

# Academic Development Interventions

## — Do We Know What Our Students Need?\*

Nombulelo Phewa (University of South Africa)

**Abstract:** This study aimed to identify Unisa students' academic needs by administering an online diagnostic assessment tool, with an intention to provide guidelines for the development of appropriate and relevant intervention programmes. The tool assessed 13 language-oriented proficiency skills, which were in turn categorised into 7 skills clusters. The latter were grouped into linguistic, discourse and visual and numerical competencies. The results revealed that not only the "underprepared" students have areas of academic development, but also the so-called prepared students do, albeit hidden by their overall achievement scores. We therefore recommend that all students be exposed to diagnostic assessments and that a close integration of learner support and academic services be established. This will ensure that the students' academic development needs are addressed at the levels of recurriculation; materials; and staff development.

Key words: academic development, diagnostic assessments, achievement assessments, language proficiency skills, academic literacies

## **1. Introduction**

This paper attempts to report an ongoing "work-in-progress" study of an academic literacies diagnostic assessment tool (ALDAT) development, administration and redevelopment, undertaken by the Directorate for Counselling, and Career Development (DCCD) of the University of Africa (Unisa). The work began in 2011 when the first version of the tool was administered to a sample of 156 second year second semester Management Accounting (ACN203S) students.

## 1.1 Aim

The purpose of the study was to develop an instrument of measurement that would enable the Academic Development Division of the DCCD to identify students' academic needs, with a view of intervening appropriately by facilitating a development of academic support systems addressing the identified needs.

Research has it that the systems which tertiary institutions put in place to support their undergraduate students can have a significant influence on how well those students achieve (Prebble T., Hargraves H., Leach L., Naidoo K., Suddaby G. & Zepke N., 2005). These researchers go on to define support systems as ...

<sup>\*</sup> This paper was presented at the 2012 HELTASA Conference in Stellenbosch, South Africa.

Nombulelo Phewa, Ed.D. Programme Manager: Academic Literacies Evaluation and information Dissemination (ALEID), University of South Africa: research areas: programme evaluation, and academic development. E-mail: nphewa@unisa.ac.za.

..."[ranging] from those which provide specific support, such as help in choosing courses, orientation programmes at the start of study or peer tutoring and other academic support, through to institutional-level behaviours and organizational features that may support, or hinder, student outcomes" (2005, p. 1).

In cognisance of the guidelines provided in the research report mentioned above, as well as other research reports' recommendations in this regard (Ludwick-Hardman & Dunlap, 2003; Bonanno & Jones; 2007), this study (assessing academic preparedness for university first entrance students) was conducted to identify some gaps in addressing academic development needs that the students may be presenting with. Another parallel and important study was conducted in assessing the predictive validity of the diagnostic assessment tool that was developed in order to provide data to be used in determining whether or not such a tool could be useful.

#### 1.2 Methodology

ALDAT was developed and piloted over two semesters. For the third pilot, the College of Economic Management Sciences (CEMS) decided to roll out the pilot by extending it to four other modules within the college. Three of these were first year level whilst one was a second year level module. These were:

- Introductory Financial Mathematics (DSC 1630);
- Economics 1A (ECS 1501);
- Taxation of Salaried Persons (TAX 1501); and
- Financial Management (FIN 2601) [second year level]

The assessment tool was developed to assess the same skills and competencies that the National Benchmark Tests (NBTs) are assessing. Unisa does not participate in the NBTs because, firstly, it is seen as being pitched at too high a level for the Unisa student profile, and secondly, the NBT results are believed to categorise students into proficient, intermediate and basic streams, thus not clearly highlighting what skills are problematic. The questions popularly posed are the following *inter alia*:

- (1) What difference is this result from the achievement assessment results?
- (2) If students fall into a particular category, does that mean that they all present with the same skills gaps?
- (3) ... and what are these [gaps]?

