

THE ENGLISH IMMERSION PROGRAM: MEASURING THE COMMUNICATION OUTCOMES

Jane Lockwood

Department of English, City University of Hong Kong

Email: lockwood@cityu.edu.hk

APA Citation: Lockwood, J. (2015). The English immersion program: Measuring the communication outcomes. *Indonesian EFL Journal*, 1(1), 98-107

Received: 10-09-2014

Accepted: 13-10-2014

Published: 01-01-2015

Abstract: This paper explores how language assessment is typically used to measure language gain as a result of the Immersion experience abroad. It also explores ways in which this might be improved. This study explores a recent experience where Australian immersion providers, participants and funders all report significant intercultural awareness raising and improved confidence in understanding and speaking in English as a result of the sojourn. However, it transpires that the immersion providers used traditional proficiency focused language assessment tools on entry and exit to measure communication outcomes across the skills of listening, speaking, reading and writing. It appears therefore, that there is a gap between what the perceived outcomes and value are of the Immersion experience, and how they are currently measured. This article reports on a small scale study exploring the perceptions of two immersion providers in Australia, one immersion coordinator in Hong Kong and four returnees on the language assessments they used and experienced, particularly probing on how well they felt these assessments measured their communication gains as a result of the immersion experience.

Keywords: *Language assessment, immersion, indigenous criteria.*

INTRODUCTION

Millions of students head off to international destinations to spend time living with 'homestay' families and studying in tertiary institution that offer different language and cultural experiences. OECD figures predict that this number will reach 8 million by 2025 (Davis 2003). In Asia, Universities have, over the last two decades, been sending increasing numbers of students to English speaking destinations such as UK, USA, Canada, Australia and New Zealand to improve their English language skills in the authentic contexts. (Bodycott & Crew 2001). In 2008-2009, the Hong Kong Institute of Education spent just under HKD2 million subsidizing short term immersion programs in such English speaking destinations as UK, Canada and Australia, but little data exist on how effective such programs are in terms of English language communication gain.

This article reports on two language immersion providers' tools and processes used to measure language proficiency gain. These measures are briefly described and

then the outcomes of key stakeholder groups including the providers, the returnees and the Hong Kong based immersion coordinator are analyzed in terms of the efficacy of such measures. The Immersion experience of Hong Kong Institute of Education undergraduate students takes place over a full semester where they live with homestay families and attend university programs in either Canada, Australia or the UK. Each student does a pre- and post - course assessment as well as the completion of an ethnographic study as part of their program.

If intercultural awareness is a key area in which sojourners are expected to make progress while abroad, then appropriate methods of assessment should be used to measure the gains they have made. While much attention has focused on the preparation, format, and the content of the study abroad programs, relatively little has been published about the modes of assessment (Jackson 2005:165). This article will report on student and staff perceptions of the communication gains and finally it

report on student and provider perceptions of the efficacy of the assessments used and suggestions for improvement.

What do English language assessment practitioners and researchers say about measuring communications skills?

English language assessment has received a great deal of attention from applied linguists over the last 35 years. In the past, teachers have mostly concerned themselves with simply testing the discrete language learning outcomes of program delivery in language learning classrooms. Commercial English language tests abound but tend to focus on the measurement of generic or academic English language proficiency gain across the skills of listening, speaking, reading and writing. These trends in turn have resulted in a preoccupation with the summative assessment comprising of a series of discrete language items that add up to an overall score. This emphasis on the quantitative score as the sole indicator of program success reflects a psychometric approach to language program evaluation and has been seriously challenged by language assessment experts (Hamp Lyons 1991; Shohamy 1992; Brindley 1995; Bachman and Palmer 1996, McNamara 1996, Douglas 1999).

More recently much of this research has centred around how language teachers and assessment providers can improve their tools and processes to reflect the whole context as well as the specific purpose for assessment. There has been a paradigmatic shift from positivist models of language assessment to more constructionist models where the language assessment tools and processes are designed to be more responsive to the particular contexts and purposes for language assessment.

