Self efficacy and mathematical proof: are undergraduate students good at assessing their own proof production ability?

Paola Iannone [1] and Matthew Inglis [2]

[1] School of Education and Lifelong Learning, University of East Anglia

p.iannone@uea.ac.uk

[2] Mathematics Education Centre, Loughborough University

m.j.inglis@lboro.ac.uk

Abstract

The aim of this research in progress is to investigate how university students assess their own proficiency in producing mathematical proofs and how this compares to their actual performance in proof tasks. Following the educational psychology literature it emerges that self-efficacy is an accurate predictor of academic achievement and the aim of this small study is to investigate self-efficacy in specifically in the context of undergraduate students' proof production. The particular focus on proof is of relevance for its implications for our understanding of undergraduate students' perception of what constitutes an acceptable proof. The study so far consists of a two-part questionnaire administered to 76 undergraduates in one university in the UK. We report the findings from this questionnaire and discuss the implications of the result on students' understanding of mathematical proof.

Background

There is strong evidence in the educational psychology literature (Bandura and Schunk, 1981; Pajares and Graham, 1999; Zimmerman, 2000) that self-efficacy (Bandura, 1977) is an accurate predictor for academic achievement. For example Pajares and Kranzler (1995) found that self-efficacy influences mathematics performance to a degree similar to general mental ability. We define *self efficacy* here, following Bandura (1977), as the judgement students make of their own capability of performing a given task (in our study the task is proof production). It is acknowledged in the mathematics education literature that students find proof one of the most difficult ideas to negotiate when starting mathematics studies at university (Moore, 1994; Weber, 2001). In this study in progress we aim to investigate whether there is a disparity between students' perceptions of their ability in producing proof and their actual proof production skills.

We have concentrated on self-efficacy (following Hackett and Betz, 1989) as this is distinguished from other measures of attitudes (e.g. confidence in learning mathematics) and it can be used as an assessment closely linked to the performance in one specific task and not a general assessment of performance of the whole filed, in this case mathematics (Fennema and Sherman, 1976). We decided to concentrate on proof production because proof is much studied in mathematics education but very seldom in terms of students' self perceptions as provers. Insights into how students' assess their performance at proof production, and comparison with their actual performance might also contribute to the debate about whether students have an accurate idea of what constitutes an acceptable proof.

The study

Participants were 76 first year undergraduates studying mathematics (or a joint degree with a substantial mathematics component) at a highly ranked UK university. Data collection took place towards the middle of the first semester, so participants had, at the time the study happened, completed 8 weeks of degree level mathematics study.

A booklet of questions was given to each participant, and they were asked to work through the questions at their own pace. The first section of the booklet consisted of 28 statements, each of which participants were asked to read and decide the extent to which they believed the statements were characteristic of them, using a five-point Likert Scale (from “extremely uncharacteristic” through to “extremely characteristic”). The order in which the statements appeared was randomised for each participant.

Of the 28 statements, ten consisted of our Proof Self-Efficacy Scale, five of which were reverse-scored. This scale was designed following Bandura (2006). A further ten statements were formed from the items of the General Self-Efficacy Scale (Schwarzer, 1995), which is designed to “assess a general sense of perceived self-efficacy with the aim in mind to predict coping with daily hassles as well as adaptation after experiencing all kinds of stressful life events”. As none of the General Self-Efficacy items are reversed scored, we added eight extra items to redress the balance between forward- and reverse-scored items (taken from the Need for Cognition scale, (Cacioppo, 1982)). The second part of the booklet consisted of four novel proof construction tasks which were designed in collaboration with a mathematics lecturer so that they would represent feasible tasks for this cohort. The order in which the proof tasks appeared was randomised for each participant. The full list of PSE items and proof tasks are given in the Appendix. Participants were asked to stop working on the proof construction tasks twenty minutes after they started.

Results

Two of the 76 participants were removed from the analysis as they did not attempt any of the proof construction questions. Each participants’ Proof Self-Efficacy Index (PSE) was calculated by adding participants’ ratings for each of the forward-scored items, plus the reflected ratings for the negative scored items, yielding a total score which could vary between 10 and 50. A General Self-Efficacy Index

