

AC 2007-874: ASSESSMENT IN A PBL CONTEXT: TOWARDS VALIDITY AND RELIABILITY

Charles Mphande, Victoria University

Paul Bronson, Victoria University

Robert ives, Victoria University

Juan SHI, Victoria University

Alec Simcock, Victoria University

Assessment in a PBL Context: Towards Validity and Reliability

Abstract

In 2006, Victoria University (VU) in Melbourne, Australia, started to teach its undergraduate engineering programs using Problem Based Learning (PBL). This was done in response to perceived inadequacies of the weaker graduates produced by a more traditional course structure. The introduction of PBL has allowed the University to address simultaneously the requirements of the Accreditation Authority, and compensate for the problems ensuing from the standards of prior education of students entering these programs. In order to assess the PBL component of the programs within the School of Electrical Engineering, it was decided that a SOLO (Structure of Observed Learning Outcomes) taxonomy should be adopted. Initial experiences of using SOLO have revealed problems, resulting in the need to develop a criterion-valid and interrater reliable basis for assessment. As a case study, this paper considers the use of rubrics as a tool to augment the existing SOLO classification of student performance. As well as guiding students as to what is expected of them, rubrics have provided better results for assessment of PBL deliverables, such as portfolios, as this approach takes into account a broader range of measurable variables. Therefore, rubrics have enabled a mechanism for improving uniformity in all assessment and for bi-directional feedback between staff and students.

Introduction

PBL approaches have been used in many different ways in higher learning to achieve various goals in study programs. In the engineering programs offered within VU's School of Electrical Engineering, the PBL approach has been adopted and adapted to answer increasing demands from accreditation bodies and industry for graduates that:

- have problem solving skills;
- have a practical orientation towards engineering;
- are able to communicate effectively within professional circles and without;
- are attuned to life-long learning ¹.

A consultant employed by the University to consider the use of PBL for the undergraduate engineering problems advised:

“... PBL provides the means to:

- *address more explicitly the essential attributes needed by engineering graduates in professional practice;*
- *enhance pedagogical effectiveness;*
- *tackle at the outset the learning difficulties faced by many commencing students”* ².

As a corollary, assessment of such learning has called for alternatives to the traditional assessments. In the context of VU, traditional assessment has taken the form of: end of semester examinations, semester tests, and assessments of laboratory based exercises, etc.

There is a large body of literature about alternative assessment in schools as well as higher education ^{3,4,5,6}. The reasons for alternative forms of assessment have been varied, ranging from dissatisfaction with the traditional assessment methods to meeting innovative

instructional approaches. In higher education, the drive towards undergraduate programs that approximate professional practice has necessitated redesign of curricula away from the traditional approaches. These innovations in programs of higher learning have brought about unique problems of how to assess student performance. Worthen *et al*⁵ and Mabry⁴ have dealt at length with the subject of alternative assessment, outlining theoretical positions and the range of assessment tools that are employed including drawbacks that are associated with the various tools. However, evaluation of the various alternative assessment tools remains problematic⁵. For instance, in regards to assessing portfolios, a popular means of assessment in higher education innovative programs in Australia and elsewhere, Worthen *et al*⁵ admit that:

Beyond general hints, little helpful counsel is provided on just how to evaluate portfolios. Unfortunately, we have no sudden wisdom in this area either. Just how to score and evaluate portfolios... is a puzzle that awaits solution...

Apart from the use of rubrics, that are widely used with performance assessments^{7,8}, Biggs & Collis⁹ and Biggs³ have suggested the Structure of the Observed Learning Outcome (SOLO) as a way of evaluating the quality of learning (for more details on SOLO, please see section **The SOLO Way of Assessment**). As a way of guiding the learning process, and reliably as well as validly assessing the learning in a PBL context within the School of Electrical Engineering at VU, this paper explores the possibility of a composite approach, combining elements of the SOLO with rubrics.

The subsequent sections of this paper examine the following. The section immediately following exemplifies the challenges of assessment faced by different staff members during the introduction of PBL at VU, and demonstrates difficulties of achieving interrater reliability (i.e, uniformity of assessment across multiple assessors) and validity of authentic assessment. Thereafter, the paper examines how and why SOLO was originally conceived, and subsequent adaptations and refinements of the taxonomy as it has been used in higher education assessment. The paper goes on to examine the merits and demerits of rubrics for learning and assessment. A concurrent comparison of SOLO and rubrics in the context of PBL at VU is included in this section. The final section advances the instructional value of rubrics, incorporating elements of the SOLO taxonomy, as well as their usefulness towards achieving interrater reliability and validity of assessment in the PBL context.

