CrossMark

Lang. Teach. (2015), **48.2**, 217–234 © Cambridge University Press 2013
doi:10.1017/S0261444813000190 First published online 13 June 2013

Plenary Speeches

Issues in post-entry language assessment in English-medium universities

John Read University of Auckland
ja.read@auckland.ac.nz

Students entering English-medium universities around the world come from diverse linguistic backgrounds and many find it difficult to cope with the language demands of their degree courses. This speech focuses first on the motivations for introducing what are known in Australia as post-entry language assessments (PELAs). I discuss in particular the Diagnostic English Language Needs Assessment (DELNA) at the University of Auckland, which has been successfully implemented throughout the university in the last ten years. This leads to a broader discussion of three significant issues in post-entry language assessment. One is the nature of the construct on which such assessments are based. A related issue is whether a PELA can appropriately be considered 'diagnostic' in nature. The third issue is how to undertake the validation of a PELA, particularly with regard to the consequential aspect: does the assessment contribute to enhancing the students' academic language ability?

1. Introduction: the Australian university context

I am going to discuss a form of assessment which has come to be known in Australia as POST-ENTRY LANGUAGE ASSESSMENT (PELA). A PELA is administered to students after they have been admitted to a tertiary institution, with a view to identifying those who are likely to struggle to meet the language demands of their degree programme and who should be encouraged or required to enhance their academic language skills. I will focus particularly on the situation in Australia and New Zealand in this presentation, for obvious reasons, but I want to make the point at the outset that the issues are truly international in scope. They apply not just to universities in the English-speaking countries and in Commonwealth nations where English has been the medium of instruction in the post-colonial period, but also to institutions in Europe, the Middle East, East Asia and other regions of the world where the pre-eminent status of English as an international language has created a great upsurge in demand for English-medium education at the tertiary level, along with pressures to introduce the language as a compulsory subject in schools at a much younger age than before.

Revised version of a plenary address given at the Inaugural Conference of the Association for Language Testing and Assessment of Australia and New Zealand (ALTAANZ), University of Sydney, Australia, 10 November 2012.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Let us start with the Australian context. I feel a little diffident as a New Zealander in dealing with this aspect of the topic – somewhat like an ornithologist from my country coming to a conference in Australia to speak about kookaburras or emus. I hope at least I can give an accurate overview of recent developments in this country. Assessment initiatives of this kind go back at least to the early 1990s, when the instrument now known as DELA (Diagnostic English Language Assessment) was developed at the University of Melbourne (see Ransom 2009 for a contemporary account of DELA). This was just a few years after the International English Language Testing System (IELTS) was introduced to Australia and quickly became the dominant means by which international students demonstrated that they had at least a threshold level of proficiency in academic English as a condition of being admitted to a university degree programme. However, over the years, alternative means of entry into undergraduate study opened up for international students, through pathways such as an upper secondary school qualification, foundation studies programmes, the intensive EAP courses known in Australia by the acronym ELICOS (English Language Intensive Courses for Overseas Students – www.elicos.com) and sub-degree programmes.

The consequences came rather dramatically to public attention in 2006, with the widespread media coverage of a paper by a demographer specialising in immigration research at Monash University in Melbourne (Birrell 2006). In his paper, Birrell pointed out that international students graduating with degrees in accounting and information technology had the right to apply for permanent residence in Australia. However, a substantial number of them were unable to meet one of the requirements, which was to achieve an overall band score of 6.0 on IELTS. Birrell traced the problem back to the fact that the students had originally entered Australia for a non-degree course on a visa which might require an IELTS score of as little as 4.0. They had then followed one or more pathways into undergraduate study without needing to achieve the normal admission standard of 6.0 or even 6.5 – and three years of English-medium study had not developed their language proficiency to that level. This also meant that their English ability was insufficient to meet the language demands of the jobs they aspired to take in Australia.

One outcome of the public debate on the matter was the National Symposium on English Language Competence of International Students (AEI 2007) in Canberra in 2007, organised by Australian Education International, an agency of the federal government. The symposium brought together specialists in English language teaching, higher education, vocational education and training, and other education sectors to review existing policies and programmes in this area and to make recommendations on how the English language proficiency of international students could be improved. This led in turn to a project by the Australian Universities Quality Agency (AUQA) to develop the Good Practice Principles for English Language Proficiency for International Students in Australian Universities (AUQA 2009). Subsequently, the agency used the principles as a tool in their academic audits of universities across the country. In August 2012, AUQA was superseded by the Tertiary Education Quality and Standards Agency (TEQSA), which now requires institutions to report key indicators that all their students are achieving their academic potential, including high levels of English proficiency. A TEQSA official told higher education providers at a briefing that same month that English language standards were 'a major, decade-long, sector-wide issue', which would receive close attention from the auditors (Lane 2012).

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

I want to highlight just four of the Good Practice Principles, which I have re-ordered here into a logical sequence:

1. Universities are responsible for ensuring that their students are sufficiently competent in English to participate effectively in their university studies.
7. Students' English language development needs are diagnosed early in their studies and addressed, with ongoing opportunities for self-assessment.
2. Resourcing for English language development is adequate to meet students' needs throughout their studies.
6. Development of English language proficiency is integrated with curriculum design, assessment practices and course delivery through a variety of methods (AUQA 2009: 4).

