

Facing up to literacy: perceptions and performance in a test of academic literacy for postgraduate students

A B S T R A C T Language proficiency and academic literacy tests such as the National Benchmark Test (NBT) and Test of Academic Literacy Levels (TALL) are already well established assessment instruments that are widely used at universities in South Africa to assess the literacy levels of first-entry students. A more recent initiative has been the institution of language testing at postgraduate level as a means of identifying students at risk of not completing their academic studies at that more advanced level. This article examines the face validity of the Test of Academic Literacy for Postgraduate Students (TALPS) amongst a cohort of postgraduate students at the University of the Free State and the perceptions of these students as to their own levels of academic literacy. A correlation is made with the students' actual performance in the TALPS as an initial step towards gaining a measure of understanding of the low levels of academic literacy of some postgraduate students. The results of the study show that, although most of the students consider the TALPS to be fair and accurate, there is a major discrepancy between the perceptions of their own academic literacy levels and their actual test performance. Possible reasons for the disparity are gleaned from the responses provided by the students in the survey questionnaire and an analysis of their test scores. It would seem that the academic literacy levels of students may not increase substantially during the undergraduate phase of study, a phenomenon that reflects back on language issues, course electives and undergraduate teaching and assessment practices.

Key words: academic literacy, language testing, face validity, test construct, postgraduate assessment

1. The assessment of academic literacy at postgraduate level

The literacy levels of both undergraduate and postgraduate students at tertiary institutions in South Africa are lower than required for academic success, largely as a result of the prevailing

conditions and standards in South African schools (Van Dyk & Weideman 2004a; Van der Slik & Weideman 2007, Bhorat & Oosthuizen 2008). An analysis of the throughput rates per faculty at the University of the Free State, calculated on the basis of information provided by the Higher Education Management Information System (HEMIS) over the five-year period 2006-2010 (DIRAP 2011), shows that on average the throughput rate for undergraduates has only been 16.62% and that for postgraduate students 33.27% (see Annexure A). While it may be accepted that under the current educational dispensation in South Africa a number of first-year students can be expected to display inadequate literacy levels for the purposes of academic study when they enter the higher education arena, it is disconcerting to find that these literacy levels may remain low during the undergraduate phase. Equally disturbing is the anomaly that students may be able to graduate at a tertiary institution with low levels of academic literacy and even be admitted to postgraduate study.

Language proficiency and academic literacy tests such as the National Benchmark Test (NBT) and Test of Academic Literacy Levels (TALL) are already well established assessment instruments that are widely used at universities in South Africa to assess the academic literacy levels of first-entry students. A more recent initiative has been the institution of language testing at postgraduate level (see Butler 2009) as a means of identifying students at risk of not completing their academic studies at that more advanced level. In this regard the contribution of Albert Weideman (e.g. Weideman 2011) and his role in the timely development and introduction of the Test of Academic Literacy for Postgraduate Students (TALPS) at the University of the Free State (UFS) in particular is to be acknowledged.

This article examines the face validity of the TALPS amongst a cohort of postgraduate students at the UFS and the perceptions of these students as to their own levels of academic literacy. A correlation is made with the students' actual performance in the test as an initial step towards gaining a measure of understanding of the low levels of academic literacy of some postgraduate students.

2. A functional framework for assessing academic literacy

Academic literacy as test construct is articulated in the TALPS through the specification of subcomponents and their operationalization into task types on the basis of the functional framework provided by Weideman and Van Dyk (see Weideman 2003a: xi; Van Dyk & Weideman 2004a, 2004b; Annexure D). Accordingly, language is not viewed restrictively in terms of sound, form and meaning assessed through the composite 'skills' of listening, speaking, reading and writing, but rather as a social instrument used to mediate and negotiate interaction within an academic context. An enriched, open view of language is adopted and what is meant by academic language ability is defined on the basis of what Blanton (1994: 8) refers to as "literate behaviours" that academic readers and writers should be capable of performing at tertiary level, or what Gee (1998: 57) calls the mastery or "full and effortless control" of language used in secondary discourses. This would include those abilities assessed in the original Placement Test in English for Educational Purposes (PTEEP), which was developed by Yeld and her associates at the University of Cape Town in 2000 as part of the Alternative Admissions Research Project (AARP) and which was the precursor of the National Benchmark Test (NBT) (Yeld *et al.* 2000, Cliff & Hanslo 2009).

Task types in the TALPS are closely aligned with the actual language tasks that postgraduate students are required to perform and are evaluated in terms of their ability to be productive, based on a quantitative system of measurement and the application of appropriate statistical procedures (Bachman 2004, Paltridge & Phakiti 2010). Authentic texts are used which engage students in tasks such as extracting information, processing it, and interpreting meaning beyond sentence level. The emphasis in academic literacy testing of this nature can be seen to fall on critical reading and analytical thinking, and tapping into several levels of cognitive processing as these are mediated through language.

