University of Glasgow

Peer Collaboration and Assessment

Sarah Honeychurch, Steve Draper,
University of Glasgow

For the slides, references, access to the students' work, etc. see:
http://www.psy.gla.ac.uk/~steve/talks/fbck1.html

University of Glasgow

Prologue

We have been interested in, and applying, various forms of collaborative learning because of the learning gains they offer.

We sketch some cases of this.

On considering the relationship of these learning activities to assessment and feedback, we have realised that in many cases they achieve the underlying purposes of Assessment and Feedback (A&F) while being at odds with many conventional statements. We will discuss this, using the cases as illustrations.

University of Glasgow

Prologue

We have been interested in, and applying, various forms of collaborative learning because of the learning gains they offer.

We sketch some cases of this.

On considering the relationship of these learning activities to assessment and feedback, we have realised that in many cases they achieve the underlying purposes of Assessment and Feedback (A&F) while being at odds with many conventional statements. We will discuss this, using the cases as illustrations.

University of Glasgow

Part A: Some cases of peer collaboration

- Peer Reviewing in Nursing
- "Jigsaw" Tutorials in Philosophy
- Online Group Work in Psychology (Strathclyde)
- Group Assessment in Psychology
- Using Technology for RPC and Student Generated MCQs (Aropa and Peerwise)
- The Diversity of Critical Voices
- Peter Elbow
- Mahara in Russian Studies
- Patchwork Text

University of Glasgow

Benefits of Peer Techniques

If RPC (reciprocal peer critiquing) is part of the design:
Peers make you think about whether to accept their view or not
Peer's voice often better (closer) to learner's idiolect
You see plenty of peers' work: enlarges your techniques

Given RPC or a huge tutoring staff then:
Plenty of work done early, done regularly, done early in course
Plenty of feedback on the work
Learn procedures (e.g. critiquing) faster by more practice
Learn them better by experiencing the integration of parts into larger wholes (in some Jigsaw related designs)

Group bonding effects: given a shared/common task.

Teaching others: promotes our own learning; and you gain confidence from having your work valued, meaningful.

University of Glasgow

Other reasons for having peer interaction as a big part of L&T

The dramatic reductions in dropout (25% dropout —> 5 %) achieved by some Australian schemes using either peer mentoring, or peer assisted learning.

Learning by teaching is probably the single most powerful learning technique for an individual. Even better if the other is actually learning from it.

Self-efficacy: helping peers gives people confidence in their own worth.

Authentic academic work: seeing your essay used and valued by other students.

Evidence from a puzzle about RPC

Morrow (2006) found strong student attitude support for RPC's benefits, but strongest for being able to see others' work.
I.e. they seem to say that getting feedback on their work is not as useful as simply seeing alternative possible ways of doing it.
That's also what I find repeatedly in oral feedback.
Price et al (2007) found the same.
This doesn't exactly match published theories of feedback.

Students believe it's useful after having experienced the process; and then act on their belief by doing it voluntarily.
But it's not clear how to measure learning gains.

Not least because the gains may only be far in the future and certainly NOT on the current piece of work.

Aronson's Jigsaw classroom

Aronson and his graduate students developed the Jigsaw Classroom learning design, originally for a special purpose: tackling the problems when US schools were forcibly desegregated. How to get the different groups of kids to work together, and stop destructive competition.

Basic answer: Make them depend on each other. Their only access to the knowledge on which their marks depend, is from other kids teaching them. Split the class into groups, each specialising on one part of the curriculum; prepare materials; present.

But this has other good effects. One of the biggest is that the work they produce is of real value to others: whereas normally all student work is artificial, with no end user.

Aronson's Actual Design (2)

The students in a history class, for example, are divided into small groups of five or six students each. Suppose their task is to learn about World War II. In one **jigsaw group**, Sara is responsible for researching Hitler's rise to power in pre-war Germany. Another member of the group, Steven, is assigned to cover concentration camps; Pedro is assigned Britain's role in the war; Melody is to research the contribution of the Soviet Union; Tyrone will handle Japan's entry into the war; Clara will read about the development of the atom bomb. Students are then tested on what they have learned about World War II from their fellow group members.

To increase the chances that each report will be accurate, the students doing the research do not immediately take it back to their jigsaw group. Instead, they meet first in "expert groups" with students who have the identical assignment (one from each **jigsaw group**).