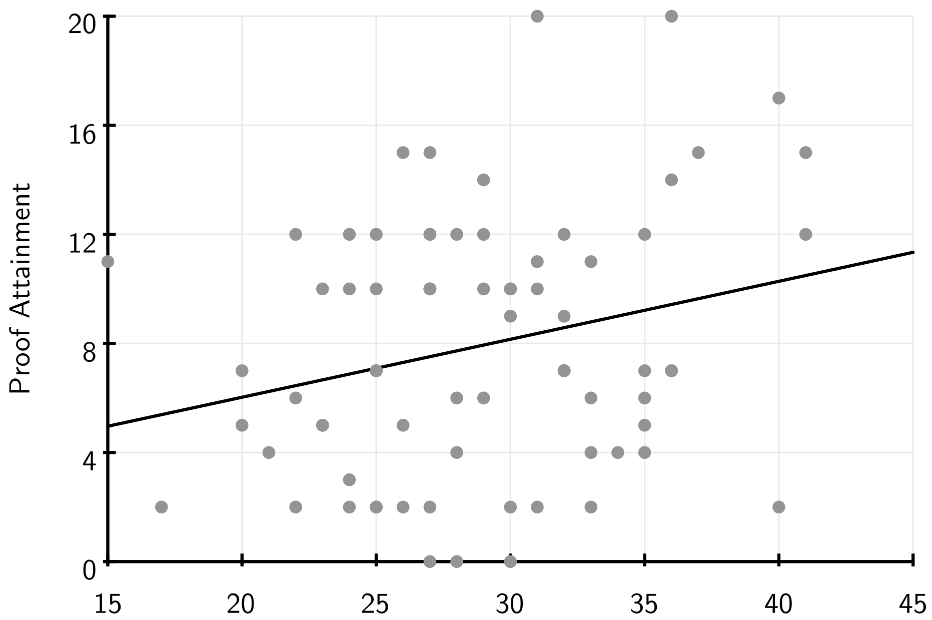

Figure 1: PSE scores against PF attainment

(GSE) was similarly calculated, again yielding a score between 10 and 50. The proof construction items were marked following a mark scheme(again in consultation with the cohort's lecturer) devised prior to data collection (a maximum of five marks were available for each part, yielding a Proof Construction Score (PCS) which could vary from 0 to 20). In Fig 1 we have the representation of the data as a scatter graph.

We first conducted a reliability analysis on the PSE scale, using the Spearman-Brown (random) split-half method. The PSE scale's coefficient was $.816$, indicating it has good internal reliability.

Our main analysis revealed that PSEs were positively correlated with PCSs, $r = .295, p = .011$. This effect retained significance when we partialled out GSE, $r = .289, p = .013$. Following Cohen's (1977) guidelines, both these r-statistics indicate a medium effect size. In summary, those participants who had low perceived proof self-efficacy tended to do worse on the proof construction tasks than those participants who had a high perceived proof self-efficacy. In other words, participants had relatively accurate self-perceptions about their abilities in proof tasks.

We also analysed the responses to the proof tasks classifying them according to the arguments that the students wrote as proofs following Stylianides and Stylianides (2009) scheme. In total the students attempted 221 proof tasks. Of these:

- 86 were correct proofs (scored full marks);
- 76 were reasonable attempts at producing a proof – containing a general argument;
- 57 were responses irrelevant to the solution of the problem;
- 2 were empirical arguments, i.e. proofs entirely based on one numerical example.

Discussion

Our data confirms that there is positive correlation between students' perception of their abilities at producing proofs and their actual proof production performance, in line with general literature about self-efficacy and mathematical ability (Hackett, 1985; Hackett and Betz, 1989). This result was surprising to us as the small interview study from which our original conjecture arose (see Iannone et al, 2009 for a description of this study) seemed to indicate the opposite. The other reason why the results were surprising is for the implication on students' understanding of what an acceptable proof is at undergraduate level. The argument has been made very strongly in the mathematics educational literature that students arrive at university and do not know what is expected from them when a proof is asked for (Moore, 1994; Healy and Hoyles, 2000). Reasons for this include: the inability to conceptualise the necessity of proof (Dreyfus, 1999), the inability to unpack logical statements (Selden and Selden, 1995) and many more. From the analysis of our data we could argue that in fact, after a few weeks of instruction, students have a fair idea of what is asked from them in terms of proof (certainly enough to know whether they are typically able to provide it).

A second observation, linked to the previous one, that can be made form the analysis of the types of proofs is that there are remarkably few (2 out of 221) empirical proofs (Harel and Sowder, 1998), by which we mean justification of the claim by a numerical example. The students in our study had only 8 weeks of instruction at university and yet the suggestion from the data is that only two students thought that a numerical example could justify a general claim.