The Unisa diagnostic assessment differs from the NBTs in that the skills assessed are clustered into what are referred to as skills clusters. Results are provided in such a manner that for each skills cluster, a percentage performance for each learner is highlighted. Reporting in this manner indicates exactly where the skills gaps lie for each learner in a particular cohort. Furthermore, now that we had the last cohort comprising of both first year and second year level students in the same college, the report can provide a comparison of skills shortages within the different levels.

ALDAT is not designed as a screening or selection device, and although, like the NBT scores, it does provide a single "achievement" score for each student, this should not be construed as a pass or fail because that is not what is sought to be achieved by this tool. At this early juncture in the development of this tool, the single score is important as it helps to portray what achievement assessment scores are capable of "disguising" in terms of what strengths and/or weaknesses in the assessed skills and competencies a learner may be presenting with.

It is believed that this assessment tool is a better fit for student placement into appropriate academic learning areas and academic development programmes. It also provided an indication of what curriculum and staff development needs must be considered, which relate to the students' identified needs.

The Unisa academic literacies diagnostic assessment tool is a classical test, whose continuous redevelopment

after each assessment period seeks to ensure that it is validated. The latter process is carried out by a computer aided data analysis programme referred to as Classical Item Analysis (CIA). After the first pilot assessment, which was written by 156 participants of ACN203S students, CIA was run after which a second version of the tool resulted. The second version was tested on 422 ACN203S students. The same CIA process was conducted which resulted into version 3 of the tool. It was at this stage that the assessment was rolled out to more modules. This third time, 485 students sat for the assessment. CIA will ensue and a fourth version will result from this process.

This is an on-line test, whose "achievement score" is immediately received by the student at a click of the "submit" button. The raw scores are exported from the test site to an excel spreadsheet, which allows for a further exportation to the CIA programme for further analysis. The results of this analysis highlight how the test questions perform, that is, how easy or difficult the questions are; how plausible or not the distracters are, and most importantly, how well discriminating the items are between top and bottom performers in the cohort. It is based on these test items performances that the test is revised to produce a new version (test redevelopment).

As has been stated previously, two parallel studies were conducted to determine the correlation of the students' performance in this test and their performance in the examinations. These were carried out at the end of the second semester 2011 and end of the first semester 2012 respectively.

Of the 156 students who participated in the diagnostic assessment in the second semester of 2011, only 138 sat for examinations. Out of the 138, only 38 passed, whereas the results of the diagnostic assessment had indicated that 125 did not seem to present with any readiness difficulties. On testing the diagnostic result against their examination result, it was found that the correlation was only 16.29%, which does not indicate a statistically significant relationship. Further, a study was conducted to determine whether or not a correlation between the examination result and that of the skills' clusters would provide any difference. There was no difference — still there was no correlation.

Of the 422 students who completed version 2 of the diagnostic test, only 340 wrote the examination. A correlation study between their performance on the diagnostic assessment *vis a vis* their examination result and the following ensued. Of the 340 students who wrote the examination, only 197 students passed, whereas 283 had, according to their results of the diagnostic test, seemingly indicated being "academically ready". On testing the diagnostic result's correlation with their examination result, it was found that the correlation is only 20.51%, which, again, does not indicate a statistically significant relationship. A further study of determining correlation between the skills cluster assessed against the examination result was conducted, just as was done in the second semester 2011, and still there was no major difference recorded in terms of correlation.

Below, I present the findings relating to both the correlation studies and the students' results of their performance in the diagnostic assessments<sup>1</sup> in versions 1, 2 and 3.