Much discussion has also taken place about the introduction of 'formative' language assessment tasks e.g. assessment for learning tasks that take place in the class room, portfolio assessment and other on-going assessment tools (Hamp Lyons 1991) that are qualitative but can also yield scores. As well, much discussion has also taken place about the veracity of reporting on language

proficiency indicators as opposed to language performance indicators as being more or less appropriate for different contexts and purposes of language training (McNamara 1996); of qualitative rather than quantitative indicators (Shohamy 1998); of tailor made assessments as opposed to standardized assessment (Douglas 1999, Pennington 1998).

The two unifying themes in the most recent language testing and educational literature revolve around the related issues of being open to different forms of assessment and being open to how these forms of assessment can be context and purpose sensitive. The language testing literature is rich in the discussions about the relative merits of using different applied linguistic frameworks to achieve construct validity and reliability in the design and use of any test. Much of the discussion has revolved around pushing criteria for language assessment beyond the traditional preoccupations with pronunciation and grammatical accuracy to a consideration of broader communicative domains such as discourse and interactive capabilities (Canale and Swain 1980; Bachman and Palmer 1982; Bachman and Savignon 1986; Bachman 1990, Davies 1988; Hughes 1989; Weir 1993, Douglas 2005).

Assessment practices that reflect communicative approaches to language training in Language for Specific Purposes (LSP) contexts is also well covered (Lumley and Brown 1996; McNamara 1997; Douglas 2000; Elder 2001), and well-documented accounts of test development for different groups of occupations and professionals abound. McNamara (1990,1996);Elder (2001) and McDowell (1995) have both developed standardised and performance-based tests for a range of teachers and health professionals in Australia and Douglas (2000) looks at specific language use situations to develop test content and test methods for highly specific LSP, such as English for pilots and air traffic control. Such frameworks, however, have yet to be fully understood in an industry context and incorporated into business language assessment practices on-site (Lockwood 2002; Lockwood 2008).

In recent discussions about language for

specific testing (Jacoby and McNamara 1999, Douglas and Myers 2000, Douglas 2001), there has been a call for what they have termed "indigenous assessment criteria". Such criterion is derived from the target language use (TLU) context. Performance assessment practices are seen as more relevant to knowing whether a candidate can handle a complex professional and/or social situations where mere language proficiency indicators are not likely to provide appropriate performance profiles. The Immersion context can be viewed as another 'indigenous assessment' site encompassing the homestay context and the university context. What is interesting about this context is the intercultural richness of the immersion experience on the one hand and the requirement, on the other hand, to report outcomes in terms of language proficiency gains to secure and maintain funding.

Much has been written on intercultural assessment (Kramasch and Sullivan 1996; Byram and Fleming 1998). All start with attempts to define what it is exactly we are trying to assess when we assess intercultural performances; is it knowledge?; is it skills?; is it behaviours?; is it motivation?; is it all of these? Some even argue the efficacy of such a goal as the perfect mastery of a second language where acculturation is the aim (Seelye, 1984). A large number of standardized measures exist to assess intercultural knowledge (Allen & Herron, 2003; Coleman, 1995; Redden, 1975; Corbitt, 1998; Pedersen, 2010); these are typically achieved through surveys, inventories, proficiency exams and multiple choice tests to quantify the amount of knowledge absorbed regarding the target culture. However, Earley and Ang (2003) suggest that the assessment of intercultural intelligence is much more than mere cognition and knowledge. They propose a three dimensional model that encompasses knowledge, behaviour and motivation. They further propose that non - psychometric methods can be applied appropriately to assess the motivational and behavioural components of cultural intelligence (CQ).

Jackson (2005) came to much the same conclusion in the study she carried out which

involved introspective, qualitative assessment of the sojourn experience as captured by a group of Hong Kong students when away on immersion. In this study students were encouraged to use first person introspective accounts (in the form of diaries) to assess the learning processes of short term sojourners as a way of measuring intercultural and communication gains. These were then assessed against standardized grading criteria (A-excellent ----F-fail).