3. Issues of validity and reliability

The confidence that may be placed in any language test is considered to be directly proportional to the evidence collected in the process to support the evaluation of the instrument's validity (Davies *et al.* 1999: 220). The latter refers to the systematic presentation of this evidence as a unity within a multiplicity of arguments setting out the relationship of the test to the definition of the ability being assessed (the construct). Rather than relying on a unitary view of validity as a comprehensive judgment founded on empirical evidence and theoretical rationales and related to the adequacy and appropriateness of inferences and actions that are based on test scores, Weideman (2009: 248) specifies certain constitutive and regulative conditions (figure 1) for language assessment, each of which is relevant and interrelated to the validation process and the TALPS in particular (also refer to Van Dyk 2010 and Rambiritch 2012).

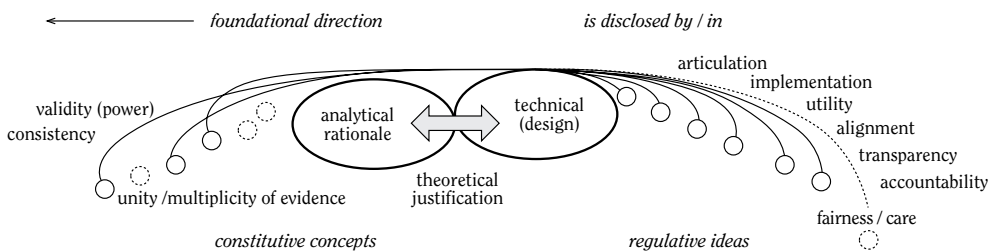


Figure 1: Constitutive concepts and regulative ideas in applied linguistic designs

In terms of the above representation, the theoretical justification for a language test such as the TALPS is to be found in the reciprocal relationship between the analytical and technical modes. The portrayed dimensions cannot be considered absolute and are mutually related. In language testing the technical (design) mode leads and qualifies the design of a solution to a language related problem, while the analytical dimension provides the foundational basis for the intervention. In the case of the TALPS, the theoretical rationale is to be found in the test construct referred to above, since it is with reference to this that evidence is systematically presented, and that scores are interpreted.

In addition to validity, reliability is considered to be the other essential variable when it comes to justifying using test scores for the purpose of making inferences. Reliability is referred to by Bachman and Palmer (1996: 19) as “consistency of measurement”. This implies that test scores may be deemed to be reliable if they remain consistent from one set of tests and tasks to another. Reliability is thus a function of score consistency between different tests and tasks.

Cronbach’s *alpha* is a commonly used reliability statistic that shows the “degree to which the observed scores represent the ‘true’ scores”, without measurement error (Van der Slik & Weideman 2005: 26). The Iteman 4.2 programme was used to analyse the results of a cohort of postgraduate students (n = 652) at the University of the Free State, who wrote the TALPS in the course of 2011, and to calculate the contribution of each item to the test as a whole, as reflected in the table that follows.

Table 1: Reliability indexes of the TALPS administered at the UFS in 2011 (Report 2012-02-16)

Score	Alpha	SEM	Split-Half Random	Split-Half (First-Last)	Split-Half (Odd-Even)	S-B Random	S-B First-Last	S-B Odd-Even
Scored items	0.931	3.549	0.835	0.654	0.876	0.910	0.790	0.934
Scrambled	0.822	0.779	0.670	0.670	0.667	0.802	0.802	0.800
Graph	0.827	1.192	0.673	0.634	0.698	0.805	0.776	0.822
Vocabulary	0.539	1.295	0.337	0.360	0.369	0.504	0.530	0.539
Text types	0.644	0.904	0.284	0.596	0.266	0.443	0.747	0.420
Understanding	0.784	1.794	0.470	0.614	0.666	0.639	0.761	0.800
Grammar	0.880	1.475	0.774	0.644	0.805	0.872	0.783	0.892
Text editing	0.891	0.924	0.751	0.722	0.759	0.858	0.839	0.863

Although the development of the TALPS is a recent initiative, its combined *alpha* of 0.931 is fully in keeping with that required by comparative international high-stakes tests, attesting to the quality and consistency of this locally designed artefact. A further analysis of the test scores of this cohort of 652 postgraduate students, reveals that the test items also have acceptable facility (Mean P) values and that they discriminate well between students of varying ability (Mean Rpbis), as evident in table 2 (for a fuller explanation see Guyer & Thompson 2011).

Table 2: Summary statistics of the TALPS as administered at the UFS in 2011 according to content domain

Score	Items	Mean	SD	Min Score	Max Score	Mean P	Mean Rpbis
All items	76	48.679	13.481	13	76	0.641	0.359
Scored Items	76	48.679	13.481	13	76	0.641	0.359
Scrambled	5	2.305	1.847	0	5	0.461	0.253
Graph	10	6.275	2.865	0	10	0.627	0.450
Vocabulary	10	7.086	1.908	1	10	0.709	0.246
Text types	5	2.436	1.515	0	5	0.487	0.214
Understanding	21	14.678	3.864	0	21	0.699	0.308
Grammar	15	8.549	4.252	0	15	0.570	0.458
Text editing	10	7.351	2.801	0	10	0.735	0.463

Now that the construct validity and reliability of the current version of the TALPS have been discussed, the perceptions of a cohort of test takers on the face validity of the test can be examined, and how these correspond with their actual test performance.