The most relevant principle from our point of view in language assessment is No. 7, but notice that even here the issue is not simply how to diagnose students' needs but also how to address them, in order to ensure effective English language development. The theme that assessment is not an end in itself is one that I will emphasise again later.

For institutions seeking to adopt the Good Practice Principles, a very useful resource is the newly published book *English language standards in higher education: From entry to exit* (Arkoudis, Baik & Richardson 2012), which presents a comprehensive account of best practice in implementing the principles. Interestingly, the authors devote comparatively little attention to language assessment in any form, which perhaps helps to keep it in perspective as just one step in the process of tackling students' academic language needs.

2. Post-entry language assessment in Australia

Let us now turn to post-entry language assessment (PELA) in particular. As I mentioned earlier, some PELA initiatives predate the developments since 2006 which I have just outlined. The fullest published description of PELA in Australia to date comes from a survey conducted in 2008 on behalf of the Association for Academic Language and Learning (AALL) (www.aall.org.au) by Katie Dunworth of Curtin University in Perth (Dunworth 2009). She reported that 14 universities were using a PELA (four university-wide and ten within specific disciplinary areas), while another twelve institutions were considering introducing such an assessment. Reasons given by respondents for implementing a PELA included the following:

- to identify those who need English language support
- to maintain/improve English language levels
- to get students to understand the importance of English language competence
- to obtain data on student English levels. (Dunworth 2009: 5)

Table 1 summarises the key features of the PELAs, as reported by the respondents. It can be seen that the typical PELA was developed in-house at the institution and was paper-based.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Table 1 The different types of PELA in Australian universities (based on Dunworth 2009: 4)

Aspect	Frequency of particular types			
Source of design	In-house	8		
	External	3		
Delivery mode	Paper-based	7		
	Online	3		
Manner of delivery	Supervised	6	Compulsory	4
	Unsupervised	4	Optional	2
Availability	At specific times	6		
	Throughout the semester/year	4		
Content of PELA	Reading	7	Writing	10
	Listening	2	Speaking	0
	(Other: vocabulary 2, grammar 3, Australian culture 1, inference 1, spelling 1)			
Target cohort	All students	7		
	EAL* students	4		

*English as an additional language

It was compulsory for the students designated to take it, who could be just the students for whom English was an additional language (EAL), or all the first-year students in a faculty or the whole university. Some of the assessments were administered on fixed dates, whereas others were available on a more flexible basis. Of the four skills, writing and reading received by far the most attention; none of the PELAs included a speaking assessment. Although these 'typical' features could be identified, the assessments were actually quite diverse in nature.

Dunworth is currently directing a project which includes a more comprehensive and up-to-date survey of post-entry assessment in Australian universities. Helen Drury has presented here a work-in-progress session (Drury et al. 2012) on behalf of the project team and reported the following interim findings for 36 of the 39 universities in the country:

- Ten had university-wide PELAs in operation or being implemented.
- Seven had both university-wide and faculty-based PELAs in operation or being implemented.
- Ten had faculty-only PELAs (usually discipline-specific).
- Five were investigating or trialling PELAs.
- Four had no PELAs operating or under consideration.

At this point, it is useful to pause and consider how a PELA differs from an international English proficiency test such as IELTS or the Test of English as a Foreign Language (TOEFL). We can usefully distinguish them on three bases: purpose, scale and interpretation of the result.

- **Purpose**: The proficiency test is used for high-stakes decisions about whether students can be admitted to the university. Although institutional policies vary, students normally

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

need to achieve a specified minimum score in order to be allowed to enrol in a degree programme. The consequences include the creation of a whole industry of test preparation and a strong incentive for students to resort to various forms of cheating in order to obtain the necessary test result. However, a PELA is administered to students who have already been admitted to the university for degree study and who will not be excluded, no matter how poorly they perform in the assessment. The students may avoid taking the PELA if they are able to, but there is much less reason for them to cheat or to obtain a fraudulent result.

- **Scale**: The major English proficiency testing programmes are industrial-scale operations with worldwide networks of test centres and sophisticated systems to develop, administer, score, report and monitor the multiple forms of their tests that are required every year. In order to maintain the confidence of universities and other test users, the test publishers must preserve consistent standards in the results they report and ensure that the integrity of their score reports is not compromised by fraud. The high costs involved are largely met by the test-takers themselves through the substantial fee they pay each time they take the test. In contrast, a PELA is typically a local initiative within a single institution, which pays for the costs of the assessment programme from its own budget as part of a commitment to addressing the language needs of its incoming students. This means that the assessment can be designed and administered in a way that is appropriate for the students entering that university, or particular academic programmes within the university. On the downside, there may be a lack of expertise in language assessment within the institution, with the result that the quality of the instrument as a valid measure of student ability may be doubtful.
- **Interpretation of the results**: Although IELTS and TOEFL report both an overall result and individual scores for the four skills, the academic and administrative staff who process international student admissions in universities tend to focus primarily on whether each applicant has achieved the minimum scores set by the institution, without further consideration. However, a PELA offers at least the potential for richer diagnostic information on individuals or groups of students, which can be used when advising them on enhancing their academic language skills or building appropriate language support into their degree programmes.