Nursing

Peer assessment exercise on individual 1,000 word essays.

Students started by saying they would give everyone As, but came to realise that this would not be helpful in a largely formative assignment: that giving accurate feedback would help the recipient student more.

Stress the difference for students in evaluating for formative purposes, vs. for a final grade.

Sarah's Philosophy Tutorials

A form of jigsaw.

Students in each tutorial-group assigned to one of 3 subgroups

Each week, the overall topic divided into 3 subtopics

Job of each group is to prepare a presentation which is taught to the larger tutorial group.

Success: attendance; students leave each week with a large set of notes on the whole topic.

Sarah's Philosophy Tutorials: what the students said

"...the notes left are excellent for revision"

"...listening to the interpretations of other students helped deepen understanding of subject materials"

"A majority had things to say that were insightful and useful, unique amongst my tutorials for other classes!"

"When questions were asked, lots of people gave contributions so you have access to a lot of different viewpoints, which is helpful"

Jigsaw Results: What the External Assessor Said

Course marks indicate that the students taught using the jigsaw technique consistently outperformed 'all other students' by more than a 5% margin and had a 93.6% pass rate compared to 84% for the latter.

Based upon 55 student feedback sheets (a good rate of 88.8% feedback) the student experience in SH's tutorials seems to have been extremely positive.

Tony Milligan, University of Aberdeen

Jim Baxter

A 2006-7 redesign of level 1 psychology at Strathclyde university was based, not on student generated content, but on organising groupwork for a giant class (N = 550) mediated by the VLE.
The groups (of about 5) each had their own space, and produced pieces of written work jointly (2 pieces every 3 weeks).
I relied on lessons I'd learned from this.

- Even in a campus university, it is often more practical for students to interact online than F2F.
- Every online action is recorded by the software. This needn't be actively monitored, yet complete records are available if there is a complaint: comprehensive but cheap policing.
- Students seeing others' work is a potent source of feedback (with no staff effort): showing them what is realistically possible, letting them evaluate their own work against others'.

Jim Baxter (2)

He allowed students to swap groups on request: virtual groups means no timetabling or room booking issues, can be done with just a minute's work.

Any complaints about loafers in a group could be addressed by looking at the VLE's records. This requires no active policing in good times, but has complete records in retrospect. Mostly failure to login established absence, but on occasion, skimming the contributions identifies students who are socially interactive but make no substantive contributions.

Also: Gordon Curry's case (Level 4 Geology) of demolishing student claims using the VLE records.

Steve's dodging of fairness complaints

I've run a kind of Jigsaw design in a final year module, where student groups must produce a joint web page for other students to benefit from. Though each student technically gets an individual mark, mostly I mark the joint product and give equal grades.

I require them to post a message or messages on the group's private forum, stating who did what. If they choose to make it look as if all did their fair share, they can't complain about equal grades; if they do complain about a group member, I pay full attention to that. (And the VLE records allow me to decide if complaints are justified.)

Aropa, Peerwise

There is good software available for managing:
1. Reciprocal Peer Critiquing (RPC), where students review other student's work e.g. Aröpa
2. Student-written MCQ test items e.g. PeerWise

The benefits of RPC are considerable. The cost for doing it maually for a group of 6 is already a bit annoying for staff. The big win with the software is to do all the admin. for huge classes, making this learning design widely practicable.

Large numbers also means quality can be self-regulated because students see multiple reviews, or multiple MCQs and see the differences in quality.

Aropa, Peerwise (2)

Diversity of judgements is a key, but generally overlooked, aim for ideal feedback, and almost never satisfied for students, though it is what you yourself can expect when your work is peer-reviewed.

Diversity versus unanimity tells you something about the importance of a point; and the diversity of interpretations of the audience.

This leads on to "Reader's Response" feedback

Peter Elbow

Most of the feedback literature only considers feedback that is authoritative, dictatorial, prescriptive.
However another kind of feedback is non-judgemental and instead phrases everything in terms of the feelings of the reviewer e.g. "I couldn't tell the direction the paper would go in from the introduction"; "I got lost in this section, and failed to follow the arugment"; "the conclusions seemed tentative: surely the results warrant more emphasis?".

