POST-ENTRY ENGLISH LANGUAGE ASSESSMENT AT THE UNIVERSITY OF AUCKLAND: ONGOING VALIDATION OF DELNA

Rosemary Erlam & Morena Botelho de Magalhães

University of Auckland

Abstract

With the increase of globalization, there has been greater demand for education in English medium universities. However, many universities have been ill-equipped to address the challenges of educating students from diverse language backgrounds, many of whom do not have the level of academic English language likely to allow for success. One way of addressing this issue is to set up a PELA, a post-entry language assessment programme, to identify students who might be at risk and direct them to appropriate support. This paper focuses on DELNA, one of the most comprehensive PELAs in Australasia, introduced to the University of Auckland in 2001 (Elder & Erlam, 2001; Read, 2008). Read (2015a) has previously presented evidence in support of the validity of DELNA but has also highlighted the need for ongoing validation. This paper describes a small-scale study, as part of a larger project, coinciding with the twenty-year anniversary of the implementation of DELNA. It investigates preliminary evidence that might support the ongoing decisions and consequences that ensue from the use of DELNA. Potential areas for improvement and the need for further investigation of other aspects of validity, in order to establish that DELNA remains fit for purpose, are highlighted.

Keywords: PELA; diagnostic assessment; language support; validation

Introduction

Universities in English speaking countries have been, for some time now, facing the challenge of dealing with student populations that are increasingly linguistically diverse. The active recruitment of international students, increasing immigration, and the adoption of policies that aim to increase participation from previously under-represented sectors of society have all been identified as contributing factors (Read, 2016). The result is that it cannot be assumed that students are equipped to cope with the English language and literacy demands of study at tertiary level. One initiative that an institution can adopt to cope with this challenge is to set up a programme to assess the language needs of those entering

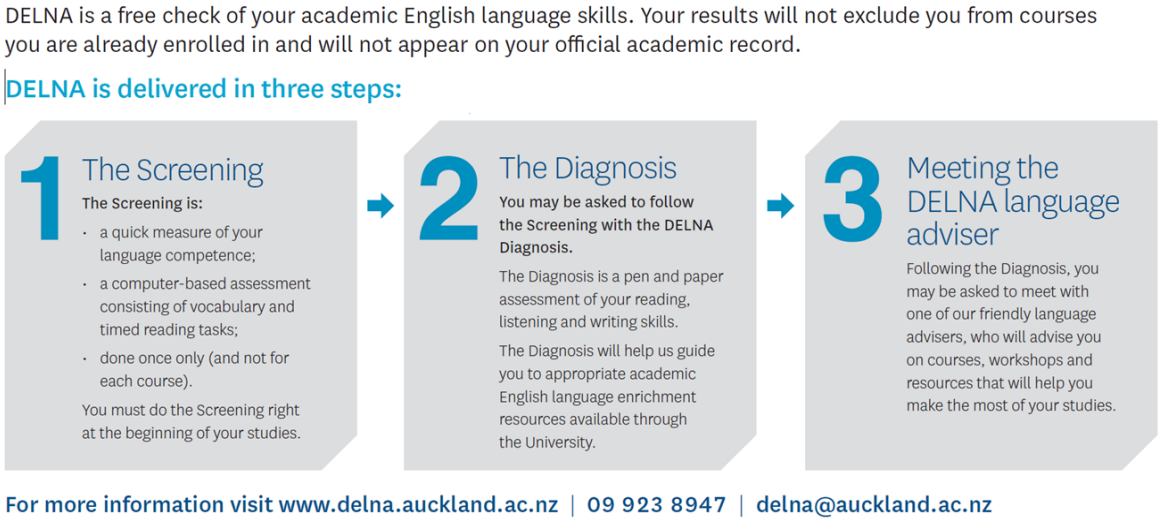
university and guide them to appropriate support. The term post-entry language assessment (PELA) is commonly used for this type of programme.

DELNA

In Aotearoa New Zealand, the University of Auckland (UoA) has been the only university to adopt a comprehensive PELA, set up in 2001 (Elder & Erlam, 2001), and known as DELNA (Diagnostic English Language Needs Assessment). Original principles under which the DELNA programme was established were that it was not to be used to determine admission to the UoA, that it would be seen as an assessment rather than a test, that it would be free of charge and that engaging with DELNA would involve personal choice (Read, 2008). While the first three principles still hold, DELNA has evolved so that it is now a requirement for many first-year undergraduate students (depending on faculty), but strongly recommended for all, and obligatory for all doctoral students (domestic and international). Other postgraduate students are also encouraged to complete the assessment programme.

