Improving formal reasoning in mathematics through tutorials*

Douglas Stebila
Dept. of Combinatorics & Optimization, University of Waterloo, Waterloo, Ontario, Canada
dstebila@uwaterloo.ca

2008 March 11

Contents

*Research project submitted for the Certificate in University Teaching at the Centre for Teaching Excellence, University of Waterloo.

1 Introduction

Although many students view mathematics as being about numbers and formulas, for mathematicians the subject is about abstract reasoning and formal logic. A primary goal in math courses then is for students to learn to approach mathematics rigourously, learning the formal tools of theorems and proofs. Only when these formal techniques are mastered can their applicability in solving real world problems be understood and appreciated.

Many mathematics courses, however, tend to emphasize the numerical and formulaic aspects of mathematics rather than the abstract reasoning, in part because it is difficult to teach and assess abstract reasoning. The settings in which students learn and solve coursework problems – via lectures and independent problem solving – do not correspond with how mathematicians solve research problems – through discussion and collaboration. The value of peer teaching and groupwork has long been known in the education literature, but it has been difficult to integrate these pedagogical techniques into traditional mathematics lecture-style courses.

In this paper, I propose the use of student-led presentations of solutions to abstract reasoning problems in tutorials as a technique to improve student skills in abstract reasoning. The proposal is motivated by the reported benefits of various peer teaching and groupwork scenarios, and encourages the use of comprehension tests and proof validation to improve formal reasoning.

2 Background

Before proceeding with the main examination of techniques useful to this proposal, it is important to review relevant background principles, namely the importance of assessment and the consideration of learning styles.

2.1 Importance of assessment

Brown et al. (1997) note the important role assessment plays in shaping what students learn. Based on various studies, they conclude that “students take their cues from what is assessed rather than from what lecturers assert as important” (Brown et al., 1997, p. 7), and moreover that changing the assessment technique can change student learning behaviour. In one reported study, for example, medical students spent more time in wards after clinical assessment was changed from pass/fail based on ward reports to a practical clinical exam. Thus, one can expect that changing assessment criteria in math courses can cause students to change their behaviour, and the goal of improving student skills at formal reasoning can be aided by appropriate assessment tools.

2.2 Learning styles

Knisley (2002) proposes a set of mathematical learning styles to reflect Kolb's model of experiential learning (Kolb (1984), Evans et al. (1998)). A student's learning style in Kolb's model is based on student preference in two factors: concrete versus abstract, and active experimentation versus reflective observation. These four learning styles have been interpreted for mathematics education by Knisley (2002) as follows:

- *Allegorizers* (concrete, reflective): These students view new ideas as reformulations of old ideas, and attempt to solve problems by applying known techniques.
- *Integrators* (concrete, active): These students compare new ideas to old ideas, and solve problems by looking for parallels with old problems.
- *Analyzers* (abstract, reflective): These students use algorithms and procedural explanations to solve problems in a logical, step-by-step progression.
- *Synthesizers* (abstract, active): These students use existing concepts to construct new approaches and solve problems by developing new strategies.

Synthesizers are at the most advanced learning stage, but instructors should address all four learning styles and try to bring students from the earlier learning stages into more advanced learning stages.

Stewart and Thomas (2007) describes three methods of thinking in mathematics: the *embodied world* (consisting of perceptions of real world objects and visuo-spatial imagery), the *symbolic world* (consisting of 'procepts', in which processes and objects are symbolized), and the *formal world* (consisting of defined objects, discussed in terms of their properties). While mathematicians usually work in the formal world, mathematics students start in the embodied world and the goal of courses is to move them through the symbolic world to the formal world. Many math courses often use assessments that test understanding of the embodied and symbolic world, but it is difficult to test the formal world. The next section describes some tools for developing and assessing students' knowledge of the formal world.

3 Tools related to formal reasoning

One tool for enhancing student understanding of theoretical techniques is proof validation, and this can be assessed using comprehension tests, both of which are described below.

A mathematical *proof* is a sequence of formal logical statements that follow one from the other until a particular statement is proved true, the final statement usually being called the *theorem*. Any sequence of formal logical statements is called an *argument*, as in Selden and Selden (2003). In other words, a proof is an argument in which each statement of the argument follows logically from the previous statements, and hence proves the theorem. However, not all arguments prove a theorem. A *false* argument is one in which the statements do not logically follow one from another, and thus do not actually prove what it may be claimed that they prove.