In an earlier study which involved Hong Kong students on immersion (Evans, Alano and Wong 2001), three levels of intercultural awareness were hypothesized as 'emergent, growing and enhanced' (p.96) and videotaped discussions Hong Kong students on English speaking Immersion programs were analyzed for both linguistic and cultural gains. Another study in the late 1990's (Murdoch and Adamson 2001) investigated the extent and range of English use before and during a 4 - week immersion experience in Australia as a way of explaining the sociolinguistic gain from this experience reported by students. Not surprisingly these studies showed considerable increase in English use in a range of social and educational settings and situations. These kinds of studies may hold the key for imagining more enriched assessment tools and processes for Immersion programs that take Hong Kong students away overseas.

METHOD

This study is a qualitative enquiry into the pre and post course language assessments of two HKIED language immersion providers domiciled in Australia. Two immersion provider coordinators provided documentation regarding the assessments they currently use, the most up to date records of the results, as well as the end of course reports. They were interviewed specifically about the assessment processes in a semi- structured interview on the phone. Four returnees also agreed to be interviewed about what they perceived to be the main gains of the immersion experience and how these were assessed and evaluated. They agreed to provide their end of program ethnographic studies and in one case, one of

the returnees volunteered her diary which she had kept during her stay overseas. Finally, the Hong Kong immersion coordinator was interviewed to elicit his views on what he felt were the gains of the immersion experience and how these match with the tools and processes used to measure communication gain. All interviewees were encouraged to provide suggestions for improved and alternative measurement and these are discussed in some detail in this article.

FINDINGS AND DISCUSSION

What do the pre and post immersion program scores indicate about the improved language levels of the students?

The English language assessment tools used by the two providers are described in detail later in this section. They are similar in that they assess the four skills of listening, speaking, reading and writing. Interestingly, the results indicate gains across all the skills and are not just confined to the skills of listening and speaking, reading and writing also appeared to report gains. No student appeared to go down in any of the skills areas, although there were differences in the percentage gains particularly in the areas of listening and speaking where one student appeared to have made gains of over 100% on the initial score for speaking. The average gains across the four skills for the two providers are tabulated below:

Table 1: Average language gains of students on immersion

	Listening	Speaking	Reading	Writing
Provider 1	20%	30%	10%	10%
Provider 2	4.3%	6.5%	3.9%	3.75%

In the data provided by both provider groups it appears that the speaking levels followed by the listening levels showed most improvement with a fairly even distribution across the writing and reading levels. What is interesting in the data is the significant difference in the reported % of gain across the skills of listening and speaking between the two provider groups. Provider 1 reports much higher post immersion gains than Provider 2. Given that both immersion programs took place in Australia at about the same time, and given that the students were not 'streamed' into levels (i.e. a more able group going to Provider 2 and a less able group going to Provider 1), one of the variables that needs to be considered is the assessment tools and processes themselves.

What do the providers do and think about the assessment processes they carry out?
Provider 1

Provider 1 complied with the tender regulation that requires pre and post course assessment by administering an (Language Proficiency Assessment for teachers of English) LPATE type test. This comprises a reading and listening comprehension that test the comprehension of authentic type texts through cloze testing, multiple choice and open-ended

questions. The open-ended questions probe reading skills that go beyond information extraction to eliciting students' interpretative, inferential and analytic skills of what they have read and listened to. The writing test is also designed to mirror the LPATE test and assesses both the ability to write extensive text as well as the ability to recognize student error and use appropriate metalanguage to explain the errors. Similarly, the speaking test mirrors the LPATE test tasks and assessment criteria. The students are asked to read aloud an extensive passage and are marked with criteria that are related to pronunciation ability as well as the ability to read aloud with meaning. The last criteria is highly contentious in that students are really being assessed on their reading comprehension ability as well as their pronunciation. The second part of the speaking test is a short semi - spontaneous presentation on a given topic. Here the students are assessed on their ability to organize what they are going to say as well as their ability to use grammatical and lexical items accurately and showing range. Finally the students are asked to participate in an open discussion with three other participants around a given topic. Here the criteria for assessment relate to their ability to interact well with their peers.

Jane Lockwood

The English Immersion Program: Measuring The Communication Outcomes

The pre and post assessment record is sent to the funding provider in compliance with the tender regulation. However none of the returnees knew what their language assessment results were, nor did they appear concerned to know, as they understood that these were not high stakes test. The LPATE test is marked on a 5 point scale with .5 levels recorded if the candidate exhibits features in the whole number score.

