The role of partially-automated assessment in assessing individual contribution to group projects

Peter Rowlett

Sheffield Hallam University

`p.rowlett@shu.ac.uk`

Twitter @peterrowlett

13th September 2016

Outline

A long time ago...
My first group work teaching

- ▶ 20 credit, year-long, compulsory second year module for maths students.
- ▶ No pre-, post- or co-requisites.
- ▶ One semester of stats and one of maths.
- ▶ Module specification called for the development of skills (modelling, communication, team-working, etc.), rather than specific syllabus.
- ▶ I ran the maths half, comprising two projects.

Group work and uneven contribution

- Are all students contributing equally, or is the group carrying "passengers" (MacBean, Graham and Sangwin, 2001)?
- "I think that's a bit unfair, every single mark that you get is dependent on other people" (p. 7).
- Hibberd (2002) recommends explicit marks for management of the group and some peer assessment of contribution (p. 168).
- Lowndes and Berry (2003) recommend groups operate a committee structure which keeps "formal minutes clearly indicating actions and progress identifiable with each team member", among other aspects (p. 21).

Peer assessment of contribution

- Fairness and reliability can be a problem. Students are not necessarily qualified to make the necessary judgements (Iannone and Simpson, 2012; p. 13).

- Concerns about efficiency — level of work not directly related to the learning outcomes — for a relatively small project.

- I chose group minutes instead, as a method to track group activity more closely aligned to learning outcome activity.

Students worked

- Details not necessarily important here, except:
  - 5 marks from 90 were for minutes of weekly team meetings.
  - An outline set of minutes were given as an example.
  - These report actions and activities against group members' names, and progress against the project plan.

Minutes

- ► Groups, particularly group chairs, were told minutes were a tool for keeping the work on track and making group members accountable.

- ► The project brief contained:

  If a team member frequently does not attend meetings or the minutes say they are not completing work, that team member may receive a reduced share of the team mark.

- ► Students were invited to bring problems of uneven contribution, evidences by minutes, to me

Minutes

- Groups, particularly group chairs, were told minutes were a tool for keeping the work on track and making group members accountable.

- The project brief contained:

  If a team member frequently does not attend meetings or the minutes say they are not completing work, that team member may receive a reduced share of the team mark.

- Students were invited to bring problems of uneven contribution, evidences by minutes, to me (no group chose to do so).

Feedback questionnaire

► Most students reported finding minutes of meetings helpful in ensuring that team members completed their assigned tasks (two of 26 disagreed).

Feedback questionnaire

- Most students reported finding minutes of meetings helpful in ensuring that team members completed their assigned tasks (two of 26 disagreed).

- However, levels of agreement with 'All team members contributed to the project equally' were mixed – fewer than half agreed.

- The two who disagreed minutes were useful (wording was "unhelpful or distracting") were both group chairs.

Minutes of meetings

- Have some value, but apparently not as a method for detecting uneven contribution.
- It seems likely that marks were awarded to students for work to which they had not contributed evenly (i.e. some of my groups were carrying 'passengers' that the minutes process did not detect).
- Student feedback:

  Always difficult to get even shares of work amongst group and yet difficult to get 'friends' turning on each other (ie challenging ineven [sic] balance of work).

Outline

Some time later...
Another shot at group work

- Again used minutes as a tool for groups to monitor activites, and assigned group management mark from this.
- In addition, used peer assessment of contribution to modify the group mark for individual group members (this time, I wrote the module learning outcomes, so I included outcomes that made this activity more appropriate to use student time on).
- Also included individual work on the same topic as the group work, running in parallel with the group work.

Individual work

- Very similar to group activity (to test same learning outcomes).
- Deadline one week ahead of group deadline, to ensure students get to grip with the mathematical aspects before producing group report.
- Similarity of tasks meant the risk of in-team plagiarism or collusion was high.

Do students copy?

- Final year undergraduates doing this project:
  - "While at university, I have copied work from other students": 22 answered 'Yes' and 19 answered 'No';
  - "While at university, other students have copied work from me": 35 answered 'Yes' and 7 answered 'No'.
- A reference group at another university gave apparently consistent answers to both questions[1].