There are three main components to DELNA as outlined in Figure 1 below, taken from an information card which is given to students.

Figure1. Description of the components of DELNA



Note: This card is now outdated given that it does not mention that the Diagnosis can be taken online.

The first component is an online Screening assessment consisting of a vocabulary task (7 minutes) and timed reading (cloze elide) task (10 minutes). Scores on this Screening assessment are reported in three categories: good, satisfactory or diagnosis required. The 'good' category indicates that students' English language

ability should be sufficient for university study and that students are able to seek support independently if need be. Students in the 'satisfactory' category would potentially benefit from some support and are sent an email encouraging them to engage independently with English language enrichment and other resources (https://www.library.auckland.ac.nz/about-us/student-learning-services).
Students who score below the minimum satisfactory band threshold fall in the 'diagnosis required' category and are asked to go on and complete the Diagnosis. The reader is directed to Elder and von Randow (2008) for a more detailed discussion of the rationale and process of setting cut scores for the Screening assessment. Elder and von Randow explain how the minimum score that would recommend further Diagnosis was initially set and subsequently lowered for operational reasons.

The Diagnosis assessment consists of listening, reading and writing tasks (either in pen and paper or online format). The listening and reading tasks are developed in conjunction with the Language Testing Research Centre at the University of Melbourne, whose DELA (Diagnostic English Language Assessment) both predated and set the foundation for the development of DELNA. In the listening and reading, various subskills are assessed, and a range of question types is used, e.g., summarising, matching ideas, information transfer, multiple choice, true/false and short answer. The writing tasks are developed in-house, with a long writing task for doctoral candidates and a short writing task for other students. In the short writing task, for instance, students are given information in a table or diagram and asked to produce between 200 and 250 words of commentary, following specific instructions (more information on the Diagnosis and examples of tasks can be found in the DELNA Handbook for students: https://cdn.auckland.ac.nz/assets/delna/delna/delna-handbook.pdf).

The results of the DELNA Diagnosis are converted into 'bands', an overall band, as well as individual bands for each of the three tasks. The bands determine to what extent the student is given information about and/or advised to take advantage of language support. The bands correspond to the following categories:

- 8 & 9, unlikely to require additional support
- 7, may benefit from some support
- 6, needs concurrent academic instruction
- 4 & 5, at severe risk and in need of urgent support

All students who score at band 6 or below are requested to see a DELNA language advisor. Students who gain an average overall band of 7 or above, including a minimum of band 7 for the writing task, receive their diagnosis results by email. However, there are exceptions. The Faculty of Education and Social Work requires all students who take the Diagnosis to meet with an advisor, for example,

and the School of Graduate Studies requires all doctoral students sitting the Diagnosis to do the same. In addition, for some students in other programmes, such as Engineering, seeing a DELNA language advisor is a requirement rather than a recommendation for those who do not achieve the minimum band 7 in writing.

All students who meet with a language advisor have the opportunity to go through their DELNA language profile and discuss their results with the advisor, as well as learn about the support options that are available. Many are recommended to enroll in an academic English language course, and for some students this is a faculty requirement.

A PELA programme must meet the needs of the institution it serves. It should adequately identify those who have significant academic language needs and facilitate that they obtain adequate support. Drawing on the Good Practice Principles for English Language Proficiency for International Students in Australian Universities (AUQA, 2009), Read (2016) identifies three principles which would specify practice for a PELA. These are that:

1. universities should take responsibility for ensuring students are sufficiently competent in English to participate effectively in their studies
2. English language needs should be diagnosed and addressed early
3. students should be responsible for further developing their English proficiency

DELNA embodies these three principles. The fact that it was set up at the UoA, with ongoing support for students assessed as 'at risk', is evidence of the University assuming responsibility for the English language needs of students. DELNA also aims to assess students as early as possible, with a major focus on making the DELNA screening process available to students during Orientation week. Finally, DELNA also places overall responsibility on students for taking the initiative to develop their English language proficiency (we will return to this point in the discussion below). At the same time, there remains a need to continue to collect evidence that would support the ongoing use of DELNA at the UoA and demonstrate to key stakeholders that it is functioning as intended, or alternatively, to suggest improvement. This is perhaps even more the case, given that DELNA has been in place at the University of Auckland for 20 years. In other words, ongoing validation of the DELNA programme is essential, to ensure that it continues to be 'fit for purpose'.