Interestingly Provider 1 felt there were extreme limitations on what this language assessment could measure as it relates to the immersion experience. Its value seemed to relate more to compliance with the tender regulation. However, this provider also said:

We knew that the assessment would not capture what we felt were the real communications gains of immersion but nonetheless we wanted something that was valid and reliable; something that would be useful to the student and would also comply with the tender regulation. After consulting (an external assessment group) we decided it would be most useful for students to do an LPATE type test as we know this is a high stakes test for the immersion students when they return to Hong Kong.

During the course of the interview, the course provider coordinator expressed her views about the gap between the current assessment procedure as described above and what she felt the real gains were in terms of communication improvement of the cohort of English major students from the Hong Kong Institute of Education. She said:

We know that students make huge gains beyond what is routinely assessed in the pre and post course test. For example the students develop great confidence in their ability to speak and understand. At the end of immersion they are able to engage much more in informal day to day interaction as a result of living with English speaking families, travelling around Australia and making friends on campus. They allow themselves more thinking time when speaking and they rephrase what they say to make their meaning clear...there has obviously been a shift from worrying about themselves as communicators to worrying about whether they are making their meaning clear to the person they are talking to...a sure sign of a good communicator. This shift from a preoccupation with the mechanics of language and grammar accuracy to real communicative ability including more confident paralinguistic and intercultural behaviours are huge.

Provider 1 suggested it would be difficult to measure, in a standardized format, this kind of shift and suggested one vehicle for tracking this change would be in the actual presentation of the ethnographic study, which is an end of course requirement, as well as the quality of observation contained therein. The provider cited one student in the 2008/2009 cohort as follows:

One of the mainland students made a remarkable transformation in his body language and communicative confidence during immersion. When it came to presenting his ethnographic study, he stood up straight (where previously he had stooped), he made eye contact and delivered his findings in such a confident and articulate way...I couldn't believe the transformation but this is not unusual.

Provider 1 suggested that perhaps attempting to develop some kind of standardized marking scheme to capture this shift either as part of the ethnographic study assignment and /or the post test in speaking would be worth considering. She suggested the videoing this final presentation may provide a source of evidence for further speaking assessment and gain. It should be noted however that Provider 1 post immersion speaking and listening scores reported very significant gain. It is the nature of this assessment construct that is being challenged for validity in this chapter

Provider 2

Provider 2 Immersion coordinator also represented a tertiary provider in Australia. Like Provider 1, Provider 2 felt there were huge gains made in confidence and intercultural awareness that translated into improvements in communications strategies not reflected in the Provider 2 pre and post course language assessments, which showed relatively small gains. This coordinator also bemoaned the lack of standardization across the immersion provider group in terms of pre and post course assessment. He reported that a core group of the immersion providers had informally agreed to administer the IELTS tests as a way of aligning, although he did not believe that this to be an ideal tool. Currently this provider uses the Professional English

Assessment for Teachers (PEAT) developed in New South Wales. This assessment has been specifically designed to evaluate whether a teacher's English proficiency is good enough to interact effectively in a school setting. The test consists of four components: Listening, Reading, Writing and Speaking and results are given in terms of Bands A, B, C or D but are ultimately expressed in terms of an overall percentage score.

Band A means the candidate can for Listening, comprehend easily and accurately in all personal and professional contexts; for Reading, are able to read all styles and forms of the language pertinent to their professional needs; for Writing, are able to write fluently and accurately on all levels normally pertinent to their personal and professional needs ;and for Speaking, are able to use language fluently and accurately on all levels normally pertinent to their personal, social, academic or professional needs. (http://www.trb.sa.edu.au/english_tests.htm (accessed 5/2/2010))

Provider 2 had two main concerns with the current tool and processes. First, he felt that students do not necessarily show their strengths soon after arrival. He felt that in the first few days many students are suffering from culture shock and a dip in their own confidence to communicate. This may be the result of being 'thrown in the deep end' with their homestay family and generally being in a strange new environment; but whatever the reasons, it was felt that the test at this stage may not be yielding results that are commensurate with what the students can really do.

