

Tension and Conflict in Assessment

Viola Wong Yuk-Yue

Hong Kong Polytechnic University

Abstract

The paper aims at bringing out some of the intricacies and delicate issues related to language assessment in a discussion that places language learning in an educational context. The paper will include a study that has been carried out on the learning of English as a second language and the means of assessing English language proficiency among a group of first-year university Chinese students. Data were collected through interviews. The study highlights relationships among the espoused aims of the institution, the philosophy of the teaching and the aspiration of students as well as the resultant interactive forces that have given momentum in the search for an appropriate mechanism for assessment. In the discussion, two key management concepts are explored: 'accountability' and 'responsibility'. The two concepts are examined with a view to clarifying the role of language educators in a specific context, such as the one in the study of the paper. Although it is language learning oriented, the paper has wider implications for assessment in education.

Introduction

Student grades and products of student learning are evidence of teaching outcomes, which are of concern not just to the students, but also to individual classroom teachers, the course designers and also the educational organization. Grades awarded to students indicate their levels of language proficiency in a programme. To teachers and course designers, the students' success rate of completing the programme can be a means of expressing their effectiveness in their teaching and designing of the courses. To an educational organization, students' performance in the programme reflects the language ability that the graduates possess. It is clear that while the teachers and course designers are expected to take up the responsibility to help students meet with the requirements of the programme, they are also expected to account to the organization (and, in fact, to society at large) for their pedagogical practices. Responsibility to students and accountability to outsiders (whoever is outside the organization) are of utmost concern to all who are involved in assessment and in instituting standards against which to evaluate students' work. It is not easy to accomplish both in a programme as the road to achieving them is often

assessment give rise to such conflicts and tensions, this paper discusses the assessment of a university English course. Three areas that are related to students' performances in the assessment have also been highlighted for discussion: language and capability, language and thought, language and self, pointing out some issues that could have significant bearing on language education.

Managing assessment

One of the great challenges faced by most teachers is to make a decision on the way(s) to assess students' learning. To evaluate different ways to assess students' performance in a subject, one needs to be clear about the assessment concept, assessment purposes, assessment accountability and responsibility, as well previous studies regarding assessment of students' performance in that particular subject.

The assessment concept

According to Craft (1992), there are three elements in assessment: judging about the performance, translation of the judgement into a grade point or numerical system, and the application of authority to legitimize the judgement, and that the assessment standardization is formal, open and public. It is formal because it makes explicit the attributes and values in the judgement as to the adequacy of performance; it is open and public because it accounts for the decisions made about the students' performance. Equity and fairness are assumed to be inherent in a healthy and sound assessment system, and the assessment itself is seen as a process of change and renewal.

Assessment purposes

Assessments of students' performance can be made for various reasons, such as maintaining students' standards, making selection, providing feedback and/or motivation to students, providing feedback to the teacher, etc. (Rowntree, 1987). The teacher's handling of assessment can be influenced by the expectations of the curriculum or syllabus designers, as well as his/her attitudes towards assessments, and his/her philosophy of teaching and learning in general. In other words, there are some teachers who are keen on getting across a body of knowledge to students (for example, grammar rules in the language); and there are some who believe that individual students should be encouraged to make their own meaning and to create new knowledge out of their own ideas and experiences (for example, using the language for self expression and communication in which accuracy of the language

used is not prominent in the learning process). The teacher's belief and assumption could affect students' perceptions of the assessments in a programme. The assessment may be seen as an objective means of determining student's present achievement and future potential, in which case the assessments are usually for selection purposes in cases like university/ school admission, or it may be regarded as a means of reflecting students' strong and weak points for developmental purposes in a programme. Since assessments play such an important role, it easily generates plenty of discussions not just in the educational context but also in society at large. Parents, employers, mass media, government officials, educators and students have their views on assessments.

Accountability and responsibility in assessment

'Accountability' may mean different things to different people (such as parents, employers, government officials) and for the sake of giving a clear focus in this paper, Kogan's definition is used. To Kogan (1986, p.26), 'accountability' has a legal overtone, whereas 'responsibility' has only a moral obligation, 'Accountability assumes institutional authority to call an individual or a group to account for their actions. It is to be contrasted with 'responsibility' which is the moral sense of duty to perform appropriately. Responsibility need not evoke the duty to answer in a legal or contractual setting, that is, to act accountably'.