Assessing Portfolios: a Case for Alternative Approaches

Diversity of approaches to PBL is a common feature, largely because of the diversity of quests that are intended to be achieved in the programs. The engineering program at VU in its first year of implementation had learning outcomes designed towards satisfying the graduate attribute requirement of both VU and Engineers Australia, the accrediting authority. Despite potential problems that are associated with portfolios^{5,10}, it has been decided at VU to adopt the portfolio as the sole assessment component for first year Electrical Engineering PBL unit. However, other formative assessment components employed in the course of solving the set problems such as: oral presentations; written technical reports; reflective journals; and records of group work and communication, including meetings, laboratory sessions and use of educational technologies, for example WebCT, are particularly useful for the supervisors to check the truth of learner claims in the portfolio. The portfolio has been considered to be the most suitable tool to assess the technical as well as the more social (generic or transferable) attributes/skills of the professional engineer mainly for three reasons.

Firstly, the portfolio appears to offer an opportunity for cumulating skills towards the competency standard required of graduate engineers. Secondly, because individual students submit their own portfolio, the use of the portfolio enables individual performance to be separated from that of their PBL team. Thirdly, the portfolio allows a delayed, and therefore guided, assessment towards the product at the end of the process. As such, a rough and ready modification of the SOLO taxonomy with its evaluative criteria, quality definitions and a scoring strategy was adopted to assess the quality of portfolios. As PBL is run collaboratively between Engineering, and Language & Communication staff, assessment is also jointly done, necessitating the identification of instruments that would promote interrater reliability. However, weaknesses of reliance on the portfolio and the holistic SOLO modification for assessment of an array of analytic learning outcomes have been noted in the first year of running, as we subsequently point out.

One of the problems is associated with the nature of the SOLO version that was put together to guide assessing the portfolio. While the SOLO is meant to examine learners' quality of work, the VU version has allowed itself to be interpreted to be a quantitative rather than a qualitative assessment. As Worthen *et al*⁵ point out, one problem in assessing by portfolio is time constraint. In this regard, it is evident that the time-constrained staff, in a bid to meet assessment deadlines, were more inclined to adding up the counts of pieces of evidence the learners gave in support of their claims than examining their quality. Indeed, there was little, if any, guidance to the evaluation of quality in the holistic SOLO version in use at VU. Because the staff were intuitively set to examine quality of evidence for each learning outcome, for which there was little guidance, if any, there have been cases of a high degree of subjectivity. As a result, consensus over grading the portfolios has been hard, rendering the interrater reliability and the validity of the assessment unsatisfactory.

Furthermore, supervisors' observations and records of student work/participation had not been built into the assessment component. This caused a dilemma for the staff. These records provide an obvious means of testing the correctness of claims made within the portfolios, and yet their application was not clearly defined. For example, if a learner was noted to be a significant participant in the team performance, but produced a weak portfolio should the rater boost that student's mark and hence assess beyond the portfolio? Varying responses to this dilemma has served as a cause of non-uniformity by the raters.

Another problem with wholly relying on the portfolio without considering other tools (so far used in the formative assessment only) has led not only to the under-rewarding of hard working students, but also the over-rewarding of undeserving learners who may have no qualms resorting to unethical means of getting an improved grade. Learners who are fluent in their written work may draw on the advantage of the group effort, to be better rewarded than they deserve.

A systematic means of incorporating staff observations and other checks needs to be incorporated into the assessment.

Other problems pertain to the learner's perception of the portfolio as a record of their learning and a product for assessment. Although learners are advised how to write their portfolios, there have been cases where claims have not been substantiated, or indeed inappropriate evidence has been given. In other cases portfolios have been submitted with missing claims relating to the learning outcomes. Although such problems may be viewed as incompetence in writing skills, we view them as a show of inadequacies in planning the program's

alignment of objectives, the learning process and its assessment, and failure to make the learner a participant in the same.

It is in light of these assessment problems in VU's recently adopted PBL approach that this paper proposes alternatives both to the use of the current modified SOLO, and the sole reliance on the portfolio. To contextualise these suggestions, we briefly examine the SOLO way of assessment as a backdrop to the proposal to adopt a composite approach, namely, rubrics informed by elements of the SOLO taxonomy.