3.1 Proof validation

The process of checking whether each statement in an argument follows from previous statements, and thus of checking whether the final statement follows from the argument, is called *proof validation*. It is a longer and more complex written and mental procedure than the written proof. But for mathematicians interested in the abstract reasoning underlying mathematics, proof validation is an essential skill; however, proof validation is often only taught implicitly.

A study by Selden and Selden (2003) examined experimentally whether undergraduate mathematics students could determine with a given argument actually proved a theorem or not. In the study of 8 undergraduate mathematics students, each student was given a mathematical statement, and then four arguments which were purported to prove the given theorem. Students were given an opportunity to review each argument, asked to indicate which ones they thought were true and which were false, asked some general questions on how they went about trying to validate the arguments, and finally asked to again state whether they thought each argument was true or false.

The study analyzed when, in the process described above, each student was correct on their judgement of the correctness of the arguments. Overall, after just having seen the arguments once, students were correct only 46% of the time, which suggests that on a test, students would not do better than simply guessing randomly. However, by the end of the procedure, the students were correct 81% of the time showing that, in the interactive environment of the interview procedure, students were over time able to much better judge the validity of an argument.

3.2 Comprehension tests

Conradie and Firth (2000) describe an assessment tool called *comprehension tests*. These assessment questions present a theorem and proof, and then ask a number of questions about the proof, such as "What proof technique was used here?", "Why does statement (3) follow from statement (2)?", and "What does the argument prove?"; a sample comprehension test question is given in Figure 1. They note the importance of properly preparing students for comprehension tests by including, for example, comprehension testing questions on assignments.

1. Read the following theorem and its proof:

   Theorem 1 *$\sqrt{20}$ is irrational.*

   Proof. Suppose there are integers m and n such that $\sqrt{20} = \frac{m}{n}$. Without loss of generality we may assume that m and n have no factors in common. Now $m^2 = 20n^2$. Hence 5 is a factor of m^2, and so 5 must be a factor of m. We can therefore write $m = 5k$, for some integer k. Then $25k^2 = m^2 = 20n^2$, or $5k^2 = 4n^2$. Hence 5 is a factor of n^2, and hence of n. But then 5 is a factor of m and n, contradicting our assumption. □

   Now answer the following questions:

   (a) What method of proof is used here?
   (b) How is $\sqrt{20}$ defined?
   (c) When is a real number irrational?
   (d) Why may we assume that m and n have no factors in common?
   (e) Given that 5 is a factor of m^2, how does it follow that 5 is a factor of m?
   (f) Why does $5k^2 = 4n^2$ imply that 5 is a factor of n^2?
   (g) Which assumption is contradicted, and how does the theorem follow from this?
   (h) The equation $m^2 = 20n^2$ also implies that 2 is a factor of m^2. Could we have used this to prove the theorem?

Figure 1: Sample comprehension test question. (Conradie and Firth, 2000, p. 227)

In describing comprehension tests, Conradie and Firth do not include a statistical study of the use of comprehension tests, but note that, in student evaluation forms, students overwhelmingly (90-95%) preferred comprehension tests to reproductions of proofs, "despite the fact that some of them remarked that it was the more difficult form of testing" (Conradie and Firth, 2000, p. 232). Moreover, the authors note that, at their institution, mathematics lecturers who have used comprehension tests "have always been surprised (if not shocked) at the results and are, in general, impressed with the insight it provides" (Conradie and Firth, 2000, p.227).

Rensaa (2007) describes an experiment that gave students a choice between a standard theorem/proof-style question and a comprehension test-style question on a final exam. Despite the fact that students had not been previously exposed to comprehension test-style questions in earlier assessment exercises in the course, 52% of respondents chose the comprehension test-style question instead of the theorem/proof-style question. Students who chose the more traditional theorem/proof-style question had better overall results on the exam, but the authors speculate this may be because students had not seen comprehension test-style questions before:

> "There is a tendency for students with the best grades to choose the familiar task [...], while students with the lowest grades go for the new type of proof explanation. It seems as if tasks where the proof is given look easier to unprepared students." (Rensaa, 2007, p. 736)

Following the exam, students were given a short questionnaire asking for their opinion on the choice option. Overwhelmingly, students responded positively about having a choice on the exam, and that having a choice may allow students to better demonstrate what they have learned; a few students noted however that making a choice requires additional time in a situation with time constraints. Opinion in favour of or opposed to the choice option was statistically independent of students' final grades, indicating that students of all skill levels enjoyed the choice option equally.

