

The Pedagogical Anatomy of Peer-Assessment: Dissecting a peerScholar Assignment

Steve Joordens², Shakinaz Desa¹ & Dwayne E. Paré²

¹University Putra Malaysia

²University of Toronto Scarborough

ABSTRACT

Peer-assessment is a process wherein students grade the work of their peers rather than relying on an expert. With the recent advent of online tools for peer-assessment [1] it is possible to utilize peer-assessment within any course context, even extremely large courses, with virtually no logistic or economic barriers. Given this potential for widespread use, the present paper highlights the pedagogical value of peer-assessment by breaking down a specific assignment and highlighting the manner in which the learning experienced by students maps onto prominent categorizations of strong pedagogy. We argue that peer-assessment provides a powerful compliment to multiple-choice testing by providing support for the sort of deep, critical and creative learning that is simply not possible to either encourage or assess via multiple-choice.

Keywords: Peer-Assessment, Peer-Scholar, Education, Technology, Critical Thinking, Online Tools for Learning, Multiple-Choice

1. INTRODUCTION

If universities were likened to factories, and one were to ask “what is the product that universities create?”, the appropriate answer would be “Scholars.” But what exactly is a scholar? Dictionaries typically define a scholar as one who is “learned”, or a “master in several disciplines” but seldom do they go further. In his poem “Scholars”, William Butler Yeats (1919) was a little more colourful.

BALD heads forgetful of their sins,
Old, learned, respectable bald heads
Edit and annotate the lines
That young men, tossing on their beds,
Rhymed out in love’s despair
To flatter beauty’s ignorant ear.

All shuffle there; all cough in ink;
All wear the carpet with their shoes;
All think what other people think;
All know the man their neighbour knows.
Lord, what would they say
Did their Catullus walk that way?

Clearly Yeats is not enamoured with the scholars of his day, and his poem may strike a little close to the mark for some of us in the business of producing scholars. Perhaps we do not ask our “young men (and women)” to “rhyme out” what they’ve heard or read in hopes that it will flatter our “ignorant ear(s)”, but it is troubling that to a large extent our assessment methods involve asking our students to simply show us that they have acquired the knowledge we have presented to them. Assignments that ask for more are increasingly infrequent, with full out extinction sometimes seeming destined.

What would Yeats have us do different? He leaves two strong clues. First, he is clearly unhappy with the notion that his scholars “think what other people think”, and he wonders what Catullus would think. Catullus was known for his blunt style of critical analysis, his so-called invectives. He certainly felt no need to think what others think, and his poetic style was new for the time. Thus, Catullus symbolizes independence, critical thinking and creativity; qualities that Yeats feels scholars should possess, but too often do not.

Almost 100 years later, and the worries of Yeats are yet more pronounced. Given increasing student to teacher ratios and reduced funding, universities have relied increasingly on assessment methods that are logistically and economically efficient. Until recently, these measures have not assessed the sorts of cognitive skills symbolized by Catullus. Given that students typically learn only what they need to learn to perform well on the specified form of assessment, the quality of their learning mirrors the quality of their assessment [2]. As a result, the students we produce are less like scholars and more like information databases.

However, to quote Bob Dylan (1964), “The times, they are a changin’”. Through the use of peer-assessment, professors can give assignments that require students to be critical and creative and, thanks to new online tools for administering these assignments (e.g., CPA, peerScholar, SWoRD), they can be graded in a way that is not only logistically and economically efficient but also rich in pedagogy. The goal of this paper is to bring this richness to light by dissecting and describing the components of one specific peer-assessment assignment in the terms of educational theory. To this end, the relevant theoretical frameworks will be briefly presented, followed by the implied autopsy.

2. THEORETICAL FRAMEWORKS FOR DEEP LEARNING

Theorists have attempted to specify the components that go into a rich learning experience, and we wish to highlight 3 specifications here; the first is Bloom’s hierarchy [3], the second is the concept of the 5Es of effective learning [4], and the third is the notion of constructivist learning [5][6]. Together, they provide a strong sense of the sorts of characteristics that would ultimately be embodied by true scholars.

When Bloom [3] provided his “Taxonomy of Education Objectives” he specified three domains of learning; the affective, the psychomotor and the cognitive. The cognitive domain has received the most subsequent attention, giving rise to several modifications over time. The modification we will focus on here is Anderson and Krathwohl’s [7] (see also [8]) categorization as depicted in Figure 1.