This kind of reader's response feedback (see Elbow's books):
- Is much less bruising to the author's emotions
- Is directly about the communicative effect, not questionable "rules" of writing
- Is direct about the different effects one bit of writing can have on different readers instead of pretending that human readers are identical clones.

Peter Elbow (2)

Maths proofs, computer programs, and calculations written in response to "problems" in science and engineering are as much human communications as an essay.

You can make them opaque by being too concise (leaving out steps), or too verbose (can't see the overall shape because too many steps are filled in);
By bad vocabulary (unconventional symbol choices, ….)

Bad punctuation / formatting (running lines together, having no indentation, ….)

Mahara: Russian studies, learning diaries

In a couple of modules, it is a requirement for students to fill in a learning diary once a week throughout the semester. This is to describe both what work has been done (which references read), and the student's thoughts or responses to them.

This is partly to get them working early and therefore more; partly so that they "grow" a personal response that can become a theme of a major piece of coursework.

The latter is a main aim of the Patchwork Text design (next slides); but the diary also achieved major improvements in time management.

It was done in Mahara, and each student made their diary visible to both the teacher and selected peers.

The "Patchwork Text" learning design

(A cousin of Jigsaw)
Every week (say), each student writes a short piece.
AND keeps a private reflective diary
The topic is typically personal: e.g. incidents, feelings, meanings from their own professional practice. Thus each student has the same brief, but quite different material.
Every week the format, genre is different e.g. short story, newspaper article, …. Or lit.review, data report, data analysis ..

They discuss their piece with members of a small group: same group each week. RPC feedback.
Overall aim is to produce a big piece / portfolio by end of term
At half term, re-read one's own patches and diary: look for "emergent" themes, to use as a structure for final portfolio.

Part B: Why "peer e-assessment" is not a good question / issue

Conventional Worries Addressed

How do you know they aren't telling each other false things?
Aronson's expert groups
Web pages that are the start, not end, of learning material
Reviewing reviews in Aröpa; Student ratings of MCQ quality

How do you mark individual contributions to group products fairly?
Cheap retrospective policing from VLE logs
Requiring the group to produce statements of division of labour

How do you get all students to contribute?
Either require them to OR law of large numbers in big classes
Seeing their work be useful (and so "authentic")
Group "pressure" from the logic of shared labour
The Vygotsky effect: 2nd type of motivation
Jigsaw: force mutual dependance

"Constructive Feedback"

Most people construe "constructive feedback" as "telling them what they should have done": a non-constructivist, didactic take on feedback. What would we really wish to happen?

That for each cycle, the learner knows what do do differently next time, and why. I.e. not to change the current product, but the internal generative mechanism that will be used next time.

Peer interaction over understanding concepts to be learned is superior to teacher exposition because it elicits reasons and discussion about reasons (cf "catalytic" assessment).

Diverse peer feedback works the same way: it focusses attention on reasons for doing things one way or another.

When such discussion is about not the content but the learning management layer (who does what), then having reasons means the learners absorb intrinsic goals, not extrinsic ones.

Marks (what peers actually do)

When is an accurate measure (marks in this case) helpful?

When you order a piece of wood from a sawmill, then absolute units are crucial. When you are shaving a bit off a door that is sticking, absolute measurements are irrelevant.

It is true that teaching (or marking) experience is needed for repeatable marks: but that means that the students getting the marks don't understand what the numbers/grades mean because they don't have that experience.

[Dates in history]

When a person, e.g. a learner, reads another's work what are they doing?

Noticing what is the same, what different between this work, and their own.

Perhaps deciding which of the two is better in each respect (difference) detected.

But most of the interest is in the difference.

What's actually important in A&F

- Dialogue
- Diversity of voices
- Reasons/abstractions, not low level directives
- Peer and self assessment (Nicol)
- Showing alternative solutions/approaches
- Increasing understanding of the assessment criteria
- Both "Reader Response" and didactic feedback voices
- Longitudinal feedback: relative to previous feedback to that learner
- Feedback compared to others in the class

Punchlines

Some of the peer collaboration designs are bloody marvellous!

Many conventional staff worries about peer assessment often turn out not to be the show stoppers we feared they were.

The A&F from peer collaboration addresses the **real** requirements better than staff can usually manage.