E-portfolios and cross-cultural assessment

Johannes C Cronjé
University of Pretoria
jcronje@up.ac.za

Abstract

It is difficult to collect evidence of adequate performance if an assessor and the assessees have different linguistic and cultural backgrounds. This paper considers three entries in student portfolios where the assessor tried to understand the students and their achievement, rather than to rate it. Three approaches were followed: positivist, post-structural /interpretive and feminist/holistic. Sources of data include artifacts produced by students for inclusion in the portfolio, assessment rubrics and assessment scores. Results indicate that portfolio assessment with criteria negotiated between the instructor/assessor, the learners and their peers allow for a rich thick evaluation that adds to the development of the learners. Detailed assessment rubrics stifled student creativity, while open-ended tasks with video feedback enhanced it. Further research needs to be conducted into standardizing the procedures of auditing assessments, rather than standardizing procedures of assessing.

1. Introduction

Digital environments make it much easier to expect students to engage in a variety of learning tasks that would have been hard to produce in a traditional environment. Instead of sticking to essays, instructors can request learners to produce slideshows, movies, spreadsheets, and even websites as evidence of their learning. The variety of possibilities makes portfolio assessment a lucrative alternative. However, while much has been written about traditional assessment by tests and essays, there are still "many gaps in the research on innovative assessment methods such as portfolios" [1]

While the Internet allows us to teach internationally and across cultural and linguistic barriers, new challenges as is pointed out by. MacKinnon & Manathunga: "If our assessment continues to be based upon a Western template of knowledge that only values Western ways of knowing and learning, all our lip service to developing interculturally competent students is meaningless"[2] .

For Delandshere, assessment is often used as a "technology" to exclude those whom the assessing authority deems unsuitable. "While there have been important technical developments (e.g. item response theory, item-biases) and technological changes (e.g. computer-adaptive testing), the procedures used have remained fundamentally the same with tests being the primary method of educational assessment"[3]. Cross-culturally test bias has been identified as a major contributor to unfair exclusion. [4] If tests are exclusive and possibly inaccurate, alternatives should be considered.

2. Background

In 2003, and 2004 three professors from the University of Pretoria presented the coursework for a Masters' degree in Information and Communication Technologies (ICT) for Education, to twelve selected, students of the Sudan University of Science and Technology (SUST) Khartoum. One professor would visit Khartoum for one week and facilitate workshops with the students. Then would follow eight weeks of independent work by the students supported over the Internet. In the follow-up visit to Khartoum students were assessed on the work of the past eight weeks briefed on the work for the next course.

3. Research problem

The main question was "How do we assess fairly a group of students with whom we do not share a common language nor a common culture?" The students could understand us reasonably well, could read English, and support each other in reading and interpreting English articles, but were shy to speak and their writing was poor. Thus, our assessment of their learning was obscured by lack of English language skills. Culturally there was the suspicion that we were foreign colonialists coming to use assessment as a "technology" [3] to enforce our ways upon them. We

had to establish culturally sensitive interpretations leading to common expectations.

4. Theoretical underpinning

The cultural divide between South Africa and Sudan is shown by four of Hofstede's [5] cultural dimensions. Table 1 shows the two countries at opposite ends of the continuum.

Table 1: Hofstede's (2000, p.xix – xx) dimensions of culture for Sudan and South Africa

Dimension	Sudan	South Africa
Power distance	80	49
Individualism vs. Collectivism	38	65
Masculinity vs. femininity	53	63
Uncertainty avoidance	68	49

Hofstede's model, however does not help us reach commonality. The more global something becomes, the less its local relevance, the more local the less globally competitive or useful it will be. Any emphasis on increasing either local relevance or global competitiveness would increase the divide. Amin identifies six "factors which students [at a bilingual university in Cameroon] consider important in the evaluation of courses" *Rapport with students; Evaluation of Assignments and feedback; Workload; Availability of Help; Course Organisation; and Language or Communications* [6]. We argue that these factors are common across all cultures.

5. Research approach and methodology

Portfolio assessment was selected because "the rich picture of teaching reality provided by a portfolio tends to make this form of assessment highly valid, because it includes longitudinal evidence from multiple sources" [7].
Johnston [1] mentions that "most assessment in practice takes place in a middle ground between the different approaches" to assessment. She calls for more "longitudinal or action research" [1] into portfolio assessment. Our action research was a disciplined self-reflective inquiry into a small scale, real-world intervention in an attempt to reform and improve practice. It strives to bridge the gap between theory and practice, and is done by a practitioner focusing on his own work, combining diagnosis with reflection and focusing on practical issues. Principal data sources are

the artifacts produced during selected courses of the programme, evaluation instruments and results. Interpretation followed a qualitative formalistic scrutiny of the data to identify trends and draw conclusions.

6. Discussion

Three cases were selected from the beginning, middle and end of the course to show how we progressed in our thinking about creating and assessing electronic portfolios to reach common understanding.

