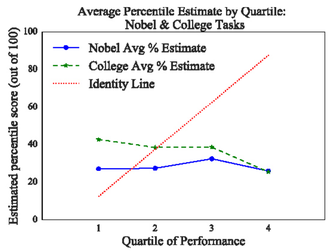




Experiment 1 Results

- Results replicated those from Burson et al. (2006) where both scores and percentile estimates were lower on the difficult versions (left) and on the sets about Nobel prize winning (right).
- Participants' actual performance was only weakly related to actual performance: more difficult questions and domains are rated as more difficult (r=.17, p<.05):

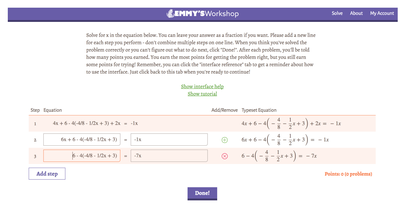




- Online participants perform similarly to participants from the original study on self-assessment in trivia-based domains by Burson et al. from 2006.

Experiment 2: Algebra

39 participants were recruited from MTurk (17 female, mean age = 33.2 years). They:



- Solved 24 problems in Emmy's Workshop,[4] an online algebra tutor
- Then estimated their performance: "How many of the 24 algebraic equations you just completed do you think you answered correctly?" (absolute self-assessment, something not asked in Experiment 1); "Compared to others, how good are you at solving algebraic equations?" (relative self-assessment).

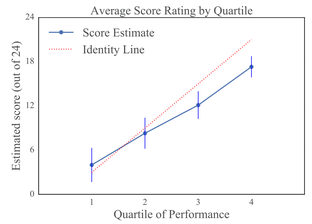
People have received feedback about both their absolute and relative math ability, such as when taking tests and when receiving percentile rankings on standardized tests. We thus would expect similar ability to self-assess in both types of assessments.

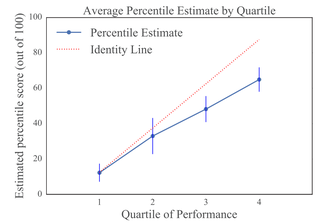
Acknowledgements

This work was supported by NSF grant DRL-1420732 to Thomas Griffiths and Anna Rafferty. Thanks to Priyanka Bhoj for assisting with analyses for Experiment 1 and to Katherine Burson for sharing her study materials.

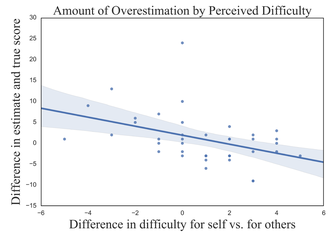
Experiment 2 Results

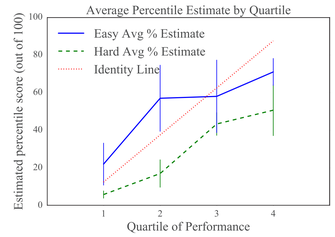
Correlation between actual performance and perceived performance was high for both absolute (left) and relative (right) assessments (r=0.66, p<.001):





The more someone found the task difficult, the more they underestimated their performance:





Conclusions

- Online participants self-assess comparably to participants of Burson et al.'s (2006) trivia-based study: they are generally miscalibrated to their ability relative to others, regardless of actual skill level
- In contrast, participants are overall very well-calibrated to their ability in algebra, both absolute and relative to others.

Future Directions

- How are learners calibrated to individual algebraic subskill ability, (e.g., distribution, arithmetic?)
- Does perceived ability influence actual ability in subsequent problem solving?

References

1 Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons.
2 Dunning, D., & Kruger, J. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments.
3 Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration.
4 Rafferty, A. N., & Griffiths, T. L. (2015). Interpreting freeform equation solving.