4. Research methodology

Postgraduate students at the university who both wrote the TALPS during the third quarter of 2011 and who participated in a reception study in the form of a survey were included in the research. Note should be taken of the fact that, although the students differed in terms of their demographics, most of them were from the Faculty of Economic and Management Science (the main user of the test at the time of the study) and that it would be preferable to be able to include more data of students from other faculties in future studies. Of the 246 students who wrote the test during this quarter, 139 completed survey questionnaires, which represents 57% of the test population. The results of an independent-samples t-test, however, show that the variation of scores was not the same for the survey and non-survey groups of students. There was a significant difference between the scores obtained by the cohort of students who participated in the survey ($M = 69.35$, $SD = 14.89$) and the scores of the remaining group of students who wrote the TALPS in the period under review ($M = 61.85$, $SD = 17.06$; $t(211) = 3.61$, $p = .0$, two-tailed; see Annexure B). This means that the students who participated in the survey cannot be considered fully representative of the remaining test population. It would seem that the survey group included students with higher academic literacy levels than those of the non-survey group.

Both qualitative and quantitative data were elicited through the survey questionnaire, although the emphasis was on the latter form of data. The responses obtained in the survey through the use of Likert scales were correlated with scores obtained in the respective test sections.

The survey questionnaire was aimed at assessing to what extent the existing version of the TALPS was well received by the test takers, especially its ratings in terms of accuracy and fairness, and how students rated their own literacy levels. The survey tests the face validity of the test on the basis of the definition of Davies *et al.* (1999: 59):

The degree to which a test appears to measure the knowledge or abilities it claims to measure, as judged by an untrained observer (such as the candidate taking the test or the institution which plans to administer it).

Face validity is important as it can influence the sustainability of the test and can also provide an indication of the amount of time, effort and resources that need to be invested in further test development. In addition to determining the face validity of the TALPS, the survey questionnaire included a section which was aimed at raising awareness of the complex nature of academic literacy and which gave the respondents a chance to reflect on their own literacy abilities and how these could impact on their academic progress.

It was postulated that the survey participants would not be convinced of the need for a test of academic literacy at postgraduate level for a number of reasons. Chan *et al.* (1997: 302) show that when examinees perform poorly in a test, they may tend to attribute their performance to low face validity of the test:

Poor performance on a test for which the content is perceived as unrelated to the content of the job is more self-serving (i.e. less ego-threatening) than when test

content is perceived as related to the content of the job. Hence test performance should positively affect face validity perceptions (Chan *et al.* 1997: 302).

Considering the stressful nature of the test, its degree of difficulty and time constraints, coupled with the possibility that test scores may be used to deny students access to postgraduate study, one would expect the TALPS to have a low face validity amongst those test takers who do not fare well in the test (see Du Plessis 2012). However, an independent-samples t-test that was conducted to compare the face validity perceptions of those who achieved under 50% with those of students who achieved above 50% showed that there was no significant difference in perception for the lower scoring group (M = 3.42, SD = 1.36) and higher scoring group (M = 3.54, SD = 1.38; $t(137) = -.42, P = .68$, two-tailed; see Annexure B). There was thus no significant correlation between lower test scores and lower face validity perceptions.

5. Results of the survey

Most of the respondents were completing their Honours (90%), four (3%) were already studying at Master's level, eight (6%) were completing a postgraduate diploma or certificate course and five (1%) were not studying at the time of the survey. Afrikaans was the home language of most of the survey participants (35%). This was followed by Sotho (16%) and Tswana (11%). Only 9% of the respondents indicated that English was their home language. Figure 2 shows that the cohort of postgraduate students is representative of many diverse language groups.

Respondents were also asked to indicate in which languages they had studied at primary and secondary school level, and what their language of instruction for their undergraduate course work at university had been. The information provided can be summarized briefly in Table 3.

From the above it can be seen how English as the language of instruction increases incrementally by around 10% in each of the education phases. Although four of the Afrikaans students objected to the fact that the test advantaged English first-language speakers, only one of the respondents appealed for an Afrikaans version of the test to be developed, a possible indication of assimilation into the hegemony of English.

On the issue of language development, 62% of the respondents indicated that they had completed development modules to strengthen their language skills while studying at university. Just under a third of the respondents (29%) had taken English as a mainstream subject. The average score obtained by the cohort of test takers for sections 1-7 was 69%, which indicates that the group showed reasonably high academic literacy levels as far as the ability to engage in critical reading and thinking was concerned. The language modules may have played

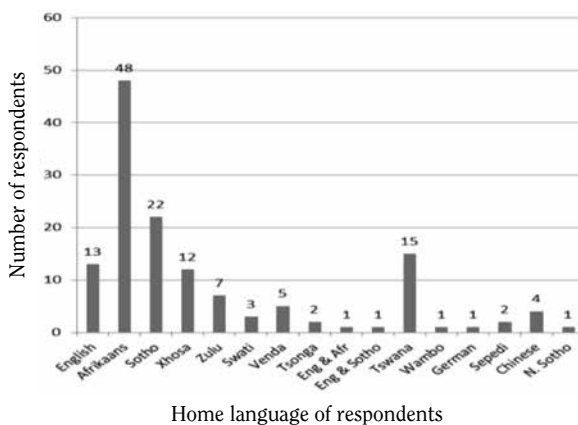


Figure 2: Representation in terms of home language

Table 3: Language of instruction

	Primary school	Secondary school	University
Afrikaans	34%	33%	22%
English	51%	62%	72%
Other	15%	5%	6%

a role in increasing the students' literacy proficiency, but at this point it would not be possible to generalize or arrive at any definite conclusion. However, the picture changes when the essay writing section is added (section 8) and at least 35% of the students are at risk of not achieving success at postgraduate level in terms of the current risk bands identified by the test developers (ICELDA 2011) as indicated in table 4.