## 1.3 Findings

Correlation studies conducted in both second semester 2011 and first semester 2012 provided statistically insignificant relationships between the diagnostic assessment results and the final examination results in the module ACN203S (16.29% and 20.51% respectively). However, it is noted that when comparing the correlation score of the second semester 2011 with that of the first semester 2012, there is a slight increase of 4.22%. Small as this difference might be, it provides hope that with the continuation of such assessments and their improvement or

<sup>&</sup>lt;sup>1</sup> The correlation analysis study only relates to the ACN203S candidates as only they participated in the assessment during the periods 2nd semester 2011 and 1st semester 2012, that is versions 1 and 2 of the test.

validation exercises carried out through the CIA analysis process, in time the assessment may record a higher correlation result. In the following sections, I present graphically the performance of all students who participated in the diagnostic assessment of the second semester 2012. This was a pool of students from the five Economic Management Sciences modules at first and second year levels (see paragraph 3). As stated earlier, this cohort consisted of 485 students.

This cohort, being a mixture of both first and second year level students in different modules, were broadly found to be in need of further development especially in reading and writing skills as portrayed by the bar chart labelled as Figure 1(a) below. The chart highlights percentages of students who scored less than 50% in the tool.

This is a deviance from the results of the earlier pilots, which suggested that those cohorts required more help with visual and numerical literacies rather than these. Secondly, the percentages of students indicating a need for help, was much lower than those portrayed by this chart (Figure 1(b)). We then assumed that since these assessments were carried out on second year level students, their skills could have been enhanced by the whole year of interventions in their first year level modules. Added to this was the fact that they had passed the first year modules, thus seemingly indicating having experienced less struggles in the entry level skills at which this tool was pitched.



Figure 1 (a) % of (Mixed) CEMS Students Scoring Below 50%, (b) ACN203S Students Scoring Below 50% (2011)

With this new set of seemingly deviant data and the assumptions made in the first two pilots, we decided to separate the first year level students' data from that of the second year level students, in a quest to find out what it would yield. Starting with the first year level cohort of DSC 1630; ECS 1501 and TAX1501, we found that about 86% of the first year students seemed to be presenting with a difficulty in sentence construction and at that writing for appropriate contexts also seemed to be problematic. They did not indicate experiencing many struggles with the rest of the linguistic related skills clusters as well as visual and numerical literacies. It is interesting to note that these students have indicated a better reading for meaning ability than the whole cohort combined. This assertion is presented by the bar chart in Figure 2.

In the current second year level cohort, we found that the assumptions made in the first two pilots of second year level students does not seem to be holding water, as these students seem to be in need of more support in writing rather than in any of the other skills clusters assessed. But then, when one considers the fact that the modules that these students have registered in are more numerical and visual oriented than writing and reading oriented, one may assume that these linguistic skills may not have been particularly been a focus of their intervention programmes. We therefore assume that this may be the reason why the students do not seem to indicate an enhancement of these during this assessment. The following bar chart presented in Figure 3 shows percentages of students in this cohort that scored below 50% for each of the skills clustered assessed by this tool. When comparing the first year level students' bar chart to that of the second year level students, it is important to remember that the two cohorts comprised of different numbers, thus the percentages may be deceiving.



Figure 2 % of First Year Level Students Scoring below 50% Per Skills Cluster Assessed



Figure 3 % of Second Year Level Students Performing below 50% Per Skills Cluster Assessed

A finding that seems to be in resonance with those of the first two pilots is that the results of this tool have indicated student's individual strengths and/or weaknesses in a manner that enables us to consider these when developing interventions. It has been found that although students may have been identified as "being at risk/unprepared" for HE learning, s/he could be having some academic strengths that are "hidden" by the total achievement score. The inverse applies to those students who may have been identified as being prepared may be presenting with some weaknesses in certain skills areas. The following table bears testimony to this assertion. The student numbers have been changed arbitrarily to protect the identifies of the students.

#### 1.4 Discussion

In this concluding section of the paper, I provide a discussion relating to predictive validity as it relates to the current study. I also provide lessons learnt so far in conducting this study, and then I make recommendations for the next phases of the study.