When discussing accountability, Norton (1997) mentions the importance of 'addressing the tension between accountability to individuals on the one hand and accountability to systems on the other' (p. 318) and suggests the importance of engaging in 'a dialogue with a broad range of stakeholders so that teaching and learning can be enhanced' (p. 319), citing assessment research carried out in the nineties by scholars such as Lacelle-Peterson & Rivera Peirce and Shohamy. Norton (1997) also quotes research in the same period to point out the increasing importance of academics and teaching professionals to shoulder the responsibility of explaining their work to laymen, 'informing the public about what they are teaching and how effective they are' (p. 317).

In an educational institute, the teaching staff are to be held accountable for the students' performance, making sure that their standards meet the expectations of the stakeholders (among whom may be the institution, the prospective employers, the parents, etc.); while at the same time they are also responsible for providing students with access to quality learning experience. To illustrate some of the problems faced by the teachers in such an endeavour, this paper discusses the challenges that arise

from assessment in a programme which teaches English as a second language in higher education.

Managing assessment for English as a second language (ESL)

Second language teaching professionals have built up a vast body of both theoretical and practical knowledge on designing and delivering language courses in the past twenty years. Different approaches to language teaching ranging from the Audiolingual to Total Physical Response, from Suggestopedia to Communicative approach have been practised by teaching professionals. There have been syllabuses such as the notional-functional syllabus, the process syllabus and the procedural syllabus. All this has pointed to a great variety of schools of thought for second language teaching and underscored the great flexibility that language course designers and teachers are applying in their courses to match their own specific situations. A designer and teacher's ideas about what and how to teach and learn (i.e. his professional world view or his pedagogic paradigm) may result in different positioning along an assessment continuum. The opposite ends of this continuum are characterized by providing valid information about the students to outside parties versus developing relationship between the students, the teachers and the subject matter. Much discussion of ways to achieve fairness and equity can be expected.

Hamp-Lyons (1997), in discussing ethics in language testing, has also drawn attention to the concepts of fairness (which highlights the proper use of tests, as well as the promotion of educationally relevant assessments) and of equity (which underscores the importance of student access to the resources, both human and material, that are crucial to their effective learning). While previous ESL assessment literature (e.g. Hamp-Lyons (Ed.) 1991; Hill & Parry (Ed.) 1994; McNamara 1996) was focusing more on valid and reliable assessments from the test designer or teacher's perspective, recent topics for assessment discussion and research has widened the scope to include issues that are related to ethics, social justice and cultural significance (Hamp-Lyons 1997; Hafernik, Messerschmitt & Vandrick 2002), fronting an obligation to appreciate, respect and help the learners in the ESL learning process. It is the argument of the writer of this paper that a discussion of accountability and responsibility issues in assessment will not be complete without taking into account the views of those who have been most affected, i.e. the students. The students' experience and comments could shed light on the students' perception of the general success of the assessment, or the effectiveness of the programme for that matter, making valuable contributions to the monitoring of the assessment process of the programme.

Study context

The language proficiency assessment that is of concern in this paper is one that is in use in a programme entitled ‘English for Academic Purposes’ or EAP run by an English centre at a university in China. (See Appendix I for the syllabus.) The programme is a mandatory English course for the first-year students at the university who learned English as a second language. There is a general belief at the university that most first-year students need help with their studies, especially English, in their transition from secondary schooling to their first year education at the university.

The discussion of this paper focuses on only one of the three assessments in the course -- an argumentative essay, which is an in-class assignment. (See Appendix II for the assignment.) Student writing skills are assessed against the band description distributed by the Centre (see Appendix III). Criterion reference is adopted and stringent measures to ensure equity and fairness in the essay assessment are taken (such as designing a web page for on-line assessment training, double or even triple marking students’ writing scripts by teachers other than the students’ own classroom teachers if the scripts have posed problems in grading or marking). The assessment has been designed to aim at reflecting the students’ ability to handle academic writing and the categories of criteria against which students’ writings are judged include ‘content’, ‘organization, cohesion and coherence’, ‘register’, ‘grammatical structures and vocabulary’, ‘range of grammatical structures and vocabulary’ and ‘academic writing conventions’. The argumentative essay is included in assessment because it is believed that the language skills and thinking skills that are involved in discursive writing are fundamental to the pursuit of university education and academic writing. In order to understand the extent to which the assessment has effectively achieved what it set out to achieve, the following questions were raised in the study:

- (1) Was the assessment reflecting students’ capability in handling writing at the university?
- (2) Was the assessment task reflecting students’ cognitive skills?
- (3) Was the assessment language requirement reflecting the students’ ability to express in an academic context?

