

Assessing individual work alongside a group project - a partially-automated approach

Peter Rowlett

Mathematics

School of Science & Technology

Nottingham Trent University

A new module

- Final year of a UK mathematics degree (not at NTU).
- Aimed to develop certain graduate skills, including:
 - real-world problem-solving skills;
 - working in depth on problems;
 - report writing;
 - working in groups;
 - communicating results to different audiences;
 - articulation of graduate skills.
- Delivery via group projects.

'Student transitions'

- (The conference theme.)
- I am not serving my students or their potential employers well if I allow them to progress through the project being carried as a 'passenger'* by their group.
- Therefore, if we care about the students' transition to employment, we must care about uneven contribution.

* MacBean, J., Graham, T. and Sangwin, C., 2001. *Guidelines for Introducing Groupwork in Undergraduate Mathematics*. Birmingham, U.K.: Maths, Stats and OR Network.

Previous experience of group work

- Previously*, I ran group projects in mathematics in the second year of a mathematics degree.
- I had groups keep minutes of meetings and told them they could use these to bring me evidence of uneven contribution.
- In evaluation, students agreed minutes were useful but many said that uneven contribution still occurred.

* Rowlett, P., 2010. Guarding art galleries: design and evaluation of an undergraduate research based group project in mathematics. *Proceedings of Annual Learning and Teaching Conference 'Developing students as scholars'*, Nottingham Trent University, 29 March 2010, pp. 12-14.

Back to the new module

- Investigate a scenario from a client, mathematically.
- For uneven contribution:
 - minutes of meetings, but previous experience says this is not enough, so:
 - peer assessment of contribution, but concern about reliability, validity (did they contribute equally to all elements?), complexity, so:
 - individual assignments.

Individual assignments

- 1. Reflective essay on 'what the client wants', to assess individual contribution on learning outcomes around:
 - 'communicating results to different audiences';
 - 'articulation of graduate skills'.

Individual assignments

- 2. Mathematical assignment, to assess learning outcomes:
 - 'real-world problem-solving skills';
 - 'working in depth on problems';
 - 'report writing'.
- Problem: Effectively, students have three problems to solve as a group and one on their own, so the temptation to treat these as four problems as a group will be very high.
- This suggests exam conditions or individualised work (randomisation via e-assessment) would be advantageous.

Exam conditions

- This would mean a class test on the general area.
- Really, the learning outcomes suited a deeper, more open form of question than would be possible in this environment.
- Deeper, more open questions doesn't very well suit a timed, high stress class test.

E-assessment

- The deeper, more open questions with discursive answers are not suitable either for automated marking.
- Also, the topic used means that solutions would involve drawing diagrams, so computer input starts to add substantial additional, irrelevant learning outcomes.

A partially-automated approach

- Ideally, we want randomisation, but can't cope with automated marking.
- A partially-automated approach would help, using computer generated individualised worksheets that are distributed, answered and marked on paper.
- A specially adapted version of Numbas¹ (Newcastle University) was used to generate individualised worksheets.

1. www.ncl.ac.uk/math/numbas

Mathematical assignment

- 2 questions, one involving one diagram randomly selected from nine, included randomised or calculated numbers.
- Answer hints were written for the marker, including numbers calculated from the randomised values.
- Saved as PDF in the browser for printing.

agp2

You can edit this question.

Author: Peter Rowlett (prowlett)

Progress:

[▶ Test Run](#) [✖ Delete](#) [⊕ Make a Copy](#) [↓ Download: standalone .zip, SCORM package, source](#)

[General](#) [Statement](#) [Variables](#) [Functions & Rulesets](#) [Parts \(0\)](#) [Advice](#) [Resources \(0\)](#) [Exams using this question](#)

[Toggle rich text editor](#) ✓

B *I* U Paragraph |

Draw a polygon using $\{p\}$ vertices for which $\{q\}$ guards are necessary to guard every point at any one time.

Path: p

Feedback

+

↑
↓

p



16

Definition

```
1 random(10..20)
```

+

↑
↓

q



5

Definition

```
1 floor(p/3)
```

Add another variable

Automatically recalculate variables:

Regenerate values

Generate exam, (starting with ID no.).

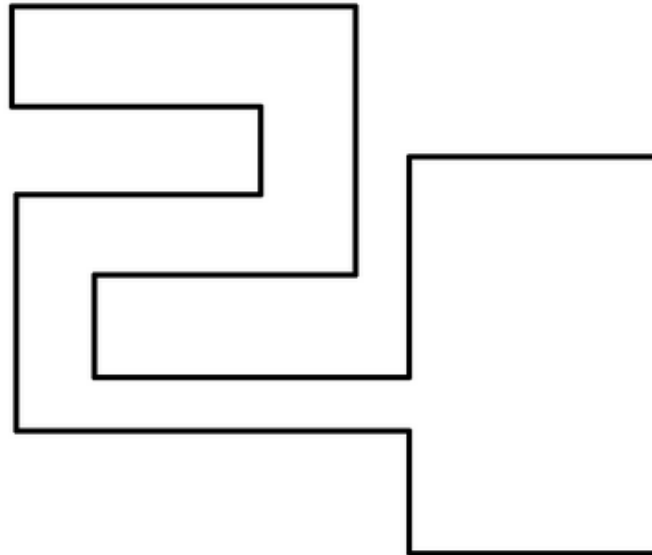
Now

- Question sheets
- Answer sheets

ID: 1

MAT30029: Individual assignment 2

1.



a) Show, by triangulating and three-colouring the polygon, how many guards are necessary to guard every point in the museum shown above at any one time.

b) In reality, given a staff of 11 guards, could you arrange for every point in the museum shown to be guarded 24 hours a day and seven days a week?