4 Peer teaching and workshops

Peer teaching – "students teaching other students" (McKeachie et al., 1986, p. 63) – is widely acknowledged as one of the most effect methods of teaching, and has been extensively studied in the literature. Peer teaching can be implemented in a variety of formats, from one-on-one partner work up to student-led discussions in tutorials and lectures. This section provides some background on peer teaching in general and then reports on various peer teaching scenarios in undergraduate mathematics education.

4.1 Peer teaching

McKeachie et al. (1986), in section V.A, report on studies of peer learning and teaching, and indicate there is a wealth of evidence that peer teaching is extremely effective for reaching students with wide ranges of learning styles. In one reported study, physics students participating in student-led discussions performed superior to traditionally-lectured students on a test containing complex problems and learning new material (although traditionally-lectured students did better on tests of simple facts and memorization). Moreover, students in such discussions asked more questions, and students with below-median ability achieved better results than in lecture-based learning. Similar results were observed in studies in other subjects reported in the same survey.

A study on small group peer teaching was done by Evans et al. (2001). In this study, instructors of two small first-year mathematics classes (18 and 12 students) integrated peer teaching exercises into their lectures. The two classes were joined and formed groups of size 2 or 3 (with one student from each class, and extras from the larger class distributed as appropriate); in one week students from one class taught their peers from the other class a pre-assigned subject, and in a subsequent week the roles were reversed. These peer teaching exercises happened during the normal lecture period as opposed to during a tutorial or specially-scheduled session. Feedback from students reported that students enjoyed the experience but were nervous about presenting their material at the right level. One student commented:

> "It was also worthwhile from the point of view that to be able to teach a subject well, you need to have a good in-depth knowledge and this meant we all went away within our group and researched the subjects required thus acquiring more than just a 'working knowledge' passed on from [the instructor]. It is also a very good way to check progress." (Evans et al., 2001, p. 165)

Constructive feedback from the students emphasized the need for clear assessment criteria. Interestingly, the instructors observed that the time it took students to peer-teach the material was less than would have happened if it had been done in a lecture setting.

4.2 Workshops and group work

The use of small group interactions was studied by vern Bonsangue (1994) in the context of enrichment workshops for minority (Black or Latino) students in first-year calculus courses. The twice-weekly 2 hour workshops of 10 to 12 students consisted of 30 minutes of individual problem solving, then peer discussions and collaboration, and finally instructor-led group discussions.

In the study, workshop students achieved higher results in first-quarter calculus (mean grade 2.69) than their Black or Latino peers who did not participate in the workshop (mean grade 1.76) and improved their results to surpass Asian and White students (mean grades 2.40 and 1.98, respectively); self-selection effects compared to other Black and Latino students seemed to be minimal as both groups scored similarly on pre-college academic measures. Moreover, workshop students had significantly higher rates of enrollment in mathematics, science, and engineering after 3 years in the longitudinal study (85%, compared to 52% for the other groups described above).

Fernández et al. (2002) reports on two different classroom formats for problem solving. In the "whole-class groupwork" format, students work in small, *ad hoc* groups in the classroom, while the instructor circulates among the groups, providing communication to the class as a whole and mediating discussions between various groups on the overall problem.

In the "office-hours groupwork" format, the instructor brought together a selected group of students for problem solving sessions. In this format, the instructor provided a problem to all of the students who would be attending, asked the students to submit solutions, provided all students with photocopies of each others' solutions as well as the instructor's, and then mediated a discussion between the students on the relative merits of each proposed solution. The office-hours groupwork format provides "students opportunities for more focused and individualized thinking" (Fernández et al., 2002, p. 259), and students reported enjoying the insight that came from these discussions.

One of the challenges of the office-hours groupwork format is that it involves only a small number of students, and thus must be repeated many times over the course of the term in order to reach all the students in a large class. As well, it does not serve as a substitute for office hours on their own. Balancing interactive teaching strategies with the restricted amount of time available to both instructors and students is a challenge, and unfortunately the authors do not discuss the amount of additional time needed to properly employ these groupwork formats.

5 Proposal: student-led proof presentations in tutorials

In the author's experience, a typical undergraduate mathematics course is delivered almost exclusively through a combination of lectures and assignments. At the University of Waterloo, for example, most undergraduate mathematics courses consist of three fifty-minute lectures per week (in sections of up to 120 students), and have between 5 and 10 assignments over a 12-week term. In first and second years, many courses also include one fifty-minute tutorial each week (in groups of up to 150 students), which is often devoted to having the lecturer address questions on the current assignment.