Provider 2 was also concerned about the areas that the test 'do not tell you about'. When probed on this point he cited the enormous confidence gain that the students experienced during the immersion experience, which he said were evidenced by a range of improved communication strategies such as initiating conversations, turn taking, changing the topic, maintaining eye contact, responding to cultural references, improved pronunciation (particularly prosodic features) and so forth. He described the linguistic and intercultural challenges in the process of carrying out the ethnographic study as follows:

To complete the ethnographic study component of the immersion program, the students first have to interview NNS students on campus...we send them out with cassette players. They then follow this up by interviewing a native speaker and then their home families. Finally they do a 4th interview of their own choice but it must be at a much deeper level.

Students finally write up and present their findings before they leave. From the data collected the students chose to focus on a broad range of ethnographic study topics such as comparative studies between Australians and Chinese on racism, public and private transport choices, part-time work choices and one study was carried out to explore the reasons for the prevalence of graffiti in Melbourne.

What do students do by way of assessment and do they think it measures the communication gain made on immersion? What suggestions do they have?

The student interviews

Three of the students interviewed had completed their immersion program in Provider 1 university. When self reporting their perceived language gains they were unanimous about their improved speaking and listening ability and general increase in self confidence and intercultural awareness when interacting in English. Whilst the homestay families varied in terms of age, social status and numbers of family members, the students attributed much of their newfound communicative confidence gains and intercultural insights to the regular interaction with their homestay families and to their close observations of how these families lived. The students all appreciated the time homestay family members took to find out, at mealtimes, about their daily activities. One of the students who was ethnically Chinese, but a native speaker of English having been educated internationally, recounted the biggest challenge as being intercultural:

I had to gauge how to deal with things going on in my homestay family like telling them that the meat portions are way too big; that I was really scared of one of their dogs and that I wanted them to pick me up on Saturday night from the

Jane Lockwood

The English Immersion Program: Measuring The Communication Outcomes

city as I was too afraid to come home on the train by myself...Even though I had the language to communicate with them perfectly well there were these kinds of situations, especially in the beginning, where I wasn't sure whether it was the right thing to do...the right thing to say.

One of the students said she was surprised at the easy intimacy displayed by her homestay 'mum' which, she said, was unusual in Chinese culture. She said she learned about how making jokes within families about individual members was a kind of endearment and was a hallmark of acceptability even when the jokes were risky. This same student observed this kind of joking among students she befriended on the campus. She said, *You know you are accepted when they make a joke about you!*

These students also said they felt different kinds of communicative gains when they went travelling. They reported serendipitous friendships struck up with other international travelers from Japan, Holland and Belgium when they travelled interstate. As English was the 'lingua franca' they reported a different kind of experience in talking with fellow tourists.

It was a different kind of talking because we both had English as a second language in common and because we had already spent about 10 weeks in Australia with our families we were able to display our knowledge of the culture...With the Japanese speakers we adjusted our English so that they could easily understand...We felt very confident at this stage of our immersion stay and very proud of our ability to communicate.

A deeper kind of cultural shift appeared to have taken place in the students when they self reported the value of attending the different programs at the provider university. One of the students articulated this well when she said:

I had learned about student-centred and enquiry based learning before but I felt what this was like in the classroom in Australia. The lecturer only took 25% of our time and we had to make our own decisions for what to do next we were left on our own a lot but for my classmates this was quite usual.

Whilst some of the students made friends on campus and these endure through

Facebook, some said they spent their social hours with each other. This finding was evident in one of earlier studies which reported:

The greatest deficiency (as shown by figures for reported use as well as reported perceptions) lies in opportunities presented by this program to expand sociolinguistic competence related to meeting, conversing and socializing with Australian peers-students of a similar age... (Murdoch and Adamson 2001:111)

One of the students commented on the odd appearance of young people in Melbourne which she found strange and quite intimidating:

On my way home I had to catch a bus at Frankston station. The young people there wore weird clothes and had lots of piercings and tattoos. They looked quite rude and I didn't dare talk to them.