Table 4: Risk bands used for the TALPS

Risk associated with level of academic literacy as measured by the Test of Academic Literacy for Postgraduate Students (TALPS)		
Mark	Code	Interpretation
0-33	1	High risk
34-55	2	Clear risk
56-59	3	Risk
60-74	4	Less risk
75 +	5	Little to no risk

The above bands have been based on years of research and the examination of test scores obtained at different levels of study (ICELDA 2011). Further research is needed to investigate whether the scores obtained in tests such as the TALPS have any predictive ability. At the moment the tests serve to indicate current literacy levels and provide an indication of how well positioned a particular test taker is to negotiate advanced academic material through the medium of English.

6. Dimensions of face validity

6.1 Reaction to TALPS prior to taking the test

The attitude of just less than half of the respondents (48.9%) towards being asked to take the test can be described as negative. This may be attributed to the fact that not enough information had been disseminated to the students on the nature of the test prior to its administration and that students feared possible exclusion from postgraduate study on the basis of the test results. In 88% of the cases, the test format had not been discussed with the students at all and only 7% of the respondents indicated that they had had access to an example test. As the number of students required to write the test increases, the status and acceptance of the TALPS may also be expected to increase simultaneously.

A total of 71% of the respondents stated that they would prefer to do a computerized version of the test rather than the paper format. An online version of the TALPS was launched in the course of 2011 at the University of the Free State. Preliminary indications are that the test works well for younger and more computer literate test takers, such as those who participated in the reception study, and provided that test takers have access to stable and sufficient bandwidth.

6.2 Anxiety experienced during the test

Any test situation generates a certain amount of anxiety, especially when the construct is unrelated to the testing of subject knowledge and students cannot prepare for the assessment. Nearly all of the respondents (93%) agreed that anxiety could impact negatively on test performance and a third of the students (32%) reported that they had experienced considerable anxiety while taking the test. Only 4% described their test anxiety as severe, which may be considered non-representative of the cohort of test takers and negligible.

6.3 Difficulty of the test

More than half of the respondents described the test as being difficult. However, only a small percentage (6%) stated that the test was very hard. The majority of respondents selected a scale of 3 or 4 to describe the test difficulty.

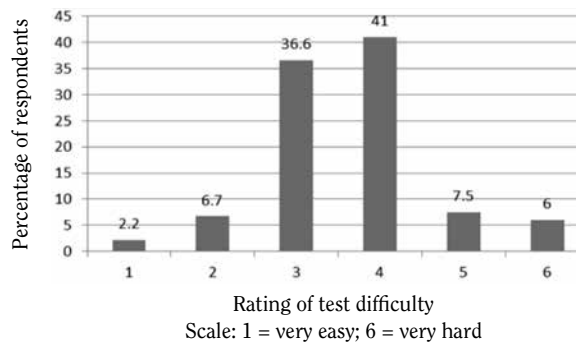


Figure 3: Perception of difficulty of test

The above graph resembles the bell curve of a normal distribution of test scores. The fact that most students found the test moderately to very difficult, suggests that the test is pitched at an appropriate level. This is further supported by the analyses of the test scores and the facility and discrimination values of the test items (see tables 1 and 2).

6.4 Time to complete the test

Most of the respondents (60%) felt that more time should be allowed for completing the test. This is to be expected, considering the pressurized nature of the test. However, as long as the TALPS continues to maintain its high reliability values and discriminates well between stronger and weaker candidates, there is no need to adjust the time allocated to complete the test. There are no indications from the analyses of test scores available up to now that the amount of time allowed for the test is unfair towards the test takers.

6.5 Accuracy of the test

The answers provided by respondents in the open sections of the survey indicate that students do not necessarily understand the nature of academic literacy and thus tend to confuse the test construct with the four language composites of listening, speaking, reading and writing. Not surprisingly about half (49.7%) of the test takers were skeptical of the test's ability to measure their academic literacy levels.

The fact that at least half of the respondents considered the test to be accurate may be viewed positively in the light of the hypothesis that respondents would be negative towards the test, and the findings of Chan *et al.* (1997: 308) that poor test performance generally leads to low face validity. If one takes into account how little information on the test was disseminated to the students before they wrote the test and the negative attitude of nearly half of the respondents prior to taking the test, the responses to the question on the test's accuracy are better than expected.