Validation of DELNA

In assessment theory, validity was originally seen as a property of a particular assessment or test. There was also an understanding that there were different types of validity. Over time, however, validity has come to be understood more as an integrated concept (Messick, 1989). Test validation can now be seen as the process of building an argument for the use of a test drawing on theoretical reasoning and empirical evidence (Read, 2015a). As part of the process of building an Assessment Use Argument, there is a need to justify that one can have confidence in the decisions made on the basis of test scores and the consequences of the test use (Knoch & Elder, 2013; Read, 2015a). Importantly, validation is understood as an ongoing process that continues even after the test has been designed and first used.

Drawing on validation research (e.g., Kane, 2012; Bachman & Palmer, 2010), Knoch and Elder (2013) proposed a framework against which PELAs can be evaluated. Knoch and Elder chose to represent this as an upside-down pyramid (see Figure 2). They stress that each part of the pyramid represents an inference and that these inferences function as bridges between "one building block of the argument and the next" (p. 51). At the bottom of the pyramid the focus of validation is on the characteristics of the items of the assessment, the way the assessment is administered and how it is scored. As we move up the pyramid attention is turned to how consistent results are across different versions of the assessment tasks and to how relevant these tasks are to the ways in which students use academic English in their university study. At the top of the pyramid the focus turns away from the assessment itself, to the relevance and appropriateness of decisions made on the basis of assessment results.

While all components of the pyramid must be 'functioning adequately' or be 'fit for purpose' for an assessment to be valid, it takes time and a large amount of research evidence to be able to establish validity at all levels of the pyramid. Drawing on Kane's work (1992), Knoch and Elder (2013) characterise validation as consisting of building a series of arguments to support the ongoing use of the assessment practice. Each argument operates at one of the levels of the pyramid, and for each argument there are a series of warrants and supporting evidence.

In this study, it is the top two components (bolded in Figure 2) of this pyramid that are the focus. While arguments at all the levels contribute to the validation process, it is important that the decisions that are made about test takers as a result of participating in a PELA be appropriate and the consequences beneficial. However, before we outline the inferences and warrants that characterize the decisions and consequences claims, we will first give an overview of existing evidence for the validity of DELNA at the first three levels of the pyramid. This

information draws on Read's work (2015a), which assembles validation evidence drawing on a range of studies conducted over the 14 years since the implementation of DELNA at the UoA. Table 1 outlines a summary of the information Read gives to support the warrants associated with the inferences of the first three stages of the pyramid. We refer the reader to Read (2015a) for a complete explanation and discussion of the process of validating DELNA, including a presentation of the warrants for each inference and, at the same time, information about where there is a need for additional evidence to establish that DELNA is fit for purpose.

Figure 2. Building blocks and inferences of a PELA interpretive argument (Knoch & Elder, 2013, p. 54)