For all of the students the pre- and post-course assessment was completed under no pressure. One student reported that they were told it was an institutional requirement only and not high stakes. None of them know their results from the test but recognized that the assessment was based around the four skills assessed in the LPATE exam. Some said the assessment was very easy and all agreed that they were not able to display the real communicative competency gains they felt they had made on the immersion. They found the question of how to mend this gap between the gain and measuring the gain better, confounding, and commented:

It's really hard to measure what we have learned in terms of socialization. I mean you know when someone is being socially or culturally inappropriate, but it is hard to measure exactly how good you have become.

Another student reported her own observations of her friend who had arrived with relatively low confidence in her speaking ability compared to when she left, and put it down to confidence more than language proficiency gain.

On the last day before we left my friend and I and her homestay mum went to a movie together. It was amazing to see her chatting with her host

mum...she could really hold her own...they were joking and messing around. When we arrived in Australia she could only talk, hesitating all the time and looking for words.

All the students commented on the value of the ethnographic project they were expected to complete during their immersion stay and cited this as perhaps evidence of one aspect of their communicative development. One student suggested 'writing about' the gain as part of their final assessment, but of course recognized its limitation as 'knowledge' evidence rather than communication skills effectiveness.

The Hong Kong coordinator

There was very broad agreement from the Hong Kong coordinator's point of view with those of the Australian providers and the returnees about the mismatch between the real communication gains and the assessment practices currently being carried out pre- and post - immersion stay. Interestingly, he made the point that these current assessment practices may reinforce in the students' minds that language gain is all about making improvements the kinds of academic language tasks in the LPATE and PEAT assessments and may inadvertently work against the development of the less formal communication skills for which the immersion experience is reported as so valuable. He agreed however that to assess the reported gains in confidence, intercultural awareness and the ability to participate in casual conversation to be extremely challenging and doubted whether they could be captured in a quantitative score. He also agreed that the lack of systematic tools and processes across the providers would mean that reported results could not be relied on as a measure of success and gain.

Rather than changing the current assessment practices by changing the tool itself, this coordinator felt something much more radical may be considered. He suggested that perhaps the Immersion experience could be built into their course become a 3 credit program of study roughly divided into 3 segments, each attracting one credit point as follows:

- (i) Pre departure studies –e.g. preparing e.g. how to keep a journal (one credit)
- (ii) Journal keeping when they are away (one credit)
- (iii) Portfolio when they return (one credit)

This would both ensure that the students would make better preparation and use of their immersion experience as well as being formally assessed as part of the degree studies. This program could also embed into it autonomous learning principles and improved skills in self and peer assessing a range of linguistic and intercultural skills.

It would appear that the overseas providers, and the students themselves, felt constrained and dissatisfied by the language assessment tools and processes they currently use to measure the communication outcomes of the immersion experience. It would appear however, that the assessments used did in fact report gain across the skills with specific reference to speaking and listening. One of the big problems seems to be the lack of a systematic proficiency test used by all providers. This would generate more reliable results than the ones reported in this study, assuming calibration and moderation processes were also in place. The other issue to emerge from this study is the need for **supplementary** information about the perceived 'other' communication gains that are not captured in the domains of the current tools being used. The funding institution itself expressed doubt about the current use of the pre- and post-course assessment tools and processes and proposed better ways to measure the outcomes of the immersion experience. Can the results of the pre and post course assessment really reflect this shift? Clearly the gains appear to impact the listening and speaking skills particularly, but go far deeper than improved language proficiency. All interviewed in this study reported that the immersion experience offers the opportunities of observing and participating in authentic English speaking contexts at home and at university. These two rich contexts throw up 'indigenous criteria' that may be able to be mapped into immersion communication assessment tools and processes. As well, it could also be argued that a scale of intercultural gain, as suggested

by Evans et al (2003) could be mapped onto such a revised assessment. However, does it need to be a test? The current literature as outlined in a previous section of this article describes other forms of assessment as measures of gain. On-going assessment processes such as diary keeping, focus group discussions that probe critical experiences at home and on campus, ethnographic studies that involve data collection and analysis on the people around them, all that yield rich qualitative information that demonstrate “all those shifts that formal pre and post course assessment doesn’t” (Provider 2). A radically revised program that is credit bearing, as proposed by the funding institutional coordinator could also embed tasks and measurements that reflect the multidimensional communication gains made on immersion.