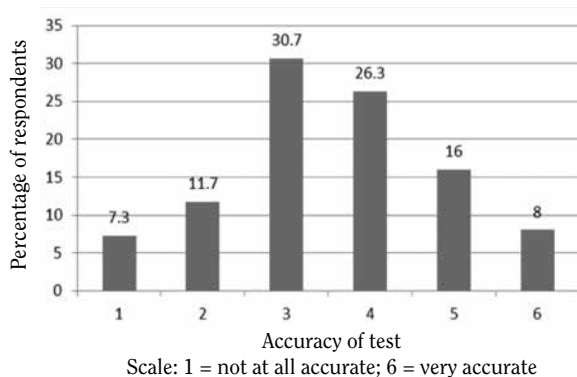


Figure 4: Perception of accuracy of test

6.6 Fairness of the test

The perception of the fairness of a test is one of the most important aspects that feature amongst the regulative conditions for language tests (see Rambiritch 2012). Encouragingly, 70% of the respondents considered the test to be fair. The main reason stated by those who disagreed that the test was fair, was the fact that students could not prepare for the test. This supports the hypothesis that students misunderstand what academic literacy entails and confuse language proficiency with achievement in a content-related subject field. Other responses which may be attributed to a lack of comprehension of the nature of academic literacy include the following statements gleaned from the open questions in the survey:

- The test should only be written by students who did not study English at school.
- Students doing honours already have an adequate knowledge of English.
- The test is not relevant to the field of study.

Six students objected to the fact that English was their second language and that the test was easier for mother-tongue speakers. Of these four were Afrikaans students and two Chinese speaking students. In the case of the Afrikaans students, providing them with an opportunity to do parts of the test or the whole test in Afrikaans would help to establish equity between the English and Afrikaans speaking students. Unfortunately no readily conceivable alternatives exist for the Chinese students, or mother-tongue speakers from other language groups, other than to ensure that they receive sufficient language support during their undergraduate studies. One respondent in the survey stated that students came from different backgrounds and as a result the test could not be considered fair. Again, providing language proficiency support at undergraduate level provides a means of addressing some of the imbalances, but no immediate redress is available in terms of dealing with the socio-economic discrepancies that exist between the different student population groups. Another respondent commented that the test was unfair, because it penalized students for incorrect language usage. This may be considered a valid objection in the instance of students who undertake their postgraduate studies in the medium of Afrikaans and who would be expected to display correct grammatical

usage when writing in Afrikaans. Here too the option of doing parts of the test in Afrikaans could resolve this issue.

6.7 Student perceptions of their own academic literacy levels

The next part of the survey questionnaire asked respondents to rate their abilities to read with understanding, apply critical thinking and produce written academic texts. The results can be summarized as follows:

Table 5: Exploring student perceptions of their own academic literacy levels

	1 Very poor	2	3	4	5	6 Very good
Reading with understanding	0%	2%	10%	44%	31%	13%
Critical thinking	0%	4%	16%	44%	26%	10%
Writing academic texts	1%	7%	35%	38%	18%	1%

Most of the respondents (88%) perceived their academic literacy levels to be above average (scale 4) to very good (scale 6) as far as their ability to read with understanding and critical thinking were concerned. They found critical thinking to be more challenging than reading with understanding, but still rated themselves above average. The students' perceptions of their ability to read with understanding were aligned to a very small degree with their actual test scores, as depicted in the chart that follows. A Pearson correlation coefficient showed that there was a weak, positive correlation between the two variables ($r = .24, n = 137, p < .01$; see Annexure C). Although the correlation is statistically significant, the practical significance is very limited, since it only explains 6% of the variance (see Pallant 2010: 125). It was not possible to make a similar comparison to reflect critical thinking ability, because this is an aspect that is difficult to measure separately and something that should be exercised when engaging in activities such as academic reading and writing. The results of the correlation are depicted graphically below. Each of the six Likert scale numerals (1-6) represents a proficiency band in percentage format (from very low to very high). The middle band (49.6-50.4%) represents a 3.5 on the scale and average ability.

The above situation changes when one examines the respondents' perceptions of their ability to produce academic texts. Here there is a definite shift towards the lower side of the scale with almost half of the respondents (43%) rating their ability to produce written texts as below average. In actual fact the average score obtained for the writing section was 30%, which shows that the students' perceptions of their ability to

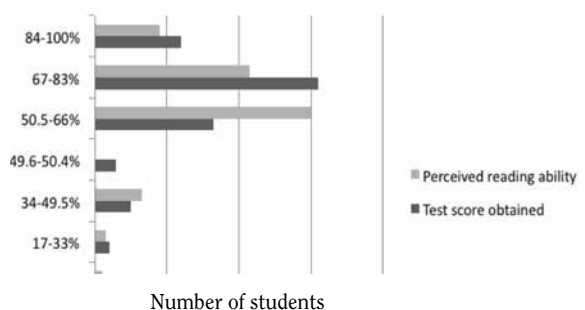


Figure 5: Perceived ability to read with understanding and actual score obtained in the TALPS

write academic texts were considerably inflated. Of the 139 students, only 15 (11%) managed to obtain a score of 50% or higher in the essay writing section. A Pearson correlation coefficient of 0.05 was obtained ($r = .05$, $n = 139$, $p = .6$), which means there is no relationship between the perceived ability and writing score variables (see Annexure C).

