Peer Assessment in Context

Albert Z. Wang, HILT

January 23, 2014
Outline

What is peer assessment?

Assessment in broader context

Coaches as judges

Assessment for new pedagogy

Peer assessment: An imperfect, but pretty good solution

Conclusion
Section 1

What is peer assessment?
A working definition

- Practice of having students assess each other’s work
- Typically accompanied with *self-assessment*
- In online contexts: a *training* phase where students learn how to grade reliably
- The perspective of this talk: peer assessment as an innovative solution to some thorny issues related to assessment more broadly
Initial motivation

No matter how well you know something, you often learn a lot when you teach it for the first time.

- Learning by teaching, or more narrowly in this case, by *assessing*
- Develop metaknowledge and ability to recognize what makes a good product
Section 2

Assessment in broader context
A thought experiment

Setup
Suppose researchers found, rigorously and unequivocally, that college had absolutely no impact on students.

- Students did not learn, mature, or develop *anything* as a result of going to college.
- If they started the same job right after high school graduation, their life outcomes are all exactly the same, but earlier.

A question
What do you think would happen?
A whimsical game-theoretic response

- Nothing.
- The point of college is to signal that you’re a good guy/gal so that you can land the job.
Why does this matter?

- a lot of emphasis on shifting away from summative to formative assessments
- But parents pay for the **summative** assessments
- Formative assessment is important, but should be kept *independent* of summative assessment
Students often only take summative assessments seriously, so formative assessment *alone* is tricky.

Can be solved with low-stakes assessments that are formative in spirit, but count for some small portion of the grade, or are "binary".

Last but not least, formative assessment is all about feedback, and *feedback is labor-intensive*.

Higher workloads are likely to limit uptake of any new practice, e.g., giving more formative feedback.
Section 3

Coaches as judges
An obvious question

In the Olympics, why aren’t the judges also coaches of at least some participants? After all, judging requires expertise that is often specific to coaches and participants.
And yet, in education

This is exactly the position of most instructors, which can introduce tensions.

➤ They are tasked with **dual roles**
  ➤ Coach: they need to help all students improve
  ➤ Judge: they need to declare which students are the best

➤ This is not in itself a problem, but often both roles are filled by the same assessments
Discriminatory multiple-choice questions

This is a multiple choice statement.

1. True
2. False
3. True declaratively, false perlocutively, because the perlocutionary aspect is false
4. True declaratively, untrue perlocutively, because perlocution is not truthy
5. Untrue performatively, false ontologically
Section 4

Assessment for new pedagogy
A lot of the recent innovation in peer assessment revolves around MOOCs, which require the generation of more feedback than a team of experts alone can manage.

- Coursera (in production)
- EdX/HarvardX (beta testing, Edge platform)
Group learning and team projects

- Group learning can be useful for learning, but individual contributions are hard to assess
- Peer assessments can be used to assign individual grades (Michaelsen and Sweet 2008)
Section 5

Peer assessment: An imperfect, but pretty good solution
Benefits for students

- Offering and receiving diverse feedback often confers direct benefits, especially for more complex skills such as *writing*
- Students develop clarity in the range of possible responses (metaknowledge)
- Helps students to become more effective self-regulated learners (monitoring and clarity in goals)
Benefits for instructors

▶ Unlike many pedagogical innovations, this one actually decreases the variable costs of the instructor
▶ More emphasis on codifying logical and coherent standards in rubrics, less on mechanical execution
▶ Makes evaluation external, so nudges the instructor toward a "coaching" role
Caveats and notes

- Student grading is biased, up for self, down for peer
  - Online systems typically include *calibration*
- The evidence is mixed and narrow:
  - Some have found no benefits to peer-assessment relative to self-assessment (Sadler and Good 2006)
  - Some have found no benefits period (Gielen et al. 2010)
  - Most of the benefit is through giving feedback and not receiving (Lundstrom and Baker 2009)
- While the use of peer assessment has been upheld by the Supreme Court, its use is understandably contentious for high-stakes assessments
In practice

  - Should not coincide with expert assessment
  - Assessment should be reciprocal
  - Small groups of 3-4
  - Where written and oral feedback are possible, they should both be used
  - Purpose of the feedback must be clear
  - When evaluative and formative roles are mixed, resistance is higher
Section 6

Conclusion
What peer assessment can help with

Partially solves some thorny issues:

- Distribution of credit for group projects (systems like WebPA)
- Scalable assessment (e.g., Coursera, HCI resources)
- Scalable rich feedback mechanism
- Clarifies the role of the instructor by offloading evaluation