---

[1]Fisher's Exact Test provides no evidence to reject the null hypothesis that the distribution of answers is independent of group with $p = 0.1513$ and $p = 0.2811$.

Individual work

- ▸ Similarity of tasks meant the risk of in-team plagiarism or collusion was high.
- ▸ Individualised assessment was used (exam conditions and e-assessment were not suitable).
- ▸ Partially-automated assessment meant questions were randomised by computer but marked by hand[2] (using Numbas 'worksheet' template).

---

[2]I claim this method has the potential to make coursework less sensitive to plagiarism while maintaining its validity, in situations where converting to written examination or randomised e-assessment would result in a less valid assessment. But that's another story.

Peer assessment of contribution

- ▶ Members of a group containing n members were asked to confidentially assign $10n$ points among the members of their group (including themselves) to represent the contribution to the project made by each group member[3].
- ▶ The mean of n scores for each student was used to scale that student's group project mark.
- ▶ Taking the mean softened extreme opinions.
- ▶ Self-ratings were checked to be consistent with the rest of the group, no problems were found.

---

[3]I got this idea from Barrie Cooper at the 'Group Work working group meeting', 13th March 2012, University of Bath. No reference as he has not written up the method.

How did it go?

- ► Peer assessment scores were in the range 9–12.
- ► The mean was 9.95 and the median and mode were both 10.

Student feedback

- ▸ Student questionnaire answers re. helpfulness of minutes in ensuring team members completed their tasks was apparently similar to last time[4].

---

[4]Fisher's Exact Test provides no evidence to reject the null hypothesis that the distribution of answers is independent of group with $p = 0.2998$.

Student feedback

- Thirty-six free-text comments about peer assessment of contribution.
- Nine offered non-specific positive encouragement. "if a team member was not contributing enough then normally no one would do anything about it".

Student feedback

- One student said "if a member or so does not contribute it should be down the other members of the group to address this themselves".
- However, my previous experience of group work and the peer assessment scores suggest this does not always happen!

Student feedback

▶ Twelve comments were about the process of assigning points to peers. "This is difficult because even though I wanted to reward others at times, it was very difficult to then remove the marks from another member of the team."

Student feedback

- Four students complained the process was sensitive to personal clashes within groups.
- Three of these had a leadership role in their group.
- "I feel that the peer assessment was marked mainly on how people got on with one another, rather than how much work they did."

Student feedback

- Three said the existence of the peer assessment caused a more even contribution.
  "Without peer assessment, people would not make much of an effort so others would have to put more effort in and everyone would get the same mark".

Student feedback

- Three said it was hard to be objective with friends.
- Two said others in their group might not be fully aware of how much they did.
- One said "I do often feel that people have a tendency to do the 'easy' thing and give everyone equal marks quite often".

Outline

Even and uneven contribution

- We might say that 'even contribution' means all group members have 'pulled their weight' and contributed equally.
- This is likely the sense in which students mean it when they fill in a peer assessment of contribution form.

Even and uneven contribution

- ▸ We might say that 'even contribution' means all group members have 'pulled their weight' and contributed equally.
- ▸ This is likely the sense in which students mean it when they fill in a peer assessment of contribution form.
- ▸ However, it is not usual to award marks at university on the basis of effort alone.
- ▸ Rather, marks should be for assessment criteria aligned to intended learning outcomes, so that the mark provides some measure of how well the student has met these outcomes (Kahn, 2002; p. 102).

An example

- ▶ A piece of work assesses two learning outcomes — one a mathematical topic and the other about report-writing.
- ▶ Two students who share the marks for a piece of work. One agrees to do the maths, the other to write the report. Both report even contribution. Do they each get the full mark for the piece of work?
- ▶ What if the mathematics is poor but the write-up excellent?

Outline

Individualised work

- Can the individualised assignment work to detect uneven contribution?
- This is individual work designed to address the same learning outcomes.
- A measure of how capable students are to contribute evenly to the group project?

Individualised work

- If all students in a group are equally capable and contribute equally, we might expect a correlation between individual coursework marks and group project marks.

- If this correlation is not present, and there is a large dispersion of marks within each group, this would suggest uneven contribution.