These three characteristics serve to show how a PELA is a distinct form of assessment, with a specific role in contemporary English-medium university programmes.

3. DELNA at the University of Auckland

I would now like to shift the focus to New Zealand and discuss one particular PELA, the Diagnostic English Language Needs Assessment (DELNA), which has operated successfully at my own institution, the University of Auckland, for the last ten years. I can only give an overview of the programme here; for more details, see Read (2008) or the DELNA website, www.delna.auckland.ac.nz. I should also preface my discussion by acknowledging that, although I have been Academic Coordinator of DELNA since 2005, most of the credit for its design and implementation belongs to my predecessor in the role, Cathie Elder, and to

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Janet von Randow, the energetic, enthusiastic and multi-tasking manager of the programme from its inception. It was through Cathie that DELNA was established as a cooperative venture between Auckland and the Language Testing Research Centre at the University of Melbourne, a relationship which continues to this day through the joint development of new listening and reading tasks. DELNA has in fact built on DELA, the Melbourne assessment that I mentioned at the outset.

Whereas the impetus since 2007 towards introducing PELAs in Australian universities has come from concerns about the English proficiency of international students, the introduction of DELNA at Auckland grew out of more general concerns over the last 15 years about the increasing linguistic diversity of the student body. As the dominant metropolitan area of the country, Auckland has long been the favoured destination of immigrants to New Zealand, and from the 1990s on the university enrolled large numbers of young permanent residents, particularly from East Asian countries, with very similar linguistic and educational backgrounds to those of international students. Many students in both groups were found to struggle with the language demands of their degree programmes, even if they had entered the university through the normal matriculation pathway from secondary school rather than coming directly from their home country. However, while international students could be required to take an English proficiency test as a condition of admission, under New Zealand law permanent residents were classified as domestic students and could not be subject to such a requirement unless it also applied to New Zealand citizens.

Other groups were also recognised as facing challenges in the area of academic literacy. An earlier wave of immigration, beginning in the 1960s, led to the establishment of large communities in Auckland of Pasifika peoples, especially from Samoa, Tonga, the Cook Islands, Tokelau and Niue, and students from these communities received special consideration for admission on equity grounds. Increasingly, they were New Zealand-born and dominant in English, but their oral fluency was not matched by a high level of competence in academic reading or writing. Mature students constituted a further target group who, under another provision in New Zealand law, could apply for special admission to the university without having met the normal matriculation requirements.

Thus, after considering various possibilities, the University opted for an assessment system which would eventually be mandatory for all first-year undergraduate students, regardless of their linguistic or educational background. Its comprehensiveness in this respect is unusual, if not unique, among universities in Australia and New Zealand. For instance, although the University of Melbourne adopted a similar policy in 2008 (Ransom 2009), it has so far been much less fully implemented.

In order to accommodate the range of students – from highly proficient native speakers to those for whom English was an inadequately acquired additional language – the assessment was designed as a two-stage process, as represented in Figure 1.The first phase, known as the Screening, was intended to provide an efficient basis for identifying students who were highly competent in the language and not in need of targeted language support. The Screening is web-based and takes less than half an hour to administer. Elder & von Randow (2008) discuss the basis on which the cut scores for the Screening were set, seeking an optimum balance between identifying all those students whose language ability should be further assessed and

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

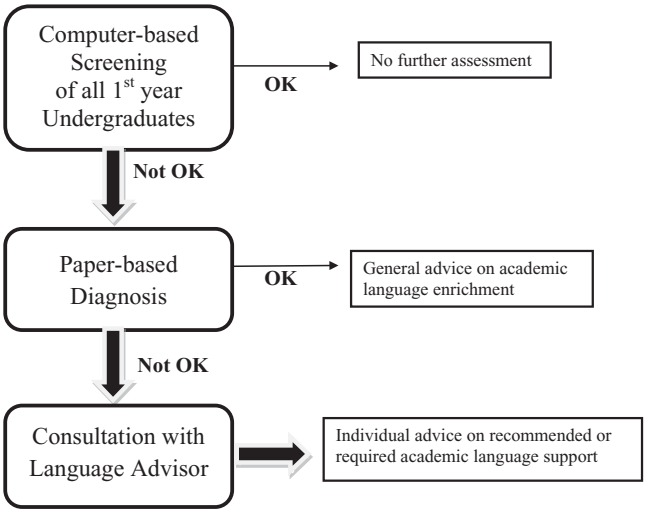


Figure 1 Overview of DELNA

Table 2 DELNA candidate statistics in 2012

SCREENING		
Total candidates	8099	
Undergraduate	7318	
Diagnosis required	2267	(28%)
DIAGNOSIS		
Total candidates	1337	(59%)
Undergraduate	1092	
Met with Language Advisor	793	(59%)

limiting, for budgetary reasons, the numbers directed to take the more expensive second phase.

Depending on the policy of their faculty or department, students placed below the cut score are either required or strongly recommended to return for the two-hour Diagnosis phase, a paper-based assessment of their academic listening, reading and writing skills. Those who obtain satisfactory scores receive generic advice by email about opportunities available on campus to enhance their academic language proficiency. Lower scoring students are called in for a meeting with the DELNA language advisor, who reviews their assessment results with them and informs them of the forms of language support they are required or recommended to take.