6.1. Case one: A mark out of 200

The first assignment for a course on evaluation of educational software was a dynamic evaluation of a piece of educational software for a target population of the students' choice.

Discussions with them showed that they were uncertain of the field and unfamiliar with the instructor. Thus, a positivist approach was followed. I created a report template with typical headings and a few useful phrases, a few citations in the text and bibliography to show how to cite and refer; and a rubric with 54 assessment statements adding up to 200 points.

Each student would select a title and target population for evaluation and submit the report to two other members of the class who would rate it according to the rubric and give feedback. The student improved the essay, self-evaluated it with the rubric and submit the essay. Table 2 compares my assessment to the self-assessment.

Table 2: Student self-assessment compared to my assessment

Student	Own grade	My grade	Difference
Md	83	86	-3
Yo	91	86	5
Et	89	85	4
Ta	68	82	-14
Om	70	82	-12
Id	76	76	0
Ha	75	75	0
Ib	70	73	-3
Ma	74	69	5
Ab	70	68	2
Ia	70	67	3
Ka	53	51	2

In only two cases did the students' grade differ from mine by more than 5%. They were two of the

better students, an interview showed that their underrating was due to lack of self-confidence.

Ka's grade was so low because he had missed the initial lectures.

The essays were disconcertingly close adaptations of the template, written to the rubric. They were good, but dull, echoing Driessen, van der Vleuten, Schuwirth, van Tartwijk, & Vermunt,'s claim that "detailed checklists can easily trivialise assessment" [8].

6.2. Case two: Spreading the sheets

With the next course. Students worked in "jig saw" groups of three. Home groups produced spreadsheets with statements and a yes/no checkbox to determine how a learning need would best be served by a tutorial, drill, simulation or game. Expert groups formulated questions, created the look and programmed the checkboxes. Once the expert groups had brought together the statements for each checklist, they negotiated in the home groups to determine the values the spreadsheet assigned. Finally the home groups worked through one another's efforts by way of summative and peer evaluation. Figure 1 shows a screenshot of a completed project.

Again the projects were very similar as students had worked in cooperative learning groups, from the same textbook content.
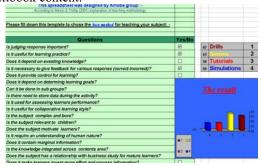

Figure 1: Screen capture of a spreadsheet task

The spreadsheet exercise showed how the home groups and expert groups negotiated meaning, and how the role of assessor had shifted. Instead of the instructor asking *Is there enough evidence that the students have mastered the content?* the students were now asking *Do we have enough knowledge to complete the task?*

6.3. Case three: At the movies

A later course required students to follow Gagne's (essentially objectivist) events of instruction to prepare and present a *constructivist* "lesson" with computer-

based support materials emphasizing the learners' construction of their own meaning. The final evaluation included a 10-minute video presentation by the students explaining what they did and what they learnt.

Students found it much less stressful to speak about the video playing behind them. An analysis of the videos showed they could identify significant events of the lesson; they could tell a story visually, perform the basics of video making.

7. Conclusions and recommendations

This article shows how we reached commonality in assessing the quality of learning across linguistic and cultural boundaries and supports Amin's cross-cultural factors of *Rapport with students; Evaluation of Assignments and feedback; Workload; Availability of Help; Course Organisation; and Language or Communications* [6].

8. References

[1] B. Johnston Summative assessment of portfolios: an examination of different approaches to agreement over outcomes *Studies in Higher Education* 29(3) 2004 pp. 395-412

[2] D. MacKinnon & C. Manathunga, "Going Global with Assessment: What to do when the dominant culture's literacy drives assessment". *Higher Education Research & Development, 22, (2)* 2003 pp. 131 – 143

[3] Delandshere, G. Implicit theories, unexamined assumptions and the status quo of educational assessment, *Assessment in Education*, 8, 2001, pp.113–133.

[4] R.J. Skiba, K. Knesting, & D.L. Bush,. "Culturally Competent Assessment: More Than Nonbiased Tests" *Journal of Child and Family Studies, 11*(1), 2002 pp. 61–78.

[5] G. Hofstede, *Culture's consequences : comparing values, behaviors, institutions and organizations across nations.* Thousand Oaks, Calif.: Sage Publications. 2001.

[6] M.E. Amin, "Six Factors of Course and Teaching Evaluation in a Bilingual University in Central Africa". *Assessment & Evaluation in Higher Education, 27*(3). 2002, pp. 281 – 291.

[7] D.E.H. Tigelaar,. D.H.J.M. Dolmans, I.H.A.P. Wolfhagen, & C.P.M. van der Vleuten, "Quality issues in judging portfolios: implications for organizing teaching portfolio assessment procedures" *Studies in Higher Education 30*(5), 2005 .pp. 595–610.

[8] E. Driessen, C. van der Vleuten, L. Schuwirth, J. van Tartwijk, & J Vermunt, J. "The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study". *Medical Education 39*: 2005, pp. 214–220.