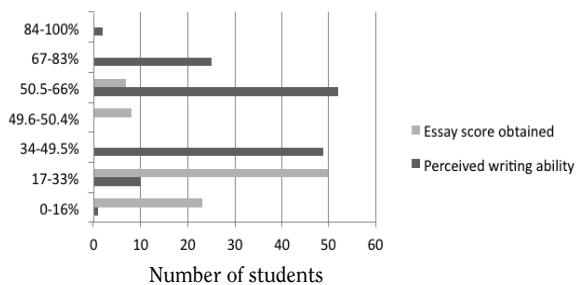


Figure 6: Comparison of perceived writing ability and actual score obtained in the TALPS

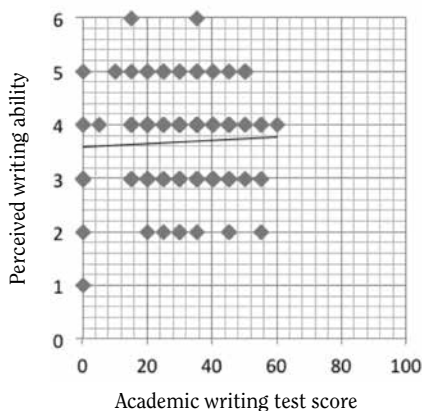


Figure 7: Scatterplot produced for writing proficiency and test performance

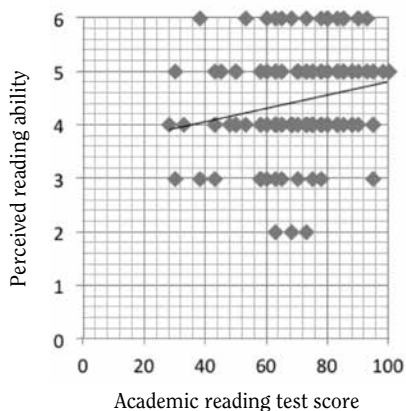


Figure 8: Scatterplot produced for reading proficiency and test performance

7. Conclusion

The analysis of the test results and survey responses show that test performance did not positively affect face validity perceptions. There was no significant difference in the perceptions of the lower and higher scoring groups. A definite discrepancy can be discerned between the perceived ability and test performance of the cohort of postgraduate students who formed part of this study. The inference may be drawn that, even after a few years of undergraduate study, many students still fail to understand the nature of academic literacy and have not been afforded the opportunity to become proficient in English, although two-thirds of the survey respondents had completed language development modules. This is further borne out by the responses of students in the survey which provide some initial reasons for the mismatch between perception and performance. More than half of the respondents (54%) stated that they used their mainstream undergraduate course marks to benchmark their literacy levels, an indication that they tend to confuse subject knowledge with academic literacy and language proficiency. A resounding 83% affirmed that the completion of secondary schooling could be seen as an indication of having attained academic literacy. As many as 50% of the students did not consider language usage within the tertiary environment to differ from that outside academe and about two-thirds (68%) in fact stated that an extensive vocabulary was unrelated

to academic literacy. All of these responses support the postulation that many students have very little understanding of a theoretical articulation of academic literacy.

What is further evident from the research is that students are being admitted to postgraduate study with low academic literacy levels. Their ability to produce authoritative texts is particularly disturbing. Apart from the fact that most of the essays produced in the TALPS were poorly structured (if at all) and riddled with grammatical errors, the more disturbing revelation at this advanced level of study is the students' inability to construct an argument and provide the necessary factual support. This powerlessness in academic writing again suggests a lack of opportunity to interpret texts and engage in critical reflection through the written medium of English at undergraduate level of study. The alignment of undergraduate teaching and learning and assessment practices with the theoretical articulation of academic literacy would appear to be the next sensible step.

REFERENCES

- Bachman, L.F. 2004. *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.
- Bachman, L.F. & Palmer, A.S. 1996. *Language testing in practice*. Oxford: Oxford University Press.
- Bhorat, H. & Oosthuizen, M. 2008. Determinants of Grade 12 pass rates in the post-apartheid South African schooling system. *Journal of African Economies*, 18(4): 634-666.
- Blanton, L.L. 1994. Discourse, artefacts and the Ozarks: Understanding academic literacy. *Journal of second language writing*, 3(1): 1-16.
- Butler, G. 2009. The design of a postgraduate test of academic literacy: Accommodating student and supervisor expectations. *Southern African linguistics and applied language studies*, 27(3): 291-300.
- Chan, D., Schmitt, N., DeShon, R.P., Clause, C.S. & Delbridge, K. 1997. Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of applied psychology*, 82(2): 300-310.
- Cliff, A. & Hanslo, M. 2009. The design and use of 'alternate' assessments of academic literacy as selection mechanisms in higher education. *Southern African linguistics and applied language studies*, 27(3): 265-276.
- Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. 1999. *Studies in language testing: Dictionary of language testing*. Cambridge: Cambridge University Press.
- DIRAP (Directorate for Institutional Research and Academic Planning). 2011. Through-put rates per faculty (HEMIS). Bloemfontein: University of the Free State.
- Du Plessis, C. 2012. The design, refinement and reception of a test of academic literacy for postgraduate students. Unpublished master's dissertation. Bloemfontein: University of the Free State.
- Gee, J.P. 1998. What is literacy? In Zamel, V. & Spack, R. (Eds). *Negotiating academic literacies: Teaching and learning across languages and cultures*. Mahwah, New Jersey: Lawrence Erlbaum Associates Inc. Pp. 51-59.
- Guyer, R. & Thompson, N.A. 2011. *User's manual for Iteman 4.2*. St Paul Minnesota: Assessment Systems Corporation.