Table 2 gives the statistics for student participation in DELNA at Auckland in 2012. It is difficult to obtain exact figures for the total number of first-year undergraduates, but we estimate that 90% of them take the Screening, so to a large degree it has been embedded as

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

one of those things that new students need to do during Orientation. Postgraduate students also feature in the table, particularly because since 2011 the School of Graduate Studies has made it mandatory for all new doctoral students to take DELNA. Of those who took the Screening, a little over a quarter (28%) fell below the cut score and were supposed to take the Diagnosis, according to university policy. Compliance with this requirement is increasing, particularly in professional degree programmes where students face sanctions for not taking the assessment. In engineering, nursing and pharmacy, students cannot proceed to the second year of their degree without taking the Diagnosis and acting on any advice they subsequently receive to enhance their academic language skills. Students of education are barred from being posted to schools for the teaching practicum if they have not met the DELNA requirement. However, students taking arts or science degrees are not subject to penalties of this kind.

Table 2 shows that, overall, about 40% did not comply with the request to take the Diagnosis. Reviews of the academic records of these non-complying students by DELNA staff have shown that they are very likely either to drop out or to obtain very low grades in their first year of study. According to the Student Counselling Service, such students may be under stress for a variety of reasons, and an unjustified fear that they will be excluded from the university on the basis of a poor DELNA result is just one factor. Conversely, those 800 or so students who complete the Diagnosis and then meet with a language advisor almost always report it as being a very positive experience for them.

Options for academic language enhancement at Auckland have not been planned or resourced quite as comprehensively as the DELNA programme. Nevertheless, there is a range of offerings, which largely predate the introduction of DELNA.

- Two university-wide resource centres: Student Learning offers support for academic learning through workshops, support groups and one-to-one tutorials; English Language Enrichment (formerly the English Language Self-access Centre) provides online and print resources, advisory sessions and language-learning groups for students with English as an additional language (EAL).
- Credit courses: in academic English (covering all four skills) and English writing, offered by the Department of Applied Language Studies and Linguistics; in scientific communication, taught by the Faculty of Science; and in technical communication for engineers, in the Faculty of Engineering.
- Discipline-specific tutorial support for EAL students, in academic programmes such as film, TV and media studies, business, law and theology.
- Structured self-study programmes for EAL students in engineering, nursing and pharmacy, and for doctoral students.

Whether participation in one or more of these language enhancement options is required or recommended for students is at the discretion of their academic programme, as is the vigour with which any policy decision is carried out. This raises the question of the effectiveness of the DELNA programme, which I will return to later.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Table 3 The DELNA assessment tasks

SCREENING	
Computer-based	(30 minutes)
Vocabulary	
Timed reading (cloze-elide)	
DIAGNOSIS	
Paper-based	(2 hours)
Listening to a mini-lecture	
Reading academic-type texts	
Writing an interpretation of a graph	

4. Issues related to PELAs

4.1 The underlying construct

I would like to move on to discuss three issues related to post-entry language assessments such as DELNA. The first is the question of the construct on which they are based. This is important in clarifying what needs to be measured and in deciding on the appropriate assessment formats. I will discuss three relevant constructs: academic language competence, academic language proficiency and academic literacies.

To help illustrate the nature of these three constructs, let me engage in a kind of reverse engineering by first discussing the assessment tasks in the two phases of DELNA, as summarised in Table 3.

4.1.1 Academic language competence

The Screening phase consists of two tasks. The first measures knowledge of academic vocabulary, based on the ability to match academic words presented in isolation with their definitions. It is a computer-based version of the test developed by Beglar & Hunt (1999) from Nation's widely used Vocabulary Levels Test (Nation 1990: 264–272). The second task is known as Timed Reading within DELNA, but in the language testing literature as the cloze-elide procedure (Davies 1975, 1990; Manning 1987). In this task, the candidates are presented with an academic-style text with one extraneous word in each line; they respond by clicking on the added word. It is a speeded task, such that only the most competent candidates are able to work through the 73 lines of text in the ten minutes allowed.

Clearly, neither of these is an authentic task in the sense of representing activities in which the students will engage during their academic studies. Rather, they are designed to measure knowledge of the structural system of the language and, in the case of the Timed Reading task, the speed of access to that knowledge. Within Bachman & Palmer's (2010) influential framework of communicative language ability, this knowledge is referred to as 'language knowledge' or 'language competence'. The argument is that students with high scores in the

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Screening have a high level of language competence, and we can infer from this that they are likely to have an adequate level of academic literacy. As Elder & von Randow (2008) point out, there is a large element of pragmatism here, given the University's policy that all first-year students should take DELNA, regardless of their language background. In functional terms, the Screening tasks need to provide an efficient, low-cost means of exempting students who are expected to cope with the language demands of their academic studies and targeting those whose language ability requires further investigation. Thus, the construct underlying the DELNA Screening can be called ACADEMIC LANGUAGE COMPETENCE.