The SOLO Way of Assessment

Biggs & Collis⁹ came up with the Structure of Observed Learning Outcomes (SOLO) partly as a way to address what they perceived as misplacement of assessment focus on descriptions of people as opposed to their responses or performances in the Piagetian tradition of stage theory. Teachers, they observed, were rather ill equipped with the Piagetian stage theory when learners' performances did not match their cognitive developmental stages. As it is more the rule than the exception that learner performances do not consistently correspond to the cognitive stages, and that there may be many factors, for instance, "motivation, prior knowledge of that particular task"⁹ that can contribute to this non-correspondence of learners' performance and their cognitive developmental stages, Biggs & Collis⁹ observed that teachers were short of explanation when their focus was on the learner's cognitive 'age'. Biggs and Collis's⁹ studies showed that assumptions of the stage theory did not hold in this regard. It was in light of this inadequacy of the stage theory that Biggs & Collis⁹ proposed the SOLO to shift the assessor's attention from learners' developmental stage, a "*purely hypothetical concept...not directly measurable*" (p22), to "*the structure of the actual responses that she gives to specific learning tasks*" (p22). Biggs & Collis⁹ concluded that:

It is because of this confusion between stages of development (about which we know very little) and particular test results (about which we know a good deal) that we have introduced the SOLO terminology (p23).

In keeping with their innovation, they came up with their terminology to describe the SOLO taxonomy they developed away from that of the Piagetian stage theory but largely parallel to its concepts.

It should be pointed out at this point that the SOLO was devised principally within the traditional teacher dependent learning/teaching concepts of education. Even though Biggs³ is a refinement of the SOLO for use in university teaching and learning contexts, it remains a taxonomy that is largely conceived for the traditional teacher-centred, subject-centred learning approaches. PBL learner/participant directed learning circumstances are radically different; thus, adoption of SOLO for use in PBL contexts will pose challenges. Hence our need for continuous revision of adaptations. For purposes of this paper, there are pertinent SOLO elements in Biggs³ that hold some potential for augmenting the learning process and its assessment, which we propose for PBL as practised at VU. The elements are part of the hierarchy of the quality of learning ranging from a quantitative phase, encompassing the prestructural, unistructural and multistructural stages, to a qualitative phase, encompassing the relational and extended abstract stages. Descriptions of the hierarchies may be useful in the criterion-referenced assessment of learning. It is our view that this hierarchy of quality of learning may be productively used in formulation of rubrics that can be more readily aligned to the problems and their stipulated deliverables, currently used as formative assessment

tools, to guide the learning process and its assessment. Thus, in the following section we argue that while SOLO enlightens us about student cognitive development, rubrics would enable us to communicate to students and supervisors expectations built into problems. As much as the rubrics would clarify goals and gradations of learning outcomes, they may be effective tools for minimising rater subjectivity.

A Case for a Composite Evaluative Tool: Analytic Rubrics Informed by SOLO

The abundant literature on rubrics (too many to list here) attests to their popularity as learning and assessment tools both in schools and higher education. In regards to this upsurge of interest, Popham¹⁰ points out *“Rubrics are all the rage these days. It is difficult to attend an educational conference without running into relentless support for the educational payoffs of rubrics.”* Although the literature points out flaws in the use of rubrics^{10, 8} there is overwhelming support in the literature for use of rubrics as learning and assessment tools in higher education where the nature of skills that are being learned is not only comprehensive but also multifaceted¹¹. It is interesting to note that changes in engineering education that have partly influenced the introduction of the PBL Engineering program at VU are not limited to Australia. Lattuca, Terenzini, Volkwein & Peterson¹² report that *“Employers [in the US] complained that new hires had poor communication and teamwork skills and did not appreciate the social and nontechnical influences on engineering solutions and quality processes...Several national reports recommend.. changes in engineering education...”* The issues that need addressing are obviously multifaceted in that they are both social and technical. Inclusion of elements that have not been part of engineering education traditionally points to the importance of the interdisciplinary nature of engineering education at the turn of the 21st century. Thus, PBL supervisors are likely to be found lacking in some of the non-technical skills now demanded of their learners. It is therefore important, in the re-skilling process of the staff to adopt learning and assessment tools that would constructively align with the PBL learning approach and promote greater uniformity in the assessment of the skills that now are considered crucial to engineering graduates.

It is our view that the multifaceted nature of the technical and social skills that are intended for the PBL Engineering graduates at VU (and elsewhere) would benefit from the use of analytic rubrics that would be founded on an understanding of qualities of cognitive performance that the SOLO taxonomy offers. Mertler⁷ points out that analytic rubrics are useful for assessment on a multidimensional level, a feature of the PBL problems. Furthermore, it is known that the use of analytic rubrics enables provision of specific feedback to learners on their performance with respect to each of the individual scoring criteria (learning outcomes) – something that does not happen when using holistic rubrics¹³. Another critical point for our advancement of SOLO based analytic rubrics is the diversity in a broad sense of the student population that VU attracts at home and abroad¹⁴. As PBL poses new challenges in addition to linguistic and socio-cultural ones, particularly to international students, a more guided approach that rubrics would afford may prove more learner friendly. In this way, it is possible to avert learner attrition, and indeed make the learning experience more enjoyable and meaningful than otherwise. Another paper¹⁵ discussed several problems that PBL as a learning approach posed to VU Engineering students in its first year of running. Of particular note was the teacher dependence that learners showed. Although this may be attributed to many factors, such as learning cultures, it came out clearly from the study that significant trepidation was experienced as learners were introduced to practices leading to learner autonomy. This was largely a result of inadequate or the absence of qualitative guidance as to what was expected of the learners in regards to deliverables they were required