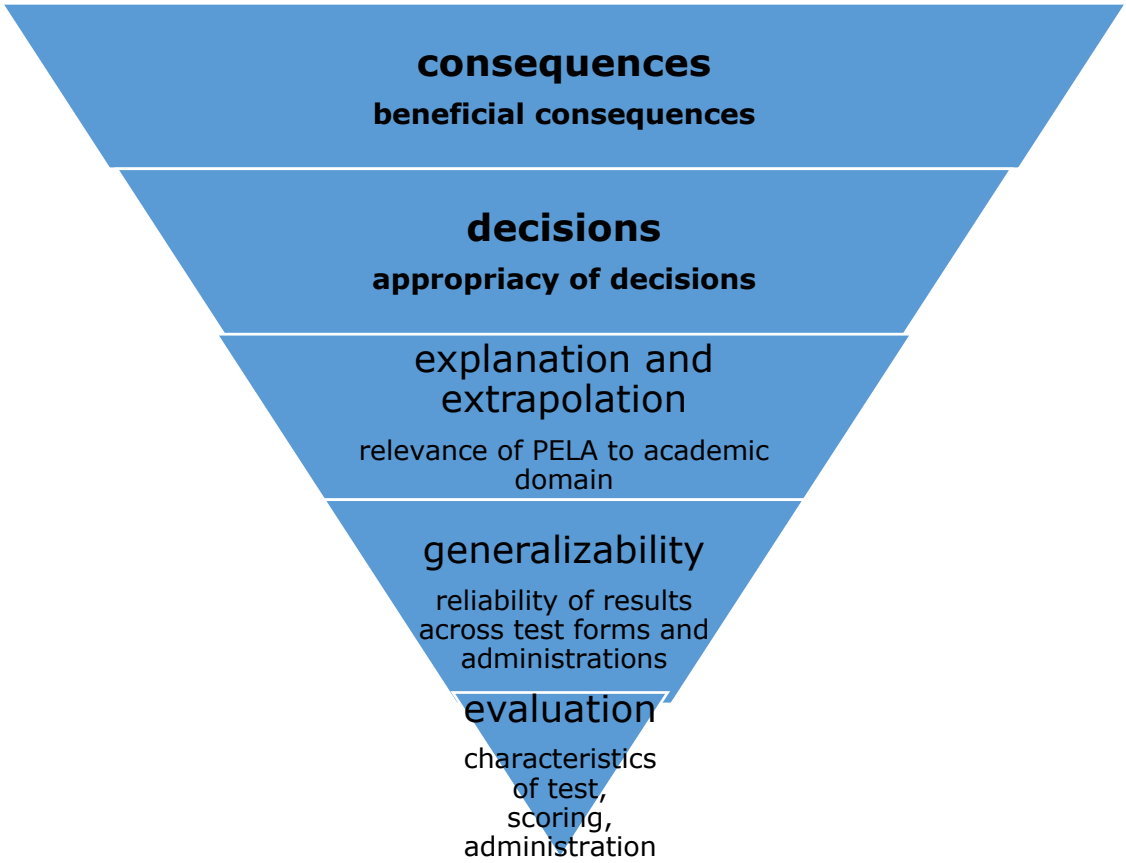


Table 1. Evidence for validity of DELNA at first three levels of the pyramid

Evaluation inference Claim: The score on the test is an adequate reflection of the observed test behavior (Knoch & Elder, 2013, p.56). Evidence: Training for raters; construction & validation of a new rating scale for writing (Knoch, 2009a, 2009b).

Student access to information about DELNA & practice material (Read, 2008).

Good item discrimination & reliability scores for screening and diagnosis assessments (Elder & Erlam, 2001; Elder & von Randow, 2008).

Generalization inference

Claim: The assessment yields results that are consistent across assessment contexts (Knoch & Elder, 2013, p.56).

Evidence: New versions of the Screening assessment tasks are modelled closely on the existing ones.

New versions of the Diagnostic assessment tasks are trialled, revised based on item analysis, reliability & Rasch analysis & finally, retrialled (Read, 2015a).

Explanation and extrapolation inferences

Claim: The assessment provides information on test-takers' skills that is in keeping with understanding of academic language proficiency and relevant to the academic domain (Knoch & Elder, 2013, p. 54).

Evidence: Twenty-three percent of students with DELNA bands (4 or 5) fail first year courses and 9% of those with higher bands (6 or 7) (Elder et al., 2007).

An integrated reading and writing task developed for first year doctoral candidates, as a more appropriate measure of writing at this level (Read, 2015a).

Given that it is the decisions and consequences arguments (the top two levels of the pyramid in Figure 2) that are the focus of the present study, we outline, in Tables 2 and 3, some of the warrants that Knoch and Elder (2013) propose as a basis for establishing validity at these two levels, along with the type of data that would be considered supporting evidence. At the same time, we include any existing DELNA evidence for these warrants, indicate those which are the specific focus of this study and those for which evidence remains yet to be collected.

The argument that needs to be made to support the *decision* inference in a PELA assessment is:

> Score-based decisions are appropriate and well communicated.

See Table 2 for potential and actual sources of evidence to support this argument.

Table 2. Warrant and supporting evidence for decision inference in a PELA validity argument

Warrant	Sources of supporting evidence	Evidence for DELNA
Students are correctly categorized based on their test scores	Review of academic outcomes for students classified above and below the cut-score	Low GPAs for students recommended to take Diagnosis but who avoid doing so (Elder & von Randow, 2008; Read, 2015a, 2015b). **Focus of present study: Research Question (RQ)1**
Test results include feedback on test performance and a recommendation	Review of policy and practice	Comprehensive feedback given to those most at risk by language advisor (Read, 2015a).
Assessment results are distributed in a timely manner	Review of practice	Screening results communicated within 24 hours, Diagnosis results within a week (Read, 2015a).
Test users understand the meaning and intended use of scores	Interviews with test users	**Evidence yet to be collected**

For a *consequences* inference, the following argument needs to be made:

> The consequences of using the PELA and the decisions informed by the PELA are beneficial to all stakeholders.