4.1.2 Academic language proficiency

Moving on to the second phase, the Diagnosis, Table 3 shows that there are three tasks, covering the skills of listening, reading and writing. These are conventional assessment tasks of the kind that can be found in communicative-style tests of English for Academic Purposes developed in the 1980s and 1990s, such as IELTS or the Test of English for Educational Purposes (TEEP) (www.reading.ac.uk/islc/english-language-courses/english-language-tests/islc-teep-about-teep.aspx). The tasks can claim to represent academic study activities, within the confines of a two-hour, paper-based assessment. In addition, they provide a basis for differentiating the students' ability across the three skill areas, in handling spoken and written input and written output, respectively. This means that they can be said to measure the construct of ACADEMIC LANGUAGE PROFICIENCY, or perhaps ACADEMIC LITERACY. The most sophisticated current definition of this type of construct provides the foundation for the validity argument developed by Chapelle, Enright & Jamieson (2008) for the internet-based TOEFL (iBT).

One obvious feature of this form of assessment, and of the construct, is that it is generic in nature, in that all the candidates are assessed using the same input material, which is intended to be of general academic interest, without reference to the students' planned field of study. This has been the predominant approach in the teaching and testing of English for Academic Purposes, reflecting the practical reality that EAP programmes are typically composed of students preparing to undertake studies in a wide range of disciplines and it is uneconomic to cater separately for specific disciplinary groupings. The same applied to IELTS, which in the early years offered a choice of three broad disciplinary modules in the reading and writing sections of the test, which were abandoned for various reasons after 1995 in favour of a single academic module for all candidates (Clapham 1996; Davies 2008, Chapter 5).

4.1.3 Academic literacies

Nevertheless, theory and research in EAP has increasingly focused on how academic disciplines form distinct discourse communities with their own preferred norms of language use which extend far beyond the obvious differences in terminology (see, for example, Hyland 2004). The leading literacy scholar Brian Street is critical of the 'autonomous' model of literacy

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

that treats reading and writing as a set of generic cognitive skills without reference to their context of use. He and Mary Lea have shown how undergraduate students taking courses in different disciplinary areas often struggle to understand their instructors' expectations with regard to academic writing in their courses, and the instructors themselves have difficulty in articulating the discourse norms of their discipline (Lea & Street 1998, 1999).

The concept of ACADEMIC LITERACIES has therefore become popular as an expression of this perspective. To illustrate how it can be operationalised as a construct, it is useful to introduce an assessment procedure which originated right here at the University of Sydney: Measuring the Academic Skills of University Students (MASUS) (Bonanno & Jones 2007). The procedure was developed by advisors at the University's Learning Centre, working in collaboration with the teaching staff of particular academic programmes. To quote from the introduction to the MASUS resource document:

> MASUS is a diagnostic assessment instrument designed to measure students' academic literacy. The whole MASUS procedure enables the creation of a literacy profile of a student cohort in terms of a set of specified criteria. This profile gives staff guidance as to where they can most effectively focus their resources in order to develop their students' literacy and generic skills. The procedure also identifies students who are at risk of failing their course because of poor literacy skills, thus enabling support to be made available to them. It has the potential to be integrated into the course curricula and assessment structure, and could be used as framework for individual student literacy profiles. (Bonanno & Jones 2007: 1)

The assessment is based on an integrated writing task (see Plakans 2012 for a discussion of integrated tasks in language assessment), which is tailored to the students' field of study. A suitable topic is formulated by the teaching staff in the programme, who also select resource materials to provide the students with background knowledge of the topic. The students write a short essay evaluating the validity of a controversial statement about the topic, and the essay is assessed using four main criteria:

- information retrieval and processing of the input material
- structure and development of the text
- control of academic style
- grammatical correctness.

For example, in an undergraduate course on design architecture, the students were taken on a tour of the campus by academic staff, who pointed out different architectural styles and uses of open space. The students also had a library session working in groups to search for references on the history of architectural styles in Australia. The third kind of input was information about referencing conventions in architecture. For the writing task, the prompt was:

> Identify two or three architectural styles present on the main campus of the University of Sydney.
>
> Do you think that having so many different architectural ideals reflected in the buildings and spaces on the campus gives an unharmonious effect?

The writing was assessed according to the above criteria by Learning Centre staff and this led to three forms of follow-up action. Students were referred to workshops at the

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Table 4 A comparison of some key features of DELNA and MASUS

DELNA	MASUS
• Focus on academic language proficiency	• Focus on academic literacies
• Centrally designed and administered for the whole institution	• Designed and delivered through specific academic programmes
• Generic assessment material	• Discipline-specific input material and tasks
• Variable individual and group opportunities for language enrichment	• Integrated with the teaching of particular courses

Language Centre as appropriate; some of the best essays were annotated and made available as exemplars of good writing in the field; and there were library sessions to develop the students' skills in information retrieval (Bonanno & Jones 2007: 10).

Thus, to return to the construct issue, Table 4 summarises the key features of MASUS in comparison with DELNA. The latter (at least in the Diagnosis phase) is based on academic language proficiency, as I have defined it, whereas MASUS clearly operationalises academic literacies, with its discipline-specific focus. If an assessment is to be applied uniformly across a whole institution, a generic approach like that of DELNA is the only feasible option. MASUS needs to be anchored in specific disciplines and requires the active involvement of academic staff, at least in the design of the writing task as well as in follow-up action. It is a model that works most effectively in professional programmes like pharmacy, accounting and law, rather than in general arts or science degrees. Students studying for professional degrees tend to form cohesive cohorts taking a common set of courses, and this makes it easier to deliver targeted language support to them. Of course, as in the DELNA case, an assessment based on academic language proficiency can also provide valuable information for staff teaching in professional programmes, but there is less integration between the assessment and the delivery of courses than in the MASUS model.