to submit at particular stages of the problem solving. We posit that there is much to be gained in adopting rubrics informed by SOLO hierarchies of cognitive performance qualities in attempts to nurture the skills that PBL intends learners to attain.

Conclusion

The paper has described the purposes of the PBL Engineering program at VU, namely, imparting technical and social skills including life-long learning to graduate engineers, and the adopted means of assessment of the learning process, the portfolio. The paper has also pointed out limitations of the portfolio's original evaluative tool, namely, an adaptation of the SOLO taxonomy. The results of such limitations in achieving interrater reliability have also been highlighted. To address inadequate or absence of alignment of learning goals and assessment of quality of learning, and interrater unreliability, rubrics informed by SOLO's hierarchy of cognitive performance have been proposed as an enhancement. It is also envisaged that the adoption of rubrics in place of the current adaptation of SOLO would be more learner friendly and avert learner attrition, especially among the international student cohort. The paper has also identified possible pitfalls in the use of the portfolio as the sole assessment tool. It has instead suggested consideration of upgrading some assessment components that are currently only used for formative assessment, to provide checks against under-rewarding deserving students and indeed over-rewarding less deserving learners. The paper acknowledges the fact that the PBL Engineering program at VU is only in its infancy, facing an array of challenges and a reflexive approach to its implementation is a sure way forward, as it would be for any other institution reaching for provision of quality higher education.

References

1. Simcock, A., 2006, "A Comparison of the PBL Models in the Schools of ACME and EE at Victoria University," Proceedings of the 17th Annual Conference of the Australasian Association for Engineering Education, Auckland, New Zealand, 10 – 13 December.
2. Parr, P., 2005, "Proposal to Introduce a Problem Based Learning Curriculum to the Bachelor of Engineering Degree Courses Commencing with First Year in 2006, Overview and Background Briefing," Submission to Higher Education Course Approvals Committee, Victoria University.
3. Biggs, J., 1999, *Teaching for Quality: What the Student Does*. Philadelphia, PA: SRHE & Open University Press.
4. Mabry, L., 1999, *Portfolios Plus: A critical Guide to Alternative Assessment*. Thousand Oaks: CA: Corwin Press, Inc.
5. Worthen, B. R., Borg, W.R. & White, K.R., 1993, *Measurement and Evaluation in the Schools*. New York: Longman.
6. Vavrus, L., 1990, Put portfolios to the test. *Instructor*, Vol. 100, No. 1, pp 48-53.
7. Mertler, C. A., 2001, *Using Performance Assessment in your Classroom*. Unpublished manuscript, Bowling Green State University.
8. Andrade, H. G., 2005, Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, Vol. 53, No 1, pp 27-30.

9. Biggs, J. B., & Collis, K. F., 1982, *Evaluating the Quality of Learning: The SOLO Taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.
10. Popham, W.J., 1997, What's wrong – and what's right – with rubrics. *Educational Leadership*. Vol. 55, pp 72-75.
11. Simon, M. & Forgette-Giroux, R., 2001, A rubric for scoring postsecondary academic skills. *Practical Assessment, Research & Evaluation*, 7(8). Retrieved May 11, 2006 from <http://PAREonline.net/getvn.asp?v=7&n=18>.
12. Lattuca, L. R., Terenzini, P. T., Volkwein, J. F. & Peterson, G.D., 2005, The changing face of engineering education. Retrieved 7th August, 2006 from <http://www.nae.edu/NAE/bridgecom.nsf/weblinks/MKEZ-6QDKCV?OpenDocument>.
13. Nitko, A. J., 2001, *Educational Assessment of Students* 3rd ed. Upper Saddle River, NJ: Merrill.
14. Keating, S. & Gabb, R., 2006, PBL in engineering: Student expectations in 2006, Postcompulsory Education Centre, Victoria University.
15. Shi, J., Mphande, C., Simcock, A., Ives, R. & Bronson, P., 2006, *Communication in a Problem Based Learning Environment: Supporting the Teaching Team in the School of Electrical Engineering*. Proceedings of the 17th Annual Conference of the Australasian Association for Engineering Education, 10 - 13 December, Auckland, New Zealand.