The typical mathematics lecture does not include much interaction. A good lecturer will use effective question-asking strategies and ask for student input at certain points in the lecture, but the lecture primarily consists of the written and oral presentation by the instructor of definitions, theorems, proofs, and examples. This author, for example, did not have a single lecture in his entire undergraduate mathematics career involving in-class group work in a math class, which is why the techniques presented in Section 4, on workshops and group work, are novel.

In order to improve students' formal reasoning skills, I propose using weekly tutorials to have students present their solutions to problems that encourage skill development in proof validation and comprehension tests.

Students will be assigned to weekly fifty-minute tutorials of approximately 30 students per tutorial. In these weekly tutorials, supervised by a graduate teaching assistant, students will present their solutions to pre-assigned problems and discuss them. In particular, at the beginning of the term a schedule will be established with 3 student presentations per week. All students will be provided with the 3 problems to be presented at the tutorials approximately one week in advance. The presenting students will be expected to prepare a solution and present it on the blackboard at the tutorial; each student will be allotted 15 minutes for the presentation and discussion of the problem, mediated by the graduate teaching assistant.

My proposal requires some additional resources compared to the existing course setup at the University of Waterloo, for example. It replaces a weekly fifty-minute tutorial for 150 students with a fifty-minute tutorial for 30 students, a five-fold increase in tutorial hours. For a course offering to 300 students, it would require the addition of 2 graduate teaching assistants (at 5 hours per week) to staff these additional tutorial hours. However, there would be no net increase in the number of scheduled hours per enrolled student.

In order for students to take the presentations seriously, it is important that they be assessed and included in the final course mark, for, as Brown et al. (1997) observed, "students take their cues from what is assessed rather than from lecturers assert as important" (p. 7). Section 5.1 includes an assessment tool for evaluating mathematics presentations by students. The presentation could be worth 5% of the final grade, for example.

One issue the resolution of which is unclear to us is how to ensure students attend the tutorials even when they are not presenting. It would be unfortunate if attendance at the tutorials degenerated to the point where only the graduate teaching assistant and the students presenting attended each week. Some instructors might choose to include a participation mark, although this author has never been fond of participation marks except as a last resort. A more palatable solution would be to arrange it so that the problems presented each week are of interest to students: if the problems presented each week are closely related to, but not identical to, problems on that week's assignment, students may see the connection between the problems and be interested in attending to get hints at how to do a similar problem. It is impossible to know *a priori* what type of attendance can be expected for a setting such as this, but with time instructors will come to understand how to motivate students' attendance.

The format proposed herein somewhat resembles the "office-hours groupwork" format of Fernández et al. (2002), but has the advantage of including a larger number of students. It includes some aspects of the same authors' "whole-class groupwork" format, and is an abbreviated form of the workshops reported on by vern Bonsangue (1994). With students doing the presentations, it includes an aspect of peer teaching, although not to the same one-on-one extent as in Evans et al. (2001).

This presentation and discussion format also has the benefit of reaching students who have various learning styles, and may in particular well-suit allegorizers and integrators who, as described in Section 2.2, both strongly depend on relating new ideas to old ideas. When students are required to give their own presentations, they need to understand the material in greater depth and may move from beginning learning stages (allegorization and integration) to more advanced stages (analysis and synthesis).

The tutorial format described above could be used with any type of problems: numerical evaluation, case studies, or formal proofs. In light of the discussion on formal reasoning in Section 3, I believe there would be great value in using problems that examine formal proofs and focusing discussion around questions related to proof validation and comprehension tests. Each problem could ask students to give a proof for a small theoretical question of an appropriate difficulty level. As part of their presentation, students could be asked to describe the method of proof, explain how the theorem follows from the argument, and answer other questions like those in the sample comprehension test in Figure 1.

The presentations provide an excellent opportunity for proof validation for students attending the presentations. The proofs being presented will be realistic and have complex arguments, since they will be developed by students. In the study of Selden and Selden (2003), the arguments that students were asked to validate were based on solutions written in another context by undergraduate students, and thus were realistic and complex arguments. By contrast, the same authors note that the few textbooks that do provide proof analysis exercises use simplistic proofs constructed so that only a single error needs to be detected, whereas student constructed proofs are often less clear and require deeper analysis. Thus, student presentations are a basis for good proof validation.