- ICELDA (Inter-institutional Centre for Language Development and Assessment). 2011. Language test development and research. [Available online: <http://www.icelda.sun.ac.za/>]. (Accessed on 2011-10-17).
- Pallant, J. 2010. *SPSS survival manual*. 4th edition. Berkshire, England: McGraw-Hill.
- Paltridge, B. & Phakiti, A. (Eds). 2010. *Continuum companion to research methods in applied linguistics*. London: Continuum International Publishing Group.
- Rambiritch, A. 2012. *Accessibility, transparency and accountability as regulative conditions for a post-graduate test of academic literacy*. Unpublished PhD. Pretoria: University of Pretoria.
- Van der Slik, F. & Weideman, A. 2005. The refinement of a test of academic literacy. *Per linguam*, 21(1): 23-35.
- Van der Slik, F. & Weideman, A. 2007. Testing academic literacy over time: Is the academic literacy of first year students deteriorating? *Ensovoort*, 11(2): 126-137.
- Van Dyk, T.J. 2010. *Konstitutiewe voorwaardes vir die ontwerp van 'n toets van akademiese geletterdheid*. Unpublished PhD. Bloemfontein: University of the Free State.
- Van Dyk, T. & Weideman, A. 2004a. Switching constructs: On the selection of an appropriate blueprint for academic literacy assessment. *SAALT Journal for language teaching*, 38(1): 1-13.
- Van Dyk, T. & Weideman, A. 2004b. Finding the right measure: From blueprint to specification to item type. *SAALT Journal for language teaching*, 38(1): 15-24.
- Weideman, A. 2003a. *Academic literacy: Prepare to learn*. Second edition. Pretoria: Van Schaik.
- Weideman, A. 2009. Constitutive and regulative conditions for the assessment of academic literacy. *Southern African linguistics and applied language studies*, 27(3): 235-251.
- Weideman, A. 2011. Academic literacy tests: design, development, piloting and refinement. *SAALT Journal for language teaching*, 45(2): 100-113.
- Yeld, N. et al. 2000. The construct of the academic literacy test (PTEEP). Mimeograph. Cape Town: Alternative Admissions Research Project, University of Cape Town.
-

ABOUT THE AUTHOR

Colleen du Plessis

Department of English
University of the Free State
Email: DuPlessisCL@ufs.ac.za

Colleen du Plessis is a junior lecturer in the Department of English at the University of the Free State and an assistant researcher of the Inter-Institutional Centre for Language Development and Assessment (ICELDA), a partnership of four multilingual South African universities. She recently completed her master's dissertation on *The design, refinement and reception of a test of academic literacy for postgraduate students*. Her fields of interest include language testing, academic literacy development and applied language studies in education.

ANNEXURE A: Through-put rates per faculty (HEMIS) – 1 August 2011 (DIRAP 2011)

	2006			2007			2008			2009			2010		
	HCs	GRs	TPRs	HCs	GRs	TPRs	HCs	GRs	TPRs	HCs	GRs	TPRs	HCs	GRs	TPRs
Post Graduate	7386	2505	33.92%	7174	2241	31.24%	7224	2385	33.01%	7512	2491	33.16%	7365	2579	35.02%
Eco & Man. Sciences	1036	374	36.10%	926	304	32.83%	887	304	34.27%	952	300	31.51%	981	309	31.50%
Education	1659	555	33.45%	1490	487	32.68%	1537	461	29.99%	1582	464	29.33%	1327	448	33.76%
Health Sciences	946	351	37.10%	919	205	22.31%	976	246	25.20%	1032	319	30.91%	1058	358	33.84%
Law	1420	518	36.48%	1395	523	37.49%	1351	524	38.79%	1420	483	34.01%	1287	490	38.07%
Nat. & Agric. Sciences	1357	404	29.77%	1477	418	28.30%	1540	527	34.22%	1614	578	35.81%	1743	598	34.31%
The Humanities	785	248	31.59%	784	248	31.63%	756	277	36.64%	744	311	41.80%	803	317	39.48%
Theology	183	55	30.05%	183	56	30.60%	177	46	25.99%	168	36	21.43%	166	59	35.54%
Under Graduate	15428	2921	18.93%	15970	2458	15.39%	16828	2728	16.21%	17108	2888	16.88%	18882	2965	15.70%
Eco. & Man. Sciences	4638	654	14.10%	4513	596	13.21%	3992	641	16.06%	3997	654	16.36%	4169	667	16.00%
Education	3043	1101	36.18%	3160	666	21.08%	4090	717	17.53%	3903	838	21.47%	4466	707	15.83%
Health Sciences	1413	250	17.69%	1417	260	18.35%	1415	284	20.07%	1375	263	19.13%	1391	261	18.76%
Law	975	100	10.26%	991	106	10.70%	898	122	13.59%	908	101	11.12%	937	124	13.23%
Nat. & Agric. Sciences	2617	410	15.67%	2817	438	15.55%	2985	462	15.48%	3248	483	14.87%	3607	553	15.33%
The Humanities	2649	389	14.68%	2974	381	12.81%	3333	485	14.55%	3574	532	14.89%	4210	631	14.99%
Theology	93	17	18.28%	98	11	11.22%	115	17	14.78%	103	17	16.50%	102	22	21.57%
Grand Total	22814	5426	23.78%	23144	4699	20.30%	24052	5113	21.26%	24620	5379	21.85%	26247	5544	21.12%