Figure 1. Anderson and Krathwohl’s [7] reinterpretation of Bloom’s [3] hierarchy of learning in the cognitive domain.

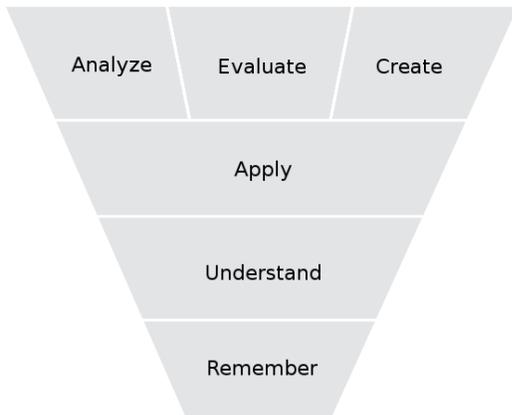


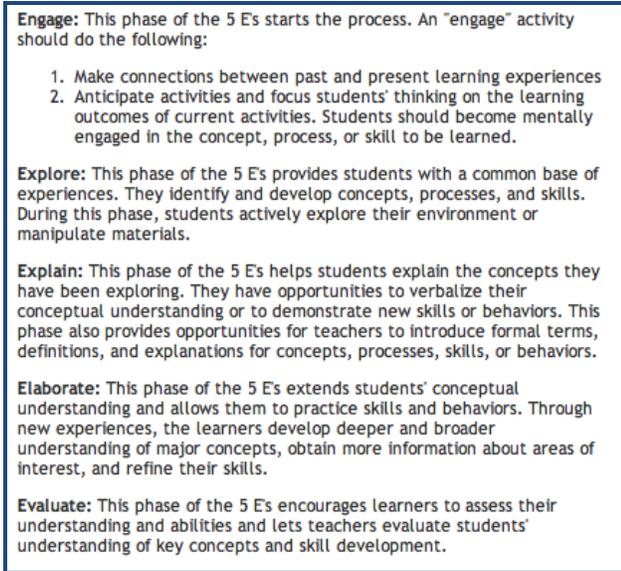
Table 1 provides a brief summary of various cognitive abilities in Anderson and Krathwohl’s [7] categorization. For current purposes, we would like the reader to imagine a fuzzy line that separates the lower three skill sets, Knowledge, Comprehension and Application, from the higher three, Analysis, Synthesis and Evaluation. The lower three essentially ask students to show that they possess some sort of knowledge, and these skill sets could be assessed via well written multiple-choice questions. However the higher three typically require students to think in a critical and creative way, and these skills can be extremely hard, if not impossible, to assess via anything other than open-ended written assignments (see [9] for a discussion of the relative merits of essay versus multiple choice assessment methods).

Table 1. Anderson and Krathwohl’s [7] terminology

Knowledge	Specifics (e.g., terminology, facts), Ways and Means (e.g., conventions, trends, categories, methodologies), Universals, Abstractions (e.g., principles, generalizations, theories)
e.g., What are the health benefits of eating apples?	
Comprehension	Translation (e.g., organizing or remapping of concepts), Interpretation (e.g., giving descriptions), Extrapolation (e.g., comparing, generalizing)
e.g., Compare the benefits of eating apples vs. oranges	
Application	Applying acquired knowledge to new situations
e.g., Which apples are best for making pies, and why?	
Analysis	Breaking information into parts and examining it critically
e.g., List four ways of serving foods made with apples, explaining (and justifying with references) the relative health benefits of each method	
Synthesis	Compiling information together in a different way by combining elements in a new pattern or proposing unique solutions (e.g., creative writing, research proposals)
e.g., Convert an “unhealthy” recipe for apply pie into a “healthy” one by replacing your choice of ingredients. Justify each change in terms of enhanced health benefits	
Evaluation	Present and defend opinions by making judgments about information, the validity of ideas or quality of work based on explicit criteria
e.g., Do you believe that serving apple pie as an after-school snack for children is healthy? Why or why not?	

Also focusing on these higher levels of learning, Bybee et al. [4] proposed their 5E learning cycle where the Es reflect Engage, Explore, Explain, Elaborate and Evaluate respectively. Details of these cognitive skills are presented in Figure 2. Briefly, they imagine the optimal learning experience as one the engages students interest, encourages them to explore the problem space to come to some position, asks them to make their position explicit via explanation, encourages them to then further elaborate on their view perhaps by bringing in new information or perspectives, and finally involves some form of evaluation of the result of this process.