One of the greatest limitations of this study has been the small scale of the investigation and the difficulty in accessing provider assessment tools and processes. Both tools had been developed for teaching contexts rather than to probe language gains from an immersion experience. It would be of great interest to know if any of the providers have attempted to map into their language assessment tools and processes any of the ‘indigenous criteria’ that Douglas (2001) argues for in any language for specific purpose (LSP) testing situation. Self-reporting, peer rating, provider rating of the ethnographic study presentation and even homestay ratings using a validated rating tool may supplement the pre and post course test as it currently exists and ultimately some kind of standardized measure may well be developed.

CONCLUSION

Whilst the tender documents clearly state a requirement for pre and post course assessments, it seems the funding authority in this study rarely demanded to see these results as proof of success of the immersion experience. Reports written by the two providers all reported language gains of the students as ascertained by the LPATE and PEAT type tests although that the percentage gains were significantly divergent. However,

the returnees’ own confidence in their own improved language skills and in their new found skills of living and travelling independently were testimony of communication improvement at quite another level. Perhaps in critiquing such tools and processes, it is not the instrument in itself that is wholly to blame, but more the limitation of its use alone as the proposed reporting and program evaluation instrument, as well as a lack of a standard tool. This study has reported on a range of approaches and types of evidence that could be used systematically to evaluate success in communication as a result of the immersion experience. Perhaps it is therefore not so much the assessment tools and processes that should come under scrutiny but more the program evaluation brief that provider institutions are being asked to comply with to demonstrate that the Hong Kong funding institutions are getting value for money.

References

- Bachman, L & Palmer, A. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly*, 16, 449-465.
- Bachman, L & Palmer, A. (1996). *Language testing in practice*. Oxford: OUP
- Bachman, L & Savignon, S. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal*. 70, 380-390
- Bachmann, L. (1990). *Fundamental considerations in language testing*. Oxford: OUP
- Bodycott, P. & Crew, V. (2001a). Short term English language immersion: Paradox, culture shock, participant adjustment. In P. Bodycott & V. Crew (Eds.), *Language and cultural immersion: Perspectives on short term study and residence abroad* (pp. 1-9). Hong Kong: Hong Kong Institute of Education.
- Bodycott, P. & Crew, V. (Ed.). (2001b). *Language and cultural immersion: Perspectives on short term study and residence abroad*. Hong Kong: Hong Kong Institute of Education.
- Brindley, G. (1995). *Language assessment in action*. Research Series 8. NCELTR
- Byram, M & Fleming, M. (1998). *Language learning in intercultural perspective: Approaches through drama and ethnography*. Cambridge. Cambridge University Press.
- Canale, M. & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1, 1-47.
- Davies, A. (1988). Communicative language testing. In Hughes, A. (ed) (1988). *Testing English for university study*. ELT Document No. 127. Modern