Approach of investigation

In order to understand the assessment issue from the students' perspective in the broad framework of learning, semi-structured interviews were held with 10 students who were divided into groups of 3 or 4. (See Appendix IV for some of the questions covered and generated in the interviews that are related to the assessment task in this paper.) Each interview lasted for about one hour. For the present study, the students' scripts and their daily in-class as well as out-of-class writings that were available for investigation were also consulted.

Students' voices in the investigation

Importance of writing at university

While most in the academic community would agree with Fulwiler (2002) that academic literacy is important because it helps them 'to discover, to communicate, and to create', first-year university students might not be able to appreciate all this importance of academic literacy. When asked in the interviews the purposes for writing in English on the campus, the students gave the following in their replies:

- Writing to meet academic requirements (to complete assignments, to answer examination questions)
- Writing to impress (to score credits) and to display knowledge
- Writing to learn the language or to prove the ability to learn in English
- Writing to make notes (copy from books, teachers' notes at lectures or tutorials)
- Writing to communicate with professors/ lecturers through e-mail

Factors affecting willingness or readiness to engage in the assessment task

(i) Experience of learning how to write

Some students mentioned a déjà vu feeling when learning the writing of argumentative essays or discursive writing in the EAP programme. They claimed they had learned it before and that the course was very boring as they felt the whole learning process was assessment-driven.

(ii) Varied discipline expectations

The students found that discursive writing in the form of essays was not needed or did not seem to be needed in their own fields of study. What frustrated some students was that what they had learned could not be used, at least not in the immediate future as some of them claimed. Face validity of the writing assessment was thus called into question among some students although there were a few students who had faith in the long term benefits.

(iii) 'Authenticity' of assessment task

The assessment task in the study asked students to launch an investigation into a current issue given by their teachers. Student comments on such a task included:

- 'A lot of people have already talked about it... Nothing new really.'
- 'Nothing much to talk about... Too schoolish the topic.'
- 'I am just a student. Who would listen to me? Why bother?'
- 'I am not studying this topic. Why should I spend time on it? I am not interested.'
- 'It's not our concern. I am only learning the language, not the subject matter in the assessment.'
- 'Difficult to get in the task. It's another exercise and another test. No different from what we did in secondary school. I was told this is what we will need in our studies. Anyway, I think I should do my best and give myself a better foundation for the future. GPA 4 is my dream.'

When 'use' is of primary concern in learning a language, 'authenticity' of use was called into question in imaginary task situations.

Academic 'socialization'

According to Gee (1990), we are an insider, colonized or an outsider in academic literacy and that discourse is an identity-kit. Acquisition of academic literacy is itself a social process. Writing in the academic community is an integral part of the way in which the culture (the culture of a discipline) is expressed, developed and maintained (Angelil-Carter, 2000). Yet, to a novice writer, this notion of academic socialization might be a little too remote as one of the students in this study said,

‘I think the teachers know much better than we do about the topic on which we write either in the examination or as practice in class. There is nothing new to say. When they read, they are not interested in what we say, I think, but rather in evaluating what we say. So I am usually very careful when I write.’

This cautious attitude is further elaborated upon by other students who commented:

- ‘I write for the teacher because when I write I would ask the teacher what I need to put into my writing.’
- ‘I dare not say anything very different from what the book says – idea- or language- wise. My opinions/ ideas as well as knowledge are limited. My ability to express is not strong.’
- ‘What’s the point of putting forward my views? After all, the language teacher is only interested in the language? The most important is to know how to obtain a high score.’

This play-safe attitude was pervasive among the students interviewed. It may also reflect the general attitude towards completing assessed tasks. It is quite obvious that teachers might have shown the students the ropes, but the students would have needed plenty of rope in the learning process. To the students, the question of how to learn the language efficiently for a specific purpose (e.g. passing the test) has taken precedence over that of how to learn the language effectively for intrinsic purposes.