Figure 2. The 5E learning cycle.



Finally, there is a third general notion called constructivist learning [5][6]. Constructivist learning is learning in which the student is asked to “construct” the knowledge necessary to perform some task using the knowledge they already have rather than having the new knowledge given to them by some expert. The notion is that when a learner figures something out for themselves, the learning is deeper than when they are given the information from an external source. Thus, assignments that support constructivist learning are viewed as pedagogically superior.

Together, the top three skill sets of the revised hierarchy [7], the 5E learning cycle [4] and the notion of constructivist learning [5][6] highlight the sorts of skills a true scholar should possess, and the sort of skills that are hard to teach and assess at the university level given the logistic and economic challenges. The following section will briefly describe advances in tools used to support peer-assessment assignments, and how they overcome the logistic and economic barriers. The final section of the paper will then dissect a peer-assessment assignment to illustrate how powerful they can be in terms of supporting the sorts of cognitive skills highlighted here.

3. TECHNOLOGICAL DEVELOPMENTS SUPPORTING PEER ASSESSMENT

Generally speaking, peer-assessment refers to contexts in which students are asked to assess the work of one-another. Thus, in place of an “expert marker”, peers are graded by peers. There are two potential benefits of peer-assessment; first, many of the logistic and economic barriers disappear when a single expert is not expected to mark the work of many, and second, the actual experience of grading the work on one’s peers can have great pedagogical value [10]. The latter benefit will form

the core of the subsequent section of this paper. This section will focus on the logistic and economic benefits.

Initially, those who used peer-assessment tended to do so in the context of the sharing of hardcopies, a process that has its own share of logistic issues. Over the last few years though a number of internet-based applications have been developed that manage the submission, distribution, grading, and feedback of peer-assessment assignments supporting these assignments in virtual any size classroom with only minor cost increments (e.g., CPR, peerScholar, SWoRD). The peerScholar system, a well researched peer-assessment tool developed in our lab, has been shown to be effective even when used with as many as 1500 students, and to provide fair assessments as indicated by both statistical comparisons to expert markers [1] and re-grade request rates which typically are in the 2% range [11]. Thus, when an assignment is given to 1500 students, approximately 30 will need to be re-marked by an expert, and this represents the typical “cost” of employing such assignments.

Given these new systems, the barriers making it difficult to assign students open-ended writing assignments are effectively removed. While the removal of barriers is extremely important, it is their reliance on peer-assessment that makes these tools especially powerful. That is, these systems are not simply solutions to a problem but, rather, they represent a true enhancement in education supported by technological innovation.

4. LET THE AUTOPSY BEGIN

The purpose of this section of the paper is to break down a typical peerScholar assignment and highlight the ways it provides a rich learning environment in terms of Bloom’s taxonomy, the 5Es of learning, and the notion of constructivism. Each peerScholar assignment has 3 phases; a composition phase, an assessment phase, and a feedback phase. We will describe each phase in more detail, emphasizing its pedagogical merits.

The Composition Phase

In one sense peerScholar is a program used to manage and administer online peer-assessment assignments. However, we have used this application in concert with certain readings that are also relevant to the pedagogical merits of the assignments as a whole.

We created our own target articles with the notion of student engagement at the forefront of our minds. There are two factors we considered highly in this regard. The first was identifying some controversial or provocative issue, one that we felt the students would enjoy reading and thinking about. The second was our use of “non-balanced (i.e., one-sided)” and if possible “extremist” arguments related to that issue. These arguments were meant to model a strong, rational presentation of some specific point of view. Thus students would read a

strongly presented argument on controversial issue with the idea that it would engage them, and motivate them. Because we only presented one side of the argument, there remained many counterarguments available. The hope was that the student would identify strongly with one of these counterarguments, and would want to present it in a strong and efficient manner (e.g., <http://psych.utoronto.ca/users/psy3001/files/Joordens.pdf>).

Thus, while virtually any open-ended writing assignment supports some level of critical thought, analysis, evaluation and synthesis, by keying in on the “Engage” E of the 5Es we tried to maximize experience with these cognitive skills. Also, by focussing on an issue we thought the students would find both interesting and provocative, we were also supporting a constructivist use of these skills. The moment students want to make their opinion heard they engage in learning in a much more personal way; one that embraces the learning experience rather than simply accepts it.