Individualised work

▶ Is it reasonable to say that if a student cannot demonstrate ability at some learning outcome in the individual coursework, it is less likely that they are the one who demonstrated that learning outcome in the project report submitted by their group one week later?

Problems

- A student might have made a valuable contribution to the group project but been unable to express this in the individual work for some reason.
- Are the two pieces of work very well aligned?
- Learning outcomes "never fully characterize a student's understanding" (Kahn, p. 93).
- Little in-group dispersion of marks might indicate widespread plagiarism, rather than a harmonious group of evenly-matched individuals.

In practice

- The raw group project marks and rankings do not correlate well with the marks and rankings for the individual assignment (Pearson's $\rho = 0.230$; Kendall's $\tau = 0.229$).

- Individual marks for each group represent a range of at least 23 marks and up to 31 marks (out of 100), and have a standard deviation of at least 8.216 and up to 11.411.

- This suggests even contribution, according to ability measured by the individual assignment, was not present within groups.

Can we use individual marks as a measure of contribution?

- The means of the individual marks for each group were compared with the raw group marks for five groups.
- Correlation: Pearson's $\rho = 0.734$.
- A reference experience comparing two phase tests on similar topics in one module for 74 first year students gave $\rho = 0.700$.
- This suggests a reasonable level of correlation between mean individual marks and group marks.
- Could the difference from the mean individual mark be a measure of individual contribution to the group work?

Problems

- Certainly the detail of how to do this.
- No student scored their group's mean individual coursework mark. Do we say no student made an even contribution?
- Could call everything within some fraction of a standard deviation (or some fraction of the total range of difference from the mean) 'even', and what is above and below this 'uneven', but what fraction?
- I tried to correlate uneven contribution by this measure with the peer assessment of correlation and got poor correlation (highest Kendall's $\tau = 0.213$). But is peer assessment measuring what we want?

To think about

▶ My view of evenness of contribution relates to whether each student has individually met the learning outcomes assessed by the group work, rather than whether all have put in the same effort.

▶ High dispersion of marks within groups for individual work indicates not all students capable of contributing evenly to group work.

▶ Is peer assessment of contribution effective, efficient and accurate for adjusting individual marks in group assignments?

▶ Concerns about peer assessment were expressed by students about objectivity, personal differences and not being fully aware of each other's contribution.

▶ Since group mean marks for the individual work were a reasonable predictor for group project marks, might it be possible to use difference from the mean as a measure of evenness of contribution?

▶ Don't we want uneven contribution, since group members should play to their strengths for the good of the whole (this is good group management)?

▶ Peer assessment was justified as contributing to learning outcomes on reflection, articulation of skills and understanding of how groups operate. But is this the most effective method for these learning outcomes?

References

▶ MacBean, J., Graham, T. & Sangwin, C. (2001). *Guidelines for Introducing Groupwork in Undergraduate Mathematics*. Birmingham, U.K.: Mathematics, Statistics and OR Network.

▶ Hibberd, S. (2002). Mathematical modelling skills. *In*: P. Kahn and J. Kyle (eds.) *Effective Teaching and Learning in Mathematics & its Applications*, pp. 158-174. London, U.K.: Kogan Page.

▶ Lowndes, V. and Berry, S., 2003. Benefits of using Industrially based Group Projects within Mathematics Programmes. *MSOR Connections*, 3(1), pp. 20-22.

▶ Iannone, P. & Simpson, A. (2012). A Survey of Current Assessment Practices. *In*: P. Iannone & A. Simpson (eds.) *Mapping University Mathematics Assessment Practices*, pp. 3-15. Norwich, U.K.: University of East Anglia.

▶ Kahn, P. (2002). Designing courses with a sense of purpose. *In*: P. Kahn and J. Kyle, eds. *Effective Teaching and Learning in Mathematics & its Applications*, pp. 92-105. London, U.K.: Kogan Page.

More from me on this

- Rowlett, P. (2014). Development and evaluation of a partially-automated approach to the assessment of undergraduate mathematics. *In*: S. Pope (ed.). *Proceedings of the 8th British Congress of Mathematics Education*. pp. 295-302. Available via: `http://bsrlm.org.uk/BCME8/BCME8-38.pdf`.