4.2 The diagnostic nature of a PELA

The second issue I want to tackle is the legitimacy of describing DELNA and other post-entry language assessments as 'diagnostic' in nature. As Alderson (2005) pointed out in his seminal book on the topic, diagnosis is a purpose of language assessment which has long received lip service from authors of introductory books on language testing but, in the absence of more than a handful of actual L2 instruments of this kind, the authors have been rather vague about the defining characteristics of diagnostic assessment. Alderson (2005) stimulated fresh interest in diagnosis through his work on developing and validating DIALANG (www.lancs.ac.uk/researchenterprise/dialang/about), the free, web-based system linked to the Common European Framework of Reference (CEFR) (Council of Europe 2001), which allows adult learners of 14 European languages to assess their proficiency in their chosen target language.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Drawing on Alderson (2005) and other recent sources (e.g. Huhta 2008; Kunnan & Jang 2009), I have compiled the following list of features of a diagnostic assessment.

- A focus on areas of difficulty and the reasons for poor performance
- An analytic approach to language ability
  - Language elements: vocabulary, grammar, phonology
  - Subskills in reading and listening; speaking and writing
- A basis for informed decisions about subsequent teaching and learning
- For adult learners, a process of raising awareness and encouraging learner autonomy.

Let us turn to the question of the extent to which DELNA is diagnostic. First, it must be acknowledged that the justification is somewhat post hoc, in the sense that the term was included in the title of the assessment (and the name of DELA at Melbourne even earlier) before the current focus on language diagnosis emerged. Nevertheless, there are several features of DELNA that can be regarded as diagnostic in nature. It has some overlap with proficiency and placement testing, but is also distinct from each of these more widely implemented purposes for language assessment.

- Like other PELA programmes, DELNA has the basic purpose of identifying students whose academic language proficiency may not be adequate AFTER they have been admitted to an academic degree programme, with a view to ensuring that the students' learning needs are addressed. This distinguishes it from a proficiency test, which simply determines whether students have achieved a pre-set standard of English proficiency as part of the admissions process. In addition, it does not function just as a means of placing students into the appropriate course or class in an English language or writing programme.
- The Screening phase of DELNA includes measures of language knowledge (vocabulary, timed reading) which are not normally used in contemporary proficiency tests, where the focus is on the test-takers' ability to perform simulated academic study tasks.
- In the Diagnosis phase, the listening and reading tasks are of the conventional type for assessing academic language proficiency, which means that they are not designed to measure specific sub-skills and the results are reported as a single band score. In contrast, the analytic rating scale for the writing task includes separate ratings for fluency, content and form and further sub-categories. The raters are also asked to record particular features of the candidates' writing which are useful as the basis for diagnostic feedback.
- As shown in Figure 1, students who perform satisfactorily in the Diagnosis phase of the assessment receive generic advice by email about opportunities available on campus to enhance their academic language skills. Those with low band scores who are considered at risk are called in for an individual face-to-face consultation with a language advisor, who discusses their results and pays particular attention to the writing sample. A number of academic programmes in the university now require these students to take a specific course or study option; otherwise, the advisor recommends appropriate courses or strategies for the students to address their academic language needs. This is therefore far more than just a placement procedure.

Overall, DELNA is most diagnostic in nature for those 800 or so students who score low on the Diagnosis phase of the assessment and who meet with the language advisor. However,

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

there is a good case for labelling it 'Diagnostic' in a broader sense, for the reasons I have just outlined.

4.3 The validity of a PELA

This brings me to the third issue related to DELNA: how do we evaluate the validity of a post-entry language assessment of this kind? This is a big topic in itself and, given the time constraints, I can present only the outline of a response.

The first priority, obviously, is to investigate the technical quality of the assessment in measurement terms. Even though the stakes are lower than in a major proficiency or achievement test, it is important that students should be correctly classified according to their results in the two phases of the DELNA assessment, and that they should be provided with well-founded diagnostic advice. The primary evidence for the quality of DELNA in these terms was obtained in the initial validation study conducted during the developmental stage (Elder & Erlam 2001), along with a subsequent investigation of the functioning of the Screening tasks in particular (Elder & von Randow 2008). This research has addressed matters such as the reliability of the measures, the correlations between Screening and Diagnosis scores, the relative performance of students from English-speaking and non-English-speaking backgrounds and the setting of appropriate cut scores.

In addition, internal comparisons of DELNA results and Grade Point Averages (GPAs) have repeatedly shown that candidates who perform poorly on DELNA are much more likely to obtain failing or low passing grades in their degree courses, or to drop out of the university altogether, during their first year of study. This applies in particular to students who are informed that they should take the Diagnosis after getting a low score on the Screening, but who avoid doing so unless their department or faculty enforces their compliance.