The Evaluation Phase

The second phase of a peerScholar assignment is the phase missing from most traditional assignments which, as we will show, is a significant loss. This is perhaps the phase where the most powerful learning occurs, and it is a kind of learning that is both rare in university contexts, and highly valuable in today’s world. Today’s world is one in which all kinds of information are instantly available to anyone who looks. However, all information is not of equal value, and the challenge of today’s world is to sort through the information, discriminating on the basis of quality. This skill, an ability to discriminate on the basis of quality, is useful in virtually every aspect of life, and yet it is one in which we typically do not give students direct experience with in the university context.

In the evaluation phase, students log back into the system wherein they are presented with the compositions as submitted by five or six of their peers. Their subset is randomly selected from the class, and anonymously presented to them, just as their composition is randomly assigned and anonymously presented to five or six of their peers. Students are asked to first examine a rubric, a rubric that was also available to them as they composed their own piece. This rubric roughly lays out the criteria of “quality” where, in our assignments, the weight of that criterion was on the strength of the counter-argument and the clarity and efficiency with which it was expressed. Students are then asked to read each of the compositions, first making a relative judgment of which is best, which second best, etc. They are then to convert this ranking into both a raw mark out of 10, and both a “positive” and a “constructive” comment intended to convey the basis of their chosen mark to the composer.

This task is clearly focused on developing analytic and evaluative skills. However, notice how the middle of the 5Es; exploring, explaining and elaborating, are also clearly at play. That is, in the first phase students came

up with their own counterargument. As they see the work of others they are in fact exploring the problem space, guided by the composer of each space. Seeing these different arguments should give them more information for elaborating on their own perspective. And of course, the need to provide comments forces them to explain why they feel a particular composition is strong or weak, which is a reflection of their learning at that time.

Once again, we see constructivism as a catalyst of sorts. That is, because students are seeing other reactions to the same controversial assignment they just completed themselves – literally a day or two earlier – they should still “care” and feel at least somewhat interested in the views and arguments of their peers. Of course, their own work is being similarly evaluated so they are quite literally embedded in the assignments. In addition, while the rubric provided a general sense of “quality”, that sense is no doubt being augmented by the pieces students see, and their own – constructivist – reactions to what makes a certain composition feel stronger than another. Thus, combined, these factors should make this a learning experience that the student partly controls and definitely feels embedded in. Things are not happening to the students, the students are making them happen. This is constructivist learning at its best.

The Feedback Phase

To be clear, students performing a peerScholar assignment might submit their composition by Monday at midnight, then complete their grading by Thursday at midnight, then log back in on Friday morning to see feedback on their own work. Thus, throughout the week the assignment is fresh in their minds, and the feedback on their work comes quickly, and in the context of the work of their peers. This timeline helps keep students engaged and allows them to better reflect on (i.e., analyze and evaluate) the feedback they are provided.

Although in general feedback phases might seem relatively non-educational, this is not the case with peerScholar assignments for the following reason. Our students are told something along the following lines. When you view the individual grades with comments, understand that the marks and comments come from your peers and that, in some cases, they might contain what we call “noise”. This noise gets averaged out in your overall mark, but remains in the individual marks. Thus, they may be slightly inaccurately or not worded as well as they could be. Thus you should consider each comment and mark while looking at your own work and decide which comments you think will help you in the future, and which may not. In addition, you want to decide whether you feel the average mark you received seems roughly fair. If you feel it is unfair, you can ask to have your composition remarked by an expert marker (e.g., a course Teaching Assistant) who will not see your original marks or comments. If you opt for a remark, your final mark will be whatever mark the expert assigns.

These instructions encourage students to perform analyses, evaluations and critiques, of the comments that are provided, and of their own work. Given the “bottom line” is the mark they receive for the assignment, they should be engaged in these analyses. In fact, this whole process involves them essentially “elaborating” on their assessment of the quality of their own work, ultimately making a relatively high stakes decisions about whether or not to ask for a re-grade. Once again, this is putting them in the driver seat with respect to the analyses and critiques being performed, providing another instance of constructivist learning.