Concluding remarks

In this research in progress we have shown that there is correlation between how university students assess their own proficiency at producing mathematical proofs and their actual performance on proof tasks. We have also found that eight weeks into a mathematics degree students seem to have accepted the fact that a numerical example cannot prove a general statement, in contrast with many claims of the literature on proof in mathematics education. This small study on its own cannot clearly fully substantiate the two claims which we have made, and further research is need to ascertain whether it is true that students have a good idea of what proof is at the beginning of a mathematics degree and that they know, at this point of their instruction, that a general argument cannot be proved by a numerical example. Further research is also needed to find out, if the two claims are true, why this is the case. Is instruction in mathematics at university level substantially changed in the past ten years? Are there any other reasons that could justify our findings? Is it really the case that being able to assess accurately one's own ability to perform an activity means that there is a good understanding of what this activity entails? We will try to answer such questions in future phases of our research.

Appendix 1 - Proof Self-Efficacy Scale

Those items marked ($*$) were reverse-scored. The order in which items were presented was randomised between-participants.

1. I can understand problems that involve mathematical proof quickly.
2. I can often come up with my own proof for statements I see in class.
3. I am normally able to fit together logical statements into a coherent mathematical proof.
4. I am good at writing mathematical proofs.
5. It is easy for me to come up with convincing and logically sound mathematical arguments.
6. When doing homework problems, I never know how to start a mathematical proof. ($*$)
7. I am better at solving equations than writing mathematical arguments. ($*$)
8. I am never certain whether a mathematical argument that I have written counts as a proof or not. ($*$)
9. I cannot understand why it is important to be able to construct mathematical proofs. ($*$)
10. When trying to write a mathematical proof, I can never find the key idea. ($*$)

Appendix 2 - Proof Construction Tasks

The order in which items were presented was randomised between-participants.

1. Prove that the sum of two odd numbers is even.

2. Prove that the sum of the first n natural numbers is equal to $\frac{1}{2}n(n+1)$.

3. Let d, a and b be integers. Prove that if $d|a$ and $d|b$ then $d^2|(a^2+b^2)$.

4. Prove that if the sum of digits of a natural number is divisible by 3 then the number itself is divisible by 3.

References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, 84, 191-215.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 41, 586-598.

Bandura, A. (2006). *Self-Efficacy Beliefs of Adolescents*. Information Age Publishing.

Bromley, D. B. (1986). *The case-study method in psychology and related disciplines*. Chichester: John Wiley & Sons.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (revised edition)*. New York:

Academic Press.

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics*, 38, 85-109.

Fennema, E. & Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitude Scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *JSAS Catalog of Selected Documents in Psychology* 6,31. (Ms.No.1225)

Ginsburg, H. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the Learning of Mathematics*, 1(1), 4-11.

Hackett, G. (1985). The role of mathematics self-efficacy in the choice of math-related majors of college women and men: A path analysis. *Journal of Counseling Psychology*, 32, 47-56.

Hackett, G & Betz E. N. (1989). An exploration of the mathematics Self-Efficacy/mathematics performance correspondence. *Journal for Research in Mathematics Education*, 20(3), 261-273.

Harel, G. & Sowder, L. (1998). Students' proof schemes: results from exploratory studies. In E. Dubinsky, A Schoenfeld and J Kaput (Eds.), *Research in collegiate mathematics education, III*. American Mathematical Society, Providence, RI, 234-283.

Healy, L. & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 31(4), 396-428.

Iannone P., Inglis, M., Mejia-Ramos J., Siemons J. & Weber K. (2009). How do undergraduate students generate examples of mathematical concepts? *Proceedings of the 33th Annual Conference of the International Group for Psychology in Mathematics Education*, Thessaloniki, Greece, (3) 217-224.

Moore, R.C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics* 27, 249-266.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology*, 20, 426-443.

Pajares F. & Graham L. (1999). Self-efficacy, motivation constructs and mathematics performance of entering middle school students. *Contemporary Educational Psychologist*, 24, 124-139.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A users portfolio. causal and control beliefs* (p. 35-37). Windsor, UK: NFER-Nelson.

Selden, A. & Selden, J. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29, 123-151.

Stylianides A. J. & Stylianides G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics* 72(2), 237-253. Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, 48(1), 101- 19.

Weber, K. (*in press*). How syntactic reasoners can develop understanding, evaluate conjectures, and generate counterexamples in advanced mathematics. *Journal of Mathematical Behavior*.

Zimmerman, B. (2000). Self-efficacy: an essential motive to learn. *Contemporary Educational Psychology*, 25, 82-91.