- English Publications
- Davis, T.M. (2003). *Atlas of student mobility*. New York: Institute of International Education
- Douglas, D. (2000) *Assessing languages for specific purposes*. Cambridge: Cambridge University Press
- Douglas, D. (2001). Language for specific purposes assessment criteria: where do they come from? *Language Testing*, 18, 171-185.
- Douglas, D. & Myers, R.K. (2000). Assessing the communication skills of veterinary students: whose criteria? In Kunnan, A. (eds.), *Fairness in validation of language assessment*. Selected papers from 19th Language Testign and research Colloquium. Studies in Language Testing 9. Cambridge. Cambridge University press. pp60-81.
- Elder, C.(2001). Assessing the language proficiency of teachers: Are there any border controls? *Language Testing*, 18 (1) 149-170.
- Evans, M; Alano, C & Wong, J (2003) Does a Short Language Course in an Immersion Setting Help teachers to gain in Communicative Competence? In P. Bodycott & V. Crew (Ed), *Language and cultural immersion Perspectives on short term study and residence abroad*. Hong Kong: Hong Kong Institute of Education.
- Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam: John Benjamins.
- Hamp-Lyons, L. (1991). Reconstructing 'academic writing proficiency'. In Hamp-Lyons, L (ed) *Assessing second language writing in academic contexts*. Norwood. N.J. Ablex
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Jackson, J. (2005). *Language, identity and study abroad*. Equinox.
- Jacoby, S & McNamara, T. (1999). Locating competence. *English for Specific Purposes*, 18(3), 213-241.
- Ife, A. (2000). Language learning and residence abroad: How self-directed are language students? *Fremdsprachen und Hochschule*, 60, 42-63.
- Kramasch, C & Sullivan, P. (1996). Appropriate pedagogy. *ELT Journal*, 50(3), 199-212
- Lockwood, J, (2002) *Language Training Design and Evaluation Processes in Hong Kong Workplaces*. Unpublished PhD. thesis. Hongkong: HKU
- Lockwood, J. (2008) What does the Business Processing Outsourcing Industry (BPO) industry want from English language assessment? *Prospect* 23,2 (2) *Special Issue* pp.67-78.
- Lumley, T. & Brown, A. (1996) Specific purpose language performance tests: task and interaction. *Australian Review of Applied Linguistics*.
- McAllister, L., Whiteford, G., Hill, B., Thomas, N., & Fitzgerald, M. (2006). Reflection in intercultural learning: Examining the international experience through a critical incident approach. *Reflective Practice*, 7(3), 367-381.
- Mc Dowell, C. (1995). Assessing the language proficiency of overseas-qualified teachers- The English Language assessment (ELSA) In Brindley, G (1995). *Language assessment in action Australia*. Sydney: NCELTR
- MacFarlane, A. & Wesche, M. B. (1995). Immersion outcomes: Beyond language proficiency. *Canadian Modern Language Review*, 51(2), 250-274.
- McNamara, T. (1990). *Assessing the second language proficiency of health professionals*. Unpublished PhD. Melbourne: University of Melbourne.
- McNamara, T. (1996) *Measuring Second Language Performance*. London: Longman
- McNamara, T. (1997) Interaction in second language performance assessment: Whose performance? *Applied Linguistics*, 8, 446-466.
- Murdoch, P., & Adamson, B. (2003). The effects of a short term study abroad programme upon sociolinguistic competence in the use of English as a second language. In P. Bodycott & V. Crew (Ed.), *Language and cultural immersion: Perspectives on short term study and residence abroad* (pp. 101-112). Hong Kong: The Hong Kong Institute of Education.
- Pedersen, P. (2010). Assessing intercultural effectiveness outcomes in a year-long study abroad program. *International Journal of Intercultural relations*. 70-80.
- Pennington, M. (1998). Designing language program evaluations in specific purposes. In Li, E. & James, G. *Testing and evaluation in second language education*. Hongkong: HKUST.
- Seelye, H.N. (1984) *Teaching culture: Strategies for intercultural communication*. Lincolnwood. Illinois: National Textbook Company.
- Sercu, L. (2004). Assessing intercultural competence: A framework for systematic test development of foreign language education and beyond. *Intercultural Education*, 15(1), 73-89.
- Shohamy, E (1992) New modes of assessment: The connection between testing and learning. In Shohamy, E & Walton, A.R. (eds) *Language assessment for feedback, testing and other strategies*. Dubuque IA: Kendall/Hunt
- Shohamy, E (1998) Alternative assessment in language testing: Applying a multiplism approach. In Li, E & James, G (eds) *testing and Evaluation in Second Language Education*. Hong Kong Language Centre, HKUST .
- Tilleman, H.H. (2000). Belief change towards self-directed learning in student teachers: Immersion in practice or reflection on action. *Teaching and Teacher Education*, 16(5-6), 575-591.
- Weir, C. (1990) *Communicative Language Testing* Hemel Hempstead, UK: prentice Hall
- Williams, T.R. (2005). Exploring the impact of study abroad on students' intercultural communication skills: Adaptability and sensitivity. *Journal of Studies in International Education*, 9(4), 356-371.