See Table 3 for warrants and evidence to support this argument. While, as we have seen in Tables 1 - 3, there is already data that builds an argument for the use of DELNA, there is need for ongoing investigation. Read (2015a) makes the case for further investigation of the decision and consequences inferences. The present study continues to investigate these two inferences and is part of a more comprehensive project that aims to investigate DELNA, drawing on interview data from students, language advisors and academic staff. The specific research questions that this, smaller scale, study addresses are presented below.

Table 3. Warrant and supporting evidence for the consequences inference in a PELA validity argument

Warrant	Sources of supporting evidence	Evidence for DELNA
All targeted test takers sit for the test	Analysis of test administration data	Compliance for Screening around 90% but lower for Diagnosis (around 60%) (Read, 2015a). **Focus of present study: RQ 2**
Students act on the test recommendation	Review of student uptake data	**Focus of present study: RQ 3**
Stakeholders' perceptions of the test and its usefulness are positive	Interviews with students & academic staff	Students report DELNA was helpful (Bright & von Randow, 2008). **More evidence yet to be collected**
Follow-up language development options provided for students are adequate	Interviews with key stakeholders, including students	**Evidence yet to be collected**

Research questions

1. What is the correlation between DELNA Screening and GPA?

If students are appropriately categorized according to their performance on the DELNA Screening, one would expect some relationship between their DELNA Screening scores and their Grade Performance Average (GPA) for the same year that they take the DELNA Screening.

2. What level of uptake is there by students who are identified as needing to take the Diagnosis?

One of the most obvious intended consequences of a PELA is that those students who are identified as potentially lacking the academic English language skills they need, do indeed take the assessment. However, convincing students, who have been identified by the Screening as needing to take the Diagnosis, to do so

has been a long-standing challenge for DELNA (Elder & von Randow, 2008; Read, 2015b) and continues to be so.

3. What uptake is there by students who are advised to meet with the DELNA language advisor?

Ensuring that those who are guided to seek follow-up support do indeed do so has, once again, been an ongoing issue for DELNA.

4. What factors might determine successful uptake of opportunity to meet with a language advisor?

In relation to the consequences inference, this question looks at what improvements might need to be made to improve uptake of support for those students identified as at risk.

Results and discussion

This study draws on DELNA Screening and Diagnosis data from 2019 and 2020 (first-year undergraduate students only). We therefore first present data showing DELNA screening results from this cohort (see Table 4). We need to point out that data from 2020 cannot be assumed to be representative of other years given factors associated with COVID. We have, nonetheless, decided to include it on the premise that a more complete picture can be gained by displaying data over a two-year period. The exception is GPA data, which, for the current study, was only collated for the year 2019.

Table 4. DELNA Screening results for first-year undergraduate students in 2019 and 2020

Screening result	2019		2020	
Diagnosis Required	2,334	35%	1,800	32%
Satisfactory	1,057	16%	897	16%
Good	3,234	49%	2,923	52%
TOTAL Screenings	6,625		5,620	

The difference in the number of screenings completed between the two years is explained by the fact that fewer students in 2020 attended Orientation week, the period where most screenings are completed. This was because of COVID 19; by the end of February 2020, the border had already closed to students from China and domestic student presence on campus during Orientation week was also negatively impacted by news of the pandemic. For the first-year undergraduate cohort, compliance with DELNA screening was 84% in 2019 and 75% in 2020.

The first research question asked whether there was a significant correlation between DELNA Screening and GPA. This question addressed the decision inference, investigating to what extent we can have confidence that DELNA is correctly identifying, from the screening process, those students who need to complete a diagnosis.

There were 5,469 undergraduate students with DELNA Screening results in 2019, for whom GPA data was also available. The correlation between DELNA Screening results and student GPA scores was statistically significant, $r = .33$. This result suggests a modest correlation between performance on the DELNA Screening and overall academic achievement, irrespective of the subject area and the language demands of programmes. As Read (2015a) documents, this finding is consistent with the results of studies investigating the validity of large-scale international proficiency tests.

Supporting evidence for decisions made on the basis of the DELNA Screening assessment is obtained by investigating academic outcomes for students classified as below and above the cut-score which determines whether students need to proceed to the Diagnosis assessment. Figure 3 shows GPA performance for students classified as below this score and thus advised to take the Diagnosis. There is a positive skewing of the GPA results of these students identified as potentially lacking adequate language skills and needing to take the DELNA Diagnosis. As in Elder and von Randow (2008), this data suggests that these students are at higher risk of failing courses.