Learning to write and writing to learn

The following is what one of the students said when asked about if he found the assessment task useful to his studies with his own department.

‘I learn English because everything in the subjects I study is in English. I want to learn how to write a good report because that is what I need to do. I still don’t understand why we are required to learn how to write essays. No arguments are required in our studies; no references are needed in our writing either. I need some language for report writing only.’

The student’s views were shared by most students in the interview. There was a general feeling that the usefulness of what one learned in the EAP was not as obvious as one would have expected. The notion ‘writing to learn’ at university, which is one of the tenets of academic writing, has obviously been narrowly interpreted by the

students in their specific contexts. The linkage between learning to write in the EAP programme and writing to learn in their disciplines has not been made very explicit on the programme, at least to some students.

Uncertain about performance evaluation

The perceived difference in the university expectations of the endeavours of learning to write and writing to learn seems to have led to some skepticism among the students about the validity of the grades they received or about the evaluation of their performance. Three of the students in the interviews made the following remarks:

- ‘I’m not sure if I really deserved a ‘B’ grade. But one thing is sure – I have tried very hard and that the teacher seemed to be satisfied with my performance. I’ve got only a ‘D’ in my writing in the public examination before I entered university. So ‘B’ was a big encouragement.’
- ‘I got only a ‘C’. I don’t know why. All the writings that I have done in other subjects were fine. No teachers said my writing was not up to scratch. I don’t know how come I had only a ‘C’, while all the writing I had with teachers in my own department got at least 7 or 8 out of 10. Is English in the English subject different from English in other subjects? Or is it because somehow the teacher didn’t like me?’
- ‘I think I’ve been very fortunate as I had a very good teacher. I know exactly what was needed in the assessment. I had a ‘B+’, which was out of my expectation. I think what I have learnt is going to be very useful because I know a lot more about writing than before. I think this will help with my writing in other subjects.’

The above three students had very different views about their results. Yet the commonality is that they viewed writing classes as a means to success in other subjects while at the same they were not certain about the qualities that are attributed to the scores that they received. To some students, assessment is not just reflecting their performance but also the relationship between them and the teachers.

‘Punishment’ for efforts to learn

There was a general feeling among students that while they learn through imitation in terms of language (i.e. mimicking the language, style or tone of a writer), they were often being accused of plagiarizing. The assessment system also

discouraged them to stretch their language abilities or to take risk with testing their own ability limit – the old bugbear being the loss of marks.

‘I still don’t quite know when I am allowed to use the original writer’s words, and when not. My language is in no way comparable to that of the writer. When I read, I usually learn the language too, not just the content. In fact, most of the time when I read passages other than the ones that are related to my formal studies, my focus has always been on language. I usually try to memorize a few sentence structures and expressions, and some vocabulary too. It, therefore, came as rather a shock to me when I was told in the EAP programme that I have to attribute what I’ve learned to the original writers and that I should only use my own words. But my own words come from the original writers’ words. I am only a beginner. I can’t help but feel that I was asked to do something impossible. That’s not fair.’

What the above student has raised is the perennial problem of the status of ‘stolen language’ in learners’ work. The strategy that the learner has been using to learn a language seems to have been discredited. There is a sense of loss as new strategies takes time to develop. The question looming large in the learners’ minds could be issues that are related to ‘language ownership’. This is especially the case with second language learner, who has never ever ‘owned’ the second language in any sense of the word. All they have is the ‘borrowed’ language from the native speaker of the language from the first day they learnt the language.

Discussion

Language and capability

To some students in the study, there did not seem to be any problems with writing at the university for the time being. The only problem seemed to be the writing in the language programme ‘EAP’ itself. The EAP programme being an aid to students’ language proficiency was not yet fully appreciated by some students. The students did not seem to think they had any great problems in dealing with the transition from school to university in academic writing either. There was no mention of difficulties in coping with the university life and studies in English. Most of the students interviewed did not come across as being very keen on the discursive writing of the assessment task. The reasons given included the nature of the task, the topic for writing and the perceived relevance of the assessment to their needs or interests. While language teachers stressed the importance of the skills