5.1 Assessment tool for presentations

Since one of the main goals of this proposal is to improve students' formal reasoning skills, the correctness of the argument and the logical flow of the argument presented are an important component of the evaluation. Also important are the presentation skills, both oral and on the blackboard, because it is important for students to be able coherently communicate their arguments.

Figure 2 proposes a possible assessment tool for student presentations of mathematical proofs. It is adapted from Brown et al. (1997), p. 159, and equally weights the mathematical correctness of the presentation with the oral and written quality of the presentation, providing room for comments. It is important to provide students with the rubric for evaluation before their presentation so that they understand what they need to do in order to succeed in the presentation.

5.2 A classroom experience

This author used some problem solving sessions for exam review while teaching first-year linear algebra courses (MATH 136 at the University of Waterloo) for mathematics students. The first time I taught the course, I provided students with the questions once they arrived at the review session, scheduled some time at the beginning for students to work on solutions individually or in small groups, and then asked for students to present their solutions to the class. Unfortunately, most students did not seem to complete the solution to even a single problem during the time allotted, leaving me as the instructor to present solutions.

The second time I taught the course, I provided students with a set of questions in advance of the review session, informing them that they would have the opportunity to present solutions to the class. This time, many more students were ready to present their solutions, and I as the instructor did not need to present any solutions, although I helped students with their presentations and responses to questions as necessary.

In addition to seeing a more favourable response from students in terms of their willingness to present, the second format was also more efficient in terms of classroom time use. However, the first format may still be useful if students become accustomed to the format; in the experience I reported, this was the only time that students had a classroom

Math Presentation Assessment

Student: **Date:**

Tutor:

Criteria	*Marks*	*Comments*
Correctness Correctness of written solution, logical flow, correct use of terms in oral presentation	____ / 10	
Visual Presentation Clarity, neatness, easy-to-follow	____ / 5	
Oral Presentation Eye contact, verbal skills, responses to questions	____ / 5	
Total mark	____ / 20	

What's good?

What needs improving?

Figure 2: Math presentation assessment instrument, adapted from Brown et al. (1997), p. 159.

setting like this the entire term. Moreover, it was not an assessed exercise, only a supplemental exercise for exam preparation. If it was assessed, and all students were expected to present at least once over the term, the differences between providing the questions in advance and providing the questions at the event may be mitigated.

Anecdotal feedback has suggested that students in the second format enjoyed the opportunity to present their solutions to their peers and, with careful mediation by the instructor, students can do so comfortably and handle mistakes and questions in a confidence-building manner. I am eager to see how this format would succeed in a weekly tutorial as I have proposed, and hope to be able to test this proposal soon.

References

Brown, G., Bull, J., and Pendlebury, M. (1997). *Assessing Student Learning in Higher Education*. Routledge, London.

Conradie, J. and Firth, J. (2000). Comprehension tests in mathematics. *Educational Studies in Mathematics*, 42(3):225–235.

Evans, N. J., Forney, D. S., and Guido-DiBrito, F. (1998). *Student development in college: Theory, research, and practice*. Jossey-Bass.

Evans, W., Flower, J., and Holton, D. (2001). Peer tutoring in first-year undergraduate mathematics. *International Journal of Mathematical Education in Science and Technology*, 32(2):161–173.

Fernández, E., Kazimir, J., Vandemeulebroeke, L., and Burgos, C. (2002). Experimenting with classroom formats to encourage problem solving. *Problems, Resources, and Issues in Mathematics Undergraduate Studies (PRIMUS)*, 12(3):247–261.

Knisley, J. (2002). A four-stage model of mathematical learning. *The Mathematics Educator*, 12(1):11–16.

Kolb, D. A. (1984). *Experiential learning: experience as the source of learning and development*. Prentice-Hall.

McKeachie, W. J., Pintrick, P. R., Lin, Y.-G., and Smith, D. A. F. (1986). *Teaching and Learning in the College Classroom: A Review of the Research Literature*. Regents of The University of Michigan.

Rensaa, R. J. (2007). A choice option between proofs in linear algebra. *International Journal of Mathematical Education in Science and Technology*, 38(6):729–738.

Selden, A. and Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 34(1):4–36.

Stewart, S. and Thomas, M. O. J. (2007). Embodied, symbolic and formal thinking in linear algebra. *International Journal of Mathematical Education in Science and Technology*, 38(7):927–937.

vern Bonsangue, M. (1994). An efficacy study of the calculus workshop model. *CBMS Issues in Mathematical Education*, 4:117–137. Research in Collegeiate Mathematics Education. I.