

Serial number 0235
r
Title A Methodology that Makes Self-Assessment an Implicit Part of the Answering Process – Preliminary Results
Session New Evidence on Learning Gain (Rogaten)
Submitter Dr. Jekaterina Rogaten
Dr. Paul Mcdermott, Dr Fabio Arico, Mr Jake Patterson

A Methodology that Makes Self-Assessment an Implicit Part of the Answering Process – Preliminary Results

Introduction

In addition to our studies at the University of East Anglia into new pedagogical practice and its effect on self-assessment and learning gain, we are also interested in the development of new self-assessment metric methodologies.

There have been many reported studies in the education research literature for the measurement of self-assessment as a function of academic performance. The most popular methodology employed across these studies is a multiple-choice quiz followed by a confidence tier questionnaire (a Likert scale rating system which “*aims to gain an understanding of how confidently students rate specific responses*” (Brandriet and Bretz, 2014)). This methodology has produced some very interesting outcomes, most of which display the Dunning-Kruger effect. For the uninitiated, Dunning, Kruger and their collaborators have argued for years that the unskilled lack the metacognitive ability to realise their incompetence (Kruger and Dunning, 1999). This “double curse” for less skilled students manifests as poor calibration between their test score and the score they predict for themselves in the type of study described above.

Although other factors including regression to the mean and variations in task difficulty have been proposed as contributors to differences in self-assessment accuracy between good and poor performers, Dunning and Kruger’s rationale remains the predominant explanation. It has been argued that a common limitation to most of the reported studies is that participants rate their performance on a single, just completed measure, typically without feedback about relative performance (Simons, 2013). However, studies in which predictions are made about future performance by people who are fully aware of their own skill level can still give results that align with the Dunning-Kruger effect (Park and Santos-Pinto, 2010). This type of result has given weight to the hypothesis that overconfidence in underperforming students is a result of irrational optimism. Interestingly, there have been studies that achieve significant increases in calibration which were a result of immediate feedback on multiple assessments and judgements (Callender *et al.*, 2016; Renner and Renner, 2001; Huff and Nietfield, 2009). Pleasingly, these results align with our own “self-efficacy as a learning gain metric” studies at UEA.

In addition to these studies we have also undertaken the task of developing novel self-assessment metric methodologies. Towards this end we have posed ourselves the question: “does the Dunning-Kruger effect persist when “reporting confidence” becomes an implicit part of the process of answering a summative assessment question?” Previous research methodology doesn’t necessarily account for the differing motivations a student will experience when answering conceptual questions and then reporting their own confidence in a separate process. We felt that if we were able to tie the motivation to correctly self-assess to the motivation to perform we would have a method that allowed us to see through the fog of subjectivity and irrational optimism.

Methodology

For a number of years we have employed an active learning approach that involves a multiple choice quiz (MCQ) on material that will have been studied prior to class. The answer format for this MCQ has been designed so that students distribute 4 marks across the answer options in a strategic manner to gain the best possible score. There is only one correct answer so if the student is 100% confident they will put all 4 points on one answer option. If they are split between two options they may place two marks on each or split the marks 3 and 1. If they have no confidence in their answer they can place one mark in each answer option to guarantee a point (see Figure 1 below). To mark the answer grid acetate is placed over the answer sheet and the values in the clear boxes are written in the points column. The points are simply added at the end to give the final grade.

Completed answer grid						Acetate answer sheet						Overlay								
Answer Sheet for Hand Scoring						Answer Sheet for Hand Scoring						Answer Sheet for Hand Scoring								
Name	John Smith				Team	3	Name	John Smith				Team	3	Name	John Smith				Team	3
Q. #	A	B	C	D	pts.	Q. #	A	B	C	D	pts.	Q. #	A	B	C	D	pts.			
1	4					1						1	4				4			
2			2	2		2						2			2	2	2			
3			3	1		3						3			3	1	3			
4	1	1	1	1		4						4	1	1	1	1	1			
5	2	1	1			5						5	2	1	1		2			
6				4		6						6				4	4			
					Total Points						Total Points						Total Points	16		

Figure 1: Example Answer Grid and Marking Method

It is our hypothesis that each answer strategy gives a clear indication of the student’s confidence, and furthermore, they do not know they are providing this self-assessment data. They have a different “motivation” as they are focussed on an effort to maximise their grade (the incentive).

Results

In our data analysis we assigned a code to each answer strategy and ranked each strategy in order of increasing entropy whereby placing 4 points on one answer corresponds to the entropy minima and placing 1 mark in each option is the entropy maxima. This allowed us to plot entropy as a function of grade. Our results showed a negative correlation between entropy and student grade with values in the range -0.4 to -0.67 (p-value <0.05). This data does not follow a Dunning-Kruger pattern and very clearly we can see that lower performing students use higher entropy strategies which we can attribute to lower confidence in their answers. In contrast higher performing students very clearly use lower entropy strategies presumably as a result of greater confidence in their own knowledge.

These are exciting results that warrant further analysis. Nonetheless, we believe we have discovered an elegant way to sidestep the problematic phenomena of poor metacognition and irrational optimism that may or may not be prevalent in lower performing students. This gives us a self-assessment metric that instantly correlates to actual student performance. What is even more pleasing is that this methodology can be seamlessly incorporated into the active learning pedagogies that we currently employ.

References

- Brandriet, A. R., Bretz, S. L. 2014. Measuring Meta Ignorance through the Lens of Self Confidence: Examining Students' Redox Misconceptions about Oxidation Numbers, Charge and Electron Transfer. *Chemical Education Research and Practice*, 15, 729-746.
- Callender, A. A., Franco-Watkins, A. M., and Roberts, A. S. 2016. Improving Metacognition in the Classroom through Instruction, Training and Feedback, *Metacognition Learning*, 11, 215-235.
- Huff, J. D., and Nietfield, J. L. 2009. Using Strategy Instruction and Confidence Judgements to Improve Metacognitive Monitoring. *Metacognition Learning*, 4, 161-176.
- Kruger, J., and Dunning, D. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence leads to inflated self-assessment. *Journal of Personality and Social Psychology*, 77, 1121-1134.
- Park, Y. J., and Santos-Pinto, L. 2010. Overconfidence in Tournaments: Evidence from the Field. *Theory and Decision*, 69, 143-166.
- Renner, C. H. and Renner, M. J. 2001. But I thought I Knew That: Using Confidence Estimation as a Debiasing Technique to Improve Classroom Performance, *Applied Cognitive Psychology*, 15, 23-32.

Simons, D. J. 2013. Unskilled and Optimistic: Overconfident Predictions Despite Calibrated Knowledge of Relative Skill. *Psychonomic Bulletin and Review*, 20, 601-607.