learned in the EAP programme and their transfer to the students' disciplines, the students' perceived realities of their learning in their own departments had been different. All this makes us wonder if the EAP assessment could reflect students' capability in handling writing at university or if one should qualify the question by asking whether the EAP assessment could reflect student's capability in handling discursive or argumentative writing at university from the language teacher's point of view. In other words, the assessment itself comes across as a specific task for a specific purpose in a specific context. Specificity of the task in assessment has spelt out the scope of assessment and any results or interpretations of the results needed to be treated with caution. If language assessment is to reflect students' ability in using the language, validity of the context and the task in which the students have been asked to display such ability should be addressed and be explained explicitly to the students in order to obtain the best performance from them.

Language and cognitive skills

Although some students mentioned the importance of teachers' guidance in coping with assessment, there was in general an absence of any detailed discussions on the levels of difficulty of the cognitive skills such as reading and writing in the assessment task. The cognitive skills are supposed to be an important component in the teaching schedule of the programme and should have been brought out clearly in the teaching and assessment itself. A general absence of detailed discussion in the interview could be due to the fact that (as some students had mentioned) the skills had been taught before in the secondary school. 'Familiarity breeds contempt' and thus there was nothing much to say about them and had escaped the students' attention. Another explanation could be that the students found the cognitive skills required manageable and was not worth mentioning as an issue. There might be other explanations; but one thing has been clear and that is the students found teachers' close guidance and the practice task (which was very similar to the assessment task) in the course book important in completing the assessment task. To what extent the assessment task could reflect the students' cognitive skills outside the familiar EAP assessment task context has been a moot question. Moreover, if some students have to struggle along to learn to express ideas in a non-native language, the role of higher order thinking and its perceived importance against language proficiency would be a point of significance to an English programme for academic purposes.

Language and 'self'

It is quite obvious from the interviews that there were two very different preferred learning strategies employed by the students: one that would learn by imitation, the other by taking risks. It seems that, according to the students, the assessment task favour none of these two strategies. To some students, mimicking other's writing is part of their development as a writer, launching them on the way to finding their own styles; to others, stretching their own limits with the language skills available and playing around with words are their preferred ways of enhancing their language repertoire. It looks as though the assessment task itself did not commend either of these learning strategies, giving the impression that the 'self' being not taken seriously into consideration in the learning process. There has also been confusion about legitimacy over mimicking the language in the EAP materials (or those given by teachers) and censure over mimicking language in other reading materials. In short, the overall impression about the written assessment among the students is that the assessment itself was having some negative effects on the students' learning of the language and performance, implying doubts about the assessment task itself being a fair judgement of the students' own language proficiency.

Conclusion

When the learning approach that is espoused and cherished by teachers and institution is autonomous and self-directed learning, the students' perception of the assessment task in this study seems to have been otherwise. The conflict and tension as experienced by the students in their learning process have highlighted the issues of the existing assessment problems. When English is seen as a subject competing with other subjects for students' attention, there is a sense of urgency among students to be 'smart' in working towards assessments efficiently and effectively. This leads to the question of how to address some possible shortcomings that accompany such assessment driven mentality. When teachers of the EAP programme are to be held responsible for the students' language learning and accountable for their time, money and resources spent on the programme, it seems an understanding of the strong and weak points of the adopted assessment procedure and design is crucial for quality assurance in assessment. The assessment designers in the study perhaps need to take into consideration more seriously students' needs, interests and ways of learning.

It would be wise to implement policies that underscore the importance of two distinct notions in assessment: language education (whereby the institution sets up sanctions and merits systems to provide an environment that is conducive to

upholding educational ideals or aims through the teaching of a language) and language learning (whereby an individual can seize upon freedom provided by the institution to make the best use of their personal attributes in the process of learning). For example, a language teacher should try his/her best to ensure that the institutionalized practices are reflecting the best of current knowledge about second language acquisition (or academic writing in the study for that matter), taking into consideration both the latest pedagogical beliefs and research on second or foreign language learning. The sanctions (such as rules regarding plagiarism) and merit system (such as the award of points) of an assessment is a reflection of the essence in a programme and also what the institution believes in and uphold.