A more global perspective is offered by Davies & Elder (2005), who took DELNA as a case study to illustrate their discussion of current conceptions of validity in language assessment. They distinguish validity – an ideal (and ultimately unattainable) attribute of language tests – from the processes of validation, which seek to obtain evidence from multiple sources of what a particular assessment is measuring. In the case of DELNA, Davies & Elder formulated questions that we might ask about the assessment as the basis for developing a validity argument, linked to Messick's (1989) influential framework of test validity. The questions in turn gave rise to ten hypotheses that could be empirically tested by the gathering of appropriate evidence. In fact, as I have just noted and as Davies & Elder explain in more detail, most of the hypotheses have already been addressed to some extent through DELNA and DELNA-related research.

The most innovative dimension in Messick's framework was his focus on the consequences of test use, as part of a more general concern about the social and ethical aspects of educational measurement. Two of Davies & Elder's (2005) hypotheses focus on this aspect of validation:

> H.9 Students diagnosed as needing support will act on the advice offered.
>
> H.10 Taking time out from academic study to attend an instructed or self-access language program will benefit students found to be at risk.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

The authors highlight the problems in obtaining the relevant evidence or, in the case of H.10, in even defining what counts as evidence. The more we move away in time and space from the actual assessment, the harder it is to say how the results of the assessment might have affected the students' study activity or their ability to meet the language demands of their degree courses. Nevertheless, it is important to establish that the assessment is having a positive impact, first by raising students' consciousness of their language needs and informing them of the opportunities available to enrich their language skills, and then by tracking their engagement with language support courses and programmes.

Given the local, institution-specific nature of most PELAs, they do not have the resources either to formulate or gather the range of evidence for the kind of comprehensive validity argument that Chapelle et al. (2008) developed for the internet-based TOEFL (iBT). Nevertheless, it is possible to conduct survey research to elicit the opinions and experiences of students, lecturers and language tutors with regard to the effectiveness of the language enrichment activities that build on the assessment results. In fact, rather than attempting a global approach to the validation of the assessment, it may be more realistic to conduct small-scale evaluations with students in particular academic programmes, particularly if the institution has followed the MASUS approach, plus the recommendations of Arkoudis et al. (2012), and implemented discipline-specific initiatives that integrate language enrichment into the delivery of degree programmes rather than providing generic courses or workshops in academic English or study skills.

5. Conclusion

To sum up, post-entry language assessments represent an institutional response to the linguistic diversity of students entering our universities, and the need to identify those students with significant academic language needs. In the Australian context, it has become an academic audit requirement for universities to show that they are making a serious effort to undertake this task, and institutions have responded in a variety of ways, depending on local circumstances and the resources available. My own institution, Auckland, stands out among Australian and New Zealand universities in the extent of its commitment to this form of assessment and the comprehensive manner in which it has been implemented over the last ten years.

I have argued that, even though it may look superficially similar, a PELA has a purpose which is distinct from that of either a proficiency or a placement test. Actually, it is difficult to generalise because post-entry assessments take a number of different forms, reflecting the kind of decision-making required to guide or direct incoming students to appropriate means of enhancing their academic language ability that are available on each campus. If we seek to analyse the construct that underlies the assessments, it may variably be defined as language competence, academic language proficiency or academic literacies, or some combination of these. In this regard, a basic decision needs to be made as to whether (as with DELNA) the assessment will be generic in nature, or whether (as with MASUS) it will be tailored to the content and discourse norms of particular academic disciplines. There is also variation in the extent to which a PELA can be regarded as diagnostic in nature.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Looking more broadly at what contributes to the validity of post-entry assessment, we can see that consequential validity in the form of a positive impact on student learning is of particular importance. Ultimately, the validity of a PELA should be investigated as much in terms of its impact on student achievement as on its effectiveness in identifying students with language needs.

Postscript

A number of members of the audience took up my offer to send them a copy of the PowerPoint file for the speech. They indicated that my analysis of this type of assessment would be helpful in their own institution as they worked through the issues involved in addressing the academic language needs of their students. Since the conference, another more comprehensive resource has become available in the form of a website called Degrees of Proficiency: www.degreesofproficiency.aall.org.au. This is the main outcome of the research project of the same name funded by the Australian Government and directed by Katie Dunworth, which produced the figures referred to in the speech for the number of PELAs currently in operation or being implemented (Drury et al. 2012). The website provides valuable information and advice about post-entry assessment, but its scope is much broader than that. To quote from the site, it 'identifies approaches, strategies and practices that are intended to develop the capacity of all university students, whatever their language background, to use English effectively in their educational environments and beyond'. It includes links to the major English proficiency tests and language proficiency frameworks, information about English language policies and programmes at particular universities, guidelines for developing institutional strategies based on good practice principles, and a glossary of relevant terms and acronyms.

One final note: my off-the-cuff remark about Australian birds at the beginning of the speech turned out to be quite apposite in ways that I had not expected. The thank-you card I received from the conference committee at the end of my session featured a kookaburra on the front. And I now realise I WAS talking about emus in the guise of English-medium universities.

References

AEI (Australian Education International) (2007). Final report: Outcomes from a national symposium: English language competence of international students. https://aei.gov.au/research/Publications/Documents/NS_Outcomes_Syposium.pdf

Alderson, J. C. (2005). *Diagnosing foreign language proficiency*. London: Continuum.