HCs = Headcounts
 GRs = Graduates
 TPRs = Through-put rates

ANNEXURE B: Results of independent samples tests using SPSS

Group Statistics: Test scores of survey respondents and non-survey test takers

SURVEY	N	Mean	Std. Deviation	Std. Error Mean
TEST SCORE	1	139	69.35	14.885
TEST SCORE	2	107	61.85	17.060

Independent Samples Test: Test scores of survey respondents and non-survey test takers

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
TESTSCORE	Equal variances assumed	4.600	.033	3.676	244	.000
TEST SCORE	Equal variances not assumed			3.612	210.981	.000

		t-test for Equality of Means			
		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
				Lower	Upper
TEST SCORE	Equal variances assumed	7.502	2.041	3.483	11.521
TEST SCORE	Equal variances not assumed	7.502	2.077	3.408	11.596

Group Statistics: Face validity perceptions of lower and higher scoring groups

Above/Below	N	Mean	Std. Deviation	Std. Error Mean
FACE VALIDITY	1.00	31	3.4194	1.36074
FACE VALIDITY	2.00	108	3.5370	1.38366

Independent Samples Test: Face validity perceptions of lower and higher scoring groups

		Levene's Test for Equality of Variances		t-test for Equality of Means			
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
FACE VALIDITY	Equal variances assumed	.001	.973	-.419	137	.676	-.11768
FACE VALIDITY	Equal variances not assumed			-.423	49.234	.674	-.11768

		t-test for Equality of Means			
		Std. Error Difference	95% Confidence Interval of the Difference		
			Lower	Upper	
FACE VALIDITY	Equal variances assumed	.28092	-.67317	.43781	
FACE VALIDITY	Equal variances not assumed	.27831	-.67690	.44154	

**ANNEXURE C:
Results of Pearson product-moment correlations using SPSS**

Correlation: Reading score and perceived ability

		Readscore	Perception
	Pearson Correlation	1	.235**
Readscore	Sig. (2-tailed)		.006
	N	137	137
	Pearson Correlation	.235**	1
Perception	Sig. (2-tailed)	.006	
	N	137	137

** . Correlation is significant at the 0.01 level (2-tailed).

Correlation: Writing score and perceived ability

		Writescore	Perception
	Pearson Correlation	1	.045
Writescore	Sig. (2-tailed)		.597
	N	139	139
	Pearson Correlation	.045	1
Perception	Sig. (2-tailed)	.597	
	N	139	139

**ANNEXURE D:
The articulation of academic literacy through task types in the TALPS**

TEST SECTION	ASPECT OF LITERACY MEASURED
<p>Section 1: Scrambled text A number of sentences that need to be re-organized into a coherent passage.</p>	<p>Textuality (knowledge of cohesion, grammar) Understanding and responding to the communicative function of the text</p>
<p>Section 2: Interpreting graphs and visual information A short text passage and accompanying graph requiring numerical calculations and visual inferences</p>	<p>Understanding genres Visual literacy Interpreting of information Extrapolation and application of information</p>
<p>Section 3: Academic vocabulary This section includes vocabulary items based on Coxhead's (2000) word list, mainly from the selection of less frequently used words</p>	<p>Advanced vocabulary knowledge Understanding and responding to the communicative function of the text</p>
<p>Section 4: Text types A selection of phrases and sentences representing different genres which have to be matched with a second group of phrases and sentences</p>	<p>Understanding genres Identifying registers Making meaning beyond sentence level</p>
<p>Section 5: Understanding texts A lengthy reading passage and series of questions to be answered</p>	<p>Critical thinking Understanding and responding to the communicative function of the text Deriving meaning beyond sentence level Extrapolating and applying information Distinguishing essential/non-essential information Drawing conclusions and making inferences</p>
<p>Section 6: Grammar and text relations A variation of cloze procedure in which certain words are deleted from a text</p>	<p>Meaning making Understanding and responding to the communicative function of the text Knowledge of cohesion</p>
<p>Section 7: Text editing A passage in which a number of grammatical errors have been made requiring correction</p>	<p>Knowledge of syntax Knowledge of morphology Knowledge of semantics</p>
<p>Section 8: Academic writing A short structured essay assignment based on information provided in the test</p>	<p>Ability to synthesize texts Making meaning beyond the level of the sentence Interpreting information Understanding and responding to the communicative function of the text Extrapolation and application of facts Knowledge of genres and registers Applying coherence Referencing</p>