Figure 3. Distribution of GPAs for Screening Band Diagnosis Required

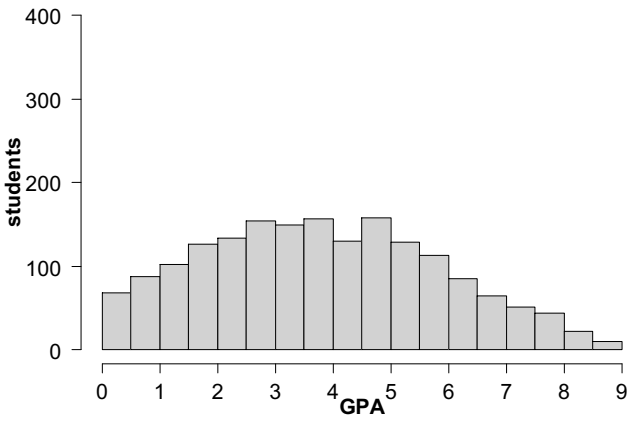
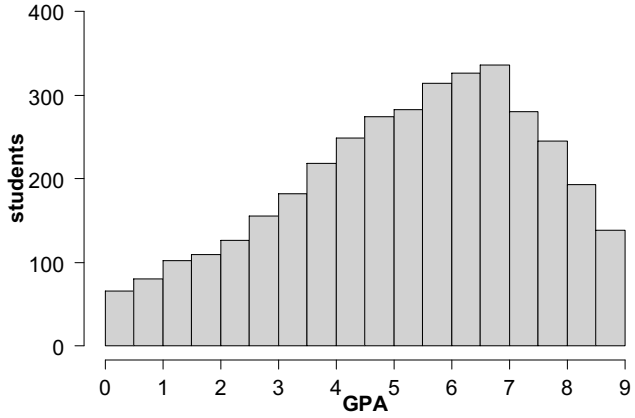


Figure 4 shows GPA performance for students classified as above the Screening cut-off score, and thus not needing to take the Diagnosis. There is a negative skewing of the GPA results of these students indicating that these students tend to perform well academically.

Overall, these results suggest that the Screening continues to identify those students who are likely to be at risk and who require further, more comprehensive, assessment.

Figure 4. Distribution of GPAs for Screening Bands Satisfactory or Good (i.e., above the cut-off score)



Research questions two and three address the issue of consequences in building an argument for the validity of DELNA. Question two asks what level of uptake there is for students who are identified as needing to take the Diagnosis.

In Table 5 we can first see that the level of uptake for the DELNA diagnosis was consistent across both years at 61% and 62%. This is very similar to the level of compliance of "around 60 per cent" that Read (2015, p. 215) reports, showing that the level of compliance remains consistently lower than ideal and is something that DELNA has historically contended with (Elder & von Randow, 2008). We will return to this issue when we look at data below which demonstrates that there is considerable variability between faculties in turns of compliance with the Diagnosis.

Table 5. Uptake of DELNA diagnosis and advisory sessions by first-year students

2019			2020		
DELNA diagnosis					
Diagnosis required	Diagnosis completed	compliance	Diagnosis required	Diagnosis completed	compliance
2334	1435	61%	1800	1111	62%
DELNA advisory session					

Advisory session required	Advisory session attended	Compliance	Advisory session required	Advisory session attended	Compliance
1373	873	63%	1107	800	72%

Research question three asks about the level of uptake that there is for students who are identified by the Diagnosis as advised to meet with the DELNA language advisor. In looking at the figures for attendance at the DELNA advisory sessions in Table 5, it is interesting to note that there is considerable variation in uptake between the two years. Uptake of the opportunity to attend an advisory session was greater in 2020, at 72%, whereas the figure for 2019 was at 63%. It would seem that this also is attributable to COVID, with this time a positive consequence for DELNA! A possible reason is due to the fact that advisory sessions needed to be adapted for online provision in 2020, and students seemed more willing to participate in these online sessions. One reason for the uptake may be the flexibility that they allowed students, and also perhaps that they were perceived as less 'threatening' than a face-to-face session. On the other hand, the DELNA team have also proposed a rather different explanation; there was an impression that students were missing the personal contact that on-campus learning provided and so were more ready to engage in a one-on-one meeting with a DELNA advisor. Whatever the explanation, the DELNA team have continued, after COVID, to offer students opportunities to attend both Diagnosis and advisory sessions online in order to cater, as much as possible, for student preference.