Arkoudis, S., C. Baik & S. Richardson (2012). *English language standards in higher education: From entry to exit*. Camberwell, Vic: ACER Press.

AUQA (Australian Universities Quality Agency) (2009). *Good Practice Principles for English language proficiency for international students in Australian universities*, Report to the Department of Education, Employment and Workplace Relations, Canberra. www.auqa.edu.au/files/otherpublications/good%20practice%20principles%20for%20english%20language%20proficiency%20report.pdf

Bachman, L. F. & A. S. Palmer (2010). *Language assessment in practice*. Oxford: Oxford University Press.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Beglar, D. & A. Hunt (1999). Revising and validating the 2000 Word Level and University Word Level Vocabulary Tests. *Language Testing* 16.2, 131–162.
Birrell, B. (2006). Implications of low English standards among overseas students at Australian universities. *People and Place* 14.4, 53–64. Melbourne: Centre for Population and Urban Research, Monash University.
Bonanno, H. & J. Jones (2007). *The MASUS procedure: Measuring the academic skills of university students. A resource document*. Sydney: Learning Centre, University of Sydney. http://sydney.edu.au/stuserv/documents/learning_centre/MASUS.pdf
Chapelle, C. A., M. K. Enright & J. Jamieson (eds.) (2008). *Building a validity argument for the Test of English as a Foreign Language*. New York: Routledge.
Clapham, C. (1996). *The development of IELTS: A study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
Davies, A. (1975). Two tests of speeded reading. In R. L. Jones & B. Spolsky (eds.), *Testing language proficiency*. Washington, DC: Center for Applied Linguistics, 119–130.
Davies, A. (1990). *Principles of language testing*. Oxford: Blackwell.
Davies, A. (2008). *Assessing academic English: Testing English proficiency 1950–1989: The IELTS solution*. Cambridge: Cambridge University Press.
Davies, A. & C. Elder (2005). Validity and validation in language testing. In E. Hinkel (ed.), *Handbook of research on second language learning*. Mahwah, NJ: Erlbaum, 795–813.
Drury, H., K. Dunworth, C. Kralik, T. Moore & D. Mulligan (2012). Degrees of proficiency: Post-entry English language assessment in Australia's universities. Work-in-progress session presented at the Inaugural Conference of the Association for Language Testing and Assessment of Australia and New Zealand, University of Sydney, 8–10 November.
Dunworth, K. (2009). An investigation into post-entry English language assessment in Australian universities. *Journal of Academic Language and Learning* 3.1, 1–13.
Elder, C. & R. Erlam (2001). *Development and validation of the Diagnostic English Language Needs Assessment (DELNA)*. Unpublished final project report. Department of Applied Language Studies and Linguistics, University of Auckland.
Elder, C. & J. von Randow (2008). Exploring the utility of a web-based English language screening tool. *Language Assessment Quarterly* 5.3, 173–194.
Huhta, A. (2008). Diagnostic and formative assessment. In B. Spolsky & F. M. Hult (eds.), *The handbook of educational linguistics*. Malden, MA: Blackwell, 469–482.
Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor, MI: University of Michigan Press.
Kunnan, A. J. & E. E. Jang. (2009). Diagnostic feedback in language assessment. In M. H. Long & C. J. Doughty (eds.), *The handbook of language teaching*. Oxford: Wiley-Blackwell, 610–625.
Lane, B. (2012). National regulator sharpens focus on English language standards. *The Australian*, 22 August. www.theaustralian.com.au/higher-education/national-regulator-sharpens-focus-on-english-language-standards/story-e6frgcjx-1226455260799
Lea, M. R. & B. Street (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education* 23.2, 157–172.
Lea, M. R. & B. Street, B. (1999). Writing as academic literacies: Understanding textual practices in higher education. In C. N. Candlin & K. Hyland (eds.), *Writing: Texts, processes and practices*. London: Longman, 62–81.
Manning, W. H. (1987). *Development of cloze-elide tests of English as a second language*. TOEFL Research Reports, No. 23. Princeton, NJ: Educational Testing Service.
Messick, S. (1989). Validity. In R. L. Linn (ed.), *Educational measurement* (3rd edn). New York: American Council on Education and Macmillan, 13–103.
Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
Plakans, L. (2012). Writing integrated items. In G. Fulcher & F. Davidson (eds.), *The Routledge handbook of language testing*. London: Routledge, 249–261.
Ransom, L. (2009). Implementing the post-entry English language assessment policy at the University of Melbourne: Rationale, processes and outcomes. *Journal of Academic Language and Learning* 3.2, 13–25.

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press

Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes* 7.2, 180–190.

JOHN READ is Associate Professor in Applied Language Studies at the University of Auckland, New Zealand, where he specialises in research on the assessment of English for academic and professional purposes and on second language vocabulary. He is the author of *Assessing vocabulary* (Cambridge, 2000) and is currently writing a book on post-entry language assessment. He is Immediate Past President of the International Language Testing Association (2013–14) and a member of the Committee of Examiners for the Test of English as a Foreign Language (TOEFL).

https://doi.org/10.1017/S0261444813000190 Published online by Cambridge University Press