Students' results are a mere snap shot of the students' abilities at a certain point of time in a particular context. The assessment itself is a human act, a human judgement of a performance where human elements (such as subjectivity) could hardly be ruled out completely in the process from assessed task design to performance evaluation and result interpretation. A language assessment can be viewed from different stakeholders' perspectives, resulting in 'conflict' and 'tension' being part and parcel of the assessment system, making changes and renewal of the assessment procedures a crucial part of the assessment development. At the same time, it is important for the assessment designers to make the stakeholders of an assessment understand the educational significance in the learning process, what has or has not been evaluated, and how the results should be interpreted.

Food for thought

There might be inherent 'tension' or 'conflict' in whatever assessments or assessment procedures one chooses to adopt. To reconcile educational objectives in learning and educational accountability and responsibility in assessing, the following questions could lend themselves to being made into some useful criteria for a fair and equitable assessment:

- (1) Do the assessment procedures and the assessment fall out from the beliefs and the philosophy held dear by the organization that institutes such assessment?
- (2) Have the assessment procedures and the assessment itself captured the latest insight into the research and development of the assessed area?
- (3) Is the assessment able to reflect what the students should be able to do in practice, not just in the restricted (here and now) assessment context?

- (4) Are the students given an opportunity to learn from the experience being assessed and to put such experience into perspective?
- (5) Are the assessment procedures conducive to students' intellectual 'growth' and self 'transformation'?

References

- Angelil-Carter, S. (2000). *Stolen Language? Plagiarism in Writing*. Pearson Education Limited.
- Craft, A. (1992). *Quality assurance in higher education: Proceedings of an international conference Hong Kong 1991*. London: Falmer Press.
- Fulweiler, T. (2002). *The Working Writer*. Prentice Hall.
- Gee, J. (1990). *Social Linguistics and Literacies: Ideology in Discourse*. Bristol: The Falmer Press.
- Hamp-Lyons, L. (Ed.) (1991). *Assessing Second Language Writing in Academic Contexts*. Ablex Publishing.
- Hamp-Lyons, L. (1997). Ethics in language testing. In *Encyclopedia of Language and Education* (Vol. 7, pp 323-333). Kluwer Academic Publishers.
- Hafernik J.J., Messerschmitt, D.S. & Vandrick, S. (2002). *Ethical Issues for ESL Faculty*. Lawrence Erlbaum Associates, Inc.
- Hill, C. & Parry, K. (1994). *From Testing to Assessment*. Longman.
- Kogan, M. (1986). *Education Accountability: an analytic overview*. Hutchinson.
- McNamara, T. (1996). *Measuring Second Language Performance*. Longman.
- Norton, B. (1997). Accountability in language assessment. In *Encyclopedia of Language and Education* (Vol. 7, pp 313-322). Kluwer Academic Publishers.
- Rowntree, D. (1987). *Assessing Students: How shall we know them?* London: Kogan Page.

Appendix I Syllabus

Part of the syllabus

Course aims

- (1) In general, to help students study effectively in the University's English medium learning environment.
- (2) More specifically, to help students to improve and develop their English language proficiency within a framework of academic contexts.

Contents

.....

Appendix II Writing an academic text with references

For this assignment students need to write a discursive essay and include in it references from sources given to them.

In the essay students will need to plan their arguments, then present and elaborate points on both sides of the argument. They also need to refer to and acknowledge appropriate sources to support information or arguments presented. Finally, they will need to give their personal opinions on the topic and provide bibliographic references.

Appendix III Assessment system band

A+

Content is highly relevant and comprehensive.

Organization, coherence and cohesion are highly effective and well achieved.

Interaction is highly effective.

Register is highly appropriate to the genre and setting.

Grammatical structures and vocabulary are mostly accurate; any errors are non-intrusive.

Range of grammatical structures and vocabulary is highly appropriate to the genre and setting.

Pronunciation is fully comprehensible; any errors are negligible.

Fluency is maintained.

Support materials are of extremely high quality.

Use of support materials is highly effective.

Academic conventions are maintained.

(There are also band descriptors for grades A, B+, B, C+, C, D+, D and F.)

Appendix IV

- (1) How do you feel about the EAP programme
- (2) Have you encountered any problems? Did you manage to overcome the problems? Any example?
- (3) Would you recommend this programme to your friends? Why/why not?
- (4) What do you like most/ least about the programme?
- (5) -----