If so, how would you arrange this?

If not, why not and how many staff would you require?

2. Draw a polygon using 16 vertices for which 5 guards are necessary to guard every point at any one time.

ID: 1

MAT30029: Individual assignment 2

1.

a) 3 guards (variant 1).

b) asking about 11 guards on staff.

2.

using 16 vertices for which 5 guards are necessary

Mathematical assignment

- 44 students took the coursework on paper.
- Each got a different question sheet.
- Each member of each group got a different diagram (by manual sorting of randomised sheets).

Evaluation

- Was the assessment reliable, valid and did it work to detect uneven contribution?
- In progress.
- Essentially I'm exploring three routes of evaluation:
 - Student feedback;
 - Comparison of different assessments;
 - Second-marker experiment.

Student feedback

- 42 of 44 students completed an end-of-module feedback questionnaire.
- 6 out of 42 students felt minutes of meetings were unhelpful or distracting, with regard to ensuring team members completed their assigned tasks, including one group leader (similar to previous experience).

Student feedback

- Students were fairly critical in 36 responses to a free-text question of the peer assessment of contribution.
- Concern a little over the precise details of the process, also reliability due to personal differences or lack of objectivity.
- “A better system is needed. I do not feel this is a fair way of assessing students to give them their final marks. I agree that low contributing individuals should be penalised however, if a dislike by a couple of people towards to one student, even if they are contributing no less than anyone else, could leave them with lower marks.” (emphasis added).

Student feedback

- 42 students answered some questions about assessment and the partially-automated approach in particular.
- Some example responses follow.

Does plagiarism happen?

Does plagiarism happen?

While at university, I have copied work from other students	Yes	22
	No	19
While at university, other students have copied work from me	Yes	35
	No	17

1 - "Strongly disagree"	2	3	4	5 - "Strongly agree"
"I disliked having different questions because I wanted to work together with another student on our answers."				
12	16	13	1	0

1 - "Strongly disagree"	2	3	4	5 - "Strongly agree"
<p>"I disliked having different questions because I wanted to work together with another student on our answers."</p>				
12	16	13	1	0
<p>"I liked having different questions because it meant I could freely discuss the work with others with no risk of plagiarism."</p>				
0	1	10	22	9

1 - "Strongly disagree"	2	3	4	5 - "Strongly agree"
"I disliked having different questions because I wanted to work together with another student on our answers."				
12	16	13	1	0
"I liked having different questions because it meant I could freely discuss the work with others with no risk of plagiarism."				
0	1	10	22	9
"I liked having different questions because it meant that no one could copy from me."				
0	4	14	17	7

1 - "Strongly disagree"	2	3	4	5 - "Strongly agree"
"I disliked having different questions because I wanted to work together with another student on our answers."				
12	16	13	1	0
"I liked having different questions because it meant I could freely discuss the work with others with no risk of plagiarism."				
0	1	10	22	9
"I liked having different questions because it meant that no one could copy from me."				
0	4	14	17	7
"If we had been set identical questions, members of our group would have copied answers from other students."				
2	5	11	15	9

Second-marker experiment

- The exam board was told, for the module, that “marks were consistent, as was the ranking.”
- I have recruited three second-marker volunteers.
- A 10% sample of student work was anonymised (5 pieces from 44 submitted).
- Second-markers were provided with grade descriptions, a mark scheme and a sample piece of marked work.

Reliability

- Results from two are in.
- Not greatly reassuring on reliability, really.
- What you expect with more subjective questions, particularly marks for the extent to which real world elements were taken into account and marks for presentation.

Validity

- Second-markers were also asked to say what they thought the assessment was assessing, in order to test validity (i.e. do they think it is assessing what I think it is assessing?).
- (Me: real-world problem solving, in-depth problem solving and writing skills.)
- One has basically my three learning outcomes reworded, the other has two of my three.

Comparison with other marks

- Reliability: Students took two effectively similar mathematical assignments, only one of which was partially-automated.
- Comparing standardised versions of these two sets of marks is in progress.
- Initial, woolly findings (of 44 students):
 - 28 students were in more or less the same position relative to the mean on both tests;
 - Nine did a bit better on the individualised test;
 - Seven did a bit better on the non-individualised test.

Comparison with other marks

- Evenness of contribution: Comparing the two individual assignment scores (students v. group average) with the peer assessment of contribution.
- Also in progress.
- Initial, woolly findings (of 44 students):
 - For 15 students, the two measures agree;
 - Otherwise, they disagree, most often with the peer assessment saying even contribution but the assignment marks saying otherwise;
 - In three cases the two say the complete opposite.

Summary

- Individual work alongside a group project to try to improve evenness of contribution and detect uneven contribution, to improve reliability of a group project.
- Partially-automated printed worksheets used as an anti-plagiarism measure - student feedback is positive.
- As for detecting unevenness of contribution - results still coming in, but it seems individual work may be giving more nuanced data than peer assessment of contribution.