Over the nearly 20-year period since DELNA was first implemented at the University of Auckland, faculties have chosen to engage with the DELNA process in different ways. While a detailed description of these differences is beyond the brief of this paper, it has come to the attention of the DELNA team that the specific initiatives of some faculties have been particularly effective, at least in terms of ensuring compliance with the DELNA process.

This information is useful in addressing the last research question, which asks what factors might determine successful uptake of opportunities to take the DELNA Diagnosis and meet with a language advisor. This question explores ways of strengthening the beneficial consequences of the use of this assessment procedure.

To explore this question, we present data in Table 6 which demonstrates different rates of compliance with the DELNA Diagnosis and advisory sessions across the University. Compliance means completion of the Diagnosis or attendance at an advisory session within the same year as the student took the Screening.

Table 6. First-year students' compliance across the University with completing the Diagnosis and attending the advisory session

Faculty	2019		2020	
	Completing diagnosis	Attending advisory session	Completing diagnosis	Attending advisory session
Arts	50%	47%	50%	60%
Business & Economics	47%	45%	47%	60%
Creative Arts & Industries	65%	62%	57%	58%
Education & Social Work	95%	69%	79%	71%
Engineering	95%	94%	94%	95%
Law	52%	52%	53%	60%
Medical & Health Sciences	77%	65%	74%	75%
Science	53%	47%	51%	63%
Total	61%	63%	62%	72%

In looking at the data in Table 6, it is clear that one faculty stands out in terms of compliance for both the Diagnosis and the advisory sessions. This is the Faculty of Engineering. For both years, students in this faculty demonstrated 94% or 95% compliance with both processes. The factor that may determine this higher level of uptake is the embedding of a compulsory non-credit bearing course which must be completed before students continue with their study beyond the first year. In order to have this course marked as completed on their academic transcripts, Bachelor of Engineering students must complete DELNA and, if required, complete self-study English language enrichment modules as recommended by the DELNA advisor. These modules are part of a course delivered by English Language Enrichment at the UoA, on the completion of which students must pass an exit test.

In second place for compliance with both processes in 2019 (95%, 69%) and for completing the Diagnosis in 2020 (79%), is the Faculty of Education and Social Work. This faculty has also embedded a compulsory zero-point credit bearing course which undergraduate students must complete before they complete their final practicum. However, compliance is lower than that of students in Engineering. This is most likely due to the fact that while students must complete this course and cannot graduate without it, there are no milestones along the way to ensure students meet the requirement in a timely manner (Keesing-Styles, 2020). This of course means that students may not get targeted English language support early on in their university study.

Another faculty which also has higher compliance is the Faculty of Medical and Health Sciences. This faculty has embedded non-credit bearing courses into three of six undergraduate programmes. Students in Nursing, Pharmacy and Medical Imaging must satisfy the requirements of a non-credit bearing course during their first year in the programme. They can only get credit for this course if they take the DELNA Screening and, if advised to do so, go on to complete the other stages of the DELNA process, including compliance with advice that they may be given. These results suggest that it is those faculties which make compliance with DELNA an integral part of their undergraduate programme requirements that have the highest rates of student 'buy-in'. Read (2015a, 2015b) also commented on how levels of participation were influenced by the extent to which faculties required compliance and imposed sanctions for non-compliance.

The Faculty of Engineering model is the most successful. Embedding a non-credit bearing course into the first-year undergraduate programme, which all students must complete (by complying with DELNA processes) before they move to the next year, is no doubt the reason for this success. On one level, it could be seen as potentially problematic when faculties mandate compliance in this way, in that one of the initial premises on which DELNA was based was that it would be students who would take responsibility for their ongoing English language proficiency development. On the other hand, students may arrive at UoA believing that they have the level of academic English language proficiency they need to succeed, solely on the basis of entry requirements. This may not, however, be the case. It is, arguably, important then that these students have the chance to have their English language needs identified in their first year of study. Making DELNA optional may mean that students who might not consider that they need support are not identified. Mandating initial compliance with DELNA does not, of course, mean that students need not take responsibility for their ongoing English language needs. In the experience of the language support staff at the UoA, it is crucially important and an ongoing challenge to make students, and indeed, some of the teaching staff, aware of the fact that it usually takes considerable time and effort for students to improve their language skills so that they are adequately equipped for academic study. There is, for example, the risk that students who are identified as having English language needs and who take courses that are recommended to them might wrongly assume that they have done all that they need to. Of course, even more important than the issue of compliance is the usefulness of the advice that students are given when they meet with a DELNA language advisor. It is crucial that this advice informs their future learning, and that the support options that are recommended to them are appropriate and help them make gains in their academic English (Knoch & Elder, 2013). These issues will be the focus of further research.