5. CONCLUSIONS

Economic and logistic pressures are pushing universities to rely more on more on multiple choice assessments of learning. This is a problem because multiple-choice tests are not suited to teaching cognitive skills such as critical thinking, analyses based on quality discriminations, and the creation of new perspectives based on a unique synthesis of information. Moreover, multiple-choice tests do not leave much room for experience with the sort of cognitive processes described within the 5E learning cycle, and they do not encourage a constructivist approach to learning. Thus, this migration from open-ended written assessments to multiple-choice tests represents a “shallowing down” of the ways we assess learning and, given that students learn what they need to learn to perform well on the assessments that are used, it also represents a “shallowing down” of learning.

New online peer-assessment programs allow the reintroduction of open-ended written assignments in a manner that is logistically and economically efficient. However, even more important than that is the fact that these assignments provide extremely rich contexts for supporting deep analysis skills. In fact, as the autopsy presented in this paper shows, they provide a far richer context than was provided by traditional “essay type” assignments. Each assignment typically offers multiple opportunities for students to analyze, evaluate, and synthesize. The first phase offers a rich opportunity for the sort of thinking outlined in the 5Es cycle of learning, an opportunity which continues into subsequent phases as students continue to elaborate, evaluate, explore and explain. Moreover, the structure of the assignments, and the readings we use, encourage students to be engaged, and to take a constructivist approach to learning. Perhaps most interesting given the current climate, the assignments also directly teach students to discriminate based on quality, a skill that is critical in this age of information overload, so critical that some think the future of the western civilization depends on it.

Thus, all things considered, peer-assessment assignments seem to represent the perfect complement to well constructed multiple-choice tests. A well constructed multiple-choice task can assess the lower levels of Anderson and Krathwohl’s [7] hierarchy and the peer-

assessment assignments can support the development of the skills reflected in the higher levels. The result would be students who not only possess knowledge, but also the skills to criticize, analyze, synthesis, and create. The sort of student even Yeats might happily call a Scholar.

6. REFERENCES

- [1] D. E. Paré, & S. Joordens (2008). Peering Into Large Lectures: Examining Peer and Expert Mark Agreement Using peerScholar, an Online Peer Assessment Tool”, **Journal of Computer Assisted Learning**, Vol. 24, No. 6, pp 526 – 540
- [2] W. Mohammed, B. Proctor & K. Murshid, K. (2009). Assessment for learning and skills development: The case of large classes. **European Journal of Engineering Education**, 34, 77-85.
- [3] Bloom, D. S. (1956) **Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1: Cognitive Domain**, MacKay, New York.
- [4] R. W., Bybee, C. E., Buchwald, S. Crissman, D. R. Heil, P. J. Kuerbis, C. Matsumoto, J. D. McInerny (1989). **Science and technology education for the elementary years: Frameworks for curriculum and instruction**. Report generated for the National Center for Improving Science Education, Andover, MA.
- [5] T.M. Duffy and D. Cunningham (1996). Constructivism: Implications for the design and delivery instruction. In: D.H. Jonasses, Editor, **The handbook of research on educational communications and technology**, Macmillan, New York, pp. 170–198.
- [6] K. Tobin and D. Tippins (1993). Constructivism as a referent for teaching and learning. In: K. Tobin, Editor, **The practice of constructivism in education**, Lawrence-Erlbaum, Hillsdale, New Jersey, pp. 3–21.
- [7] L. W. Anderson, & D. R. Krathwohl, (2001). **A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives**. New York, USA: Addison-Wesley Longman
- [8] L. W. Anderson, and L. A. Sosniak (1994). **Bloom’s Taxonomy: A Forty Year Retrospective**. University of Chicago Press, Chicago.
- [9] T. A. Ackerman, & P. L. Smith (1988) “A Comparison of the Information Provided by Essay, Multiple-Choice, and Free-Response Writing Tests” **Applied Psychological Measurement**, Vol 12, No. 2, pp 117-128.
- [10] S. Bloxham & A. West (2004) “Understanding the Rules of the Game: Marking Peer Assessment as a Medium for Developing Students’ Conceptions of Assessment. **Assessment & Assessment in Higher Education**, Vol 29, No. 6, pp 721-733.
- [11] S. Joordens & D. E. Paré (2009) “Insuring satisfaction (and fairness) while enhancing education: Adding a re-grade option to peerScholar assignments” [In preparation]