Conclusion

This study aimed to provide support for decisions and consequences that follow on from the administration of the PELA (Post Entry Language Assessment) that is in place at the University of Auckland, known as DELNA. Specifically, it provides some evidence to show that the Screening assessment, administered to first year students, continues to identify those who may lack adequate academic English language skills and who should take the Diagnosis.

Data, which shows student attendance at the Diagnosis assessment and advisory sessions, is mixed, and demonstrates that compliance with these processes continues to be a challenge for DELNA. Making these sessions available online, as a result of COVID, corresponded to higher levels of uptake. Online delivery of both of these will therefore continue to be an option that DELNA offers students; a demonstration of how DELNA continues to be flexible to adapt to student preferences and needs.

Data indicating student compliance with the Diagnosis and advisory sessions showed large variation according to faculty. Those faculties which took measures to make completing the DELNA process compulsory in the first year of study had the highest rates of compliance. Even more successful was the initiative that makes enrolment in the second year of study contingent upon this compliance. Enabling students to become aware of their academic English language needs early in their study at UoA, and helping them access appropriate support, is crucial to the success of any PELA. It also part of a process of helping students take responsibility for their ongoing academic English language needs.

Looking forward, this study demonstrates that it is important that DELNA, with the support of University management, encourages faculties to take initiatives to ensure that students engage with the DELNA process early in their University studies. It also highlights the importance of ongoing validation of the programme so that it is responsive to stakeholder needs. The next phase of this study will continue to look at the extent to which an argument can be built for the ongoing use of DELNA, in particular making space for perspectives that are missing in this stage of the project, that is, the voices of the students and other stakeholders.

References

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].

http://hdl.voced.edu.au/10707/208559

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.

Bright, C., & von Randow, J. (2008). *Missed opportunities: EAL students reflect on their first-year language experiences at an English-medium university*. Pacific Rim First Year in Higher Education Conference, Hobart, Australia. http://www.fyhe.com.au/past_papers/papers08/FYHE2008/content/pdfs/5d.pdf

Elder, C., Bright, C., & Bennett, S. (2007). The role of language proficiency in academic success: Perspectives from a New Zealand university. *Melbourne Papers in Language Testing*, *12*(1), 24–58.

Elder, C., & Erlam, R. (2001). *Development and validation of the Diagnostic English Language Needs Assessment (DELNA): Final report*. The University of Auckland, Department of Applied Language Studies and Linguistics.

Elder, C., & von Randow, J. (2008). Exploring the utility of a web-based English language screening tool. *Language Assessment Quarterly*, *5*(3), 173–194. https://doi.org/10.1080/15434300802229334

Kane, M. (1992). An argument-based approach to validation. *Psychological Bulletin*, *112*, 527–535.

Kane, M. (2012). Articulating a validity argument. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 34–47). Routledge.

Keesing-Styles, L. (2020). *EDSW DELNA analysis and recommendations* [Unpublished].

Knoch, U. (2009a). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, *26*(2), 275–304. https://doi.org/10.1177/0265532208101008

Knoch, U. (2009b). *Diagnostic writing assessment: The development and validation of a rating scale*. Peter Lang.

Knoch, U., & Elder, C. (2013). A framework for validating post-entry language assessments (PELAs). *Papers in Language Testing and Assessment*, *2*(2), 48–66.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). American Council on Education and Macmillan.

Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes*, *7*(3), 180–190. https://doi.org/10.1016/j.jeap.2008.02.001

Read, J. (2015a). *Assessing English proficiency for university study*. Palgrave Macmillan.

Read, J. (2015b). Issues in post-entry language assessment in English-medium universities. *Language Teaching*, *48*(2), 217–234.

https://doi.org/10.1017/S0261444813000190

Read, J. (2016). Some key issues in post-admission language assessment. In J. Read (Ed.), *Post-admission language assessment of university students* (pp. 3–20). Springer International Publishing. https://doi.org/10.1007/978-3-319-39192-2_1