Student numbers	Lexical items	Sentence construction	Understanding purpose	Reading for meaning	Appropriacy to context	Visual literacy	Numerical literacy	Total score
0001	60	56	69	50	67	63	73	74
0002	60	67	66	63	33	75	73	63
0003	68	33	63	74	100	81	82	74
0004	70	78	66	75	100	75	91	75
0005	75	56	72	72	67	75	82	72
0006	28	11	25	20	0	25	18	29
0007	70	44	72	51	100	88	82	75
0008	75	33	72	49	67	81	82	72
0009	75	67	75	76	100	88	100	76
0010	65	33	75	49	100	88	100	72
0011	68	44	38	27	33	56	55	40

Table 1 Students' Scores in Skills Clusters Assessed Including Total Scores

### 2. Predictive Validity

The Academic Literacies Diagnostic Assessment tool used in these studies is a very newly developed tool and is thus still under redevelopment for validation. Specifically, the results which are the focus of this report emanate from the third version of the tool, whereas the predictive validity results come out of the first and second versions' results.

Test redevelopment is a process that reoccurs after every period of administration, by exporting the assessment scores from the development administrative site to an excel spreadsheet. This is done so as to accommodate easy exportation to the CIA programme site, where analysis of the tool's performance automatically takes place. The results of the CIA inform us *inter alia*, whether or not the test items work, that is:

- Their level of difficulty;
- Their ability to discriminate between the top and bottom performers;
- The plausibility and therefore attractiveness of the distracters; and
- The acceptability or not of statistical measurements of the test constructs. Based on the results of CIA regarding the test items' performance, all those items that seem to be leaving a

lot to be desired in terms of bullet points above, are reviewed, rephrased, changed or omitted completely from the next version of the tool. All of this is done in order to improve and therefore validate the tool.

It is not surprising therefore to find that the correlation results provided by colleagues from CEMS portray the tool as having a "non-predictive validity" because predictive validity is defined as that which "implies the prediction of future academic performance of students against their scores on a testing instrument" (Visser & Hanslo, 2005). Having said that, I need to add that, in my opinion and from experience in diagnostic test development, an instrument's predictive validity is developmental. It is something that gets stronger as the instruments gets continually improved — in other words predictive validity is a quality issue.

Visser and Hanslo (2005) also state that "predictive studies in the testing arena most often take the form of correlations, Z scores and regression analysis and that the use of these statistics in Higher Education seem straightforward but this is seldom the case" (1). I note that our colleagues only make mention of correlation studies to the exclusion of other stats methods, and the question is whether or not triangulation of methods in this regard could have yielded a wider comparison and therefore a more "open" conclusion. This may or may not be the issue as research has shown that statistical methods of the likes of correlation studies, Z scores and regression analysis are fraught with challenges (Shuttleworth, 2009; Cliff, Hanslo et al., 2002; Yeld & Visser, 2000; Griesel, 1999; Linn, 1989). In order to mediate the challenges as reported in research, Visser and Hanslo (2005), in the concluding section of their research report, advocate the use of ...

"...[A] survival analysis approach to investigate predictive validity [as it] overcomes many of the challenges associated with the traditional statistical approaches such as correlations and regression analyses" (9).

Elsewhere in the research world, a study was conducted that seems to mediate the challenges of predictive validity by proposing a different approach to test development. This study draws a link between a test designed to measure the outcomes of a programme of instruction and a diagnostic instrument ...

"...[W]here a number of items allotted to each item type usually reflects the amount of time devoted to mastery of material tapped by this item type during the course instruction" (Kennet-Cohen et al., 2003, p. 3).

Contrary to the study by Kennet-Cohen and her colleagues, the Unisa Academic Literacies Diagnostic Assessment Tool (ALDAT) is a generic instrument and has not been designed to measure the outcomes of any specific programme of instruction. Due to the generic status of the ALDAT, it was not clear during its development how many items should have been written for each type in order to adequately represent a particular body of content (Thorndike, 1982, cited in Kennet-Cohen et al., 2003, p. 3). Thus the test developers made a decision regarding the number of items allotted to each type as a problem of maximization of criterion-related validity under certain constraints. It is a known factor that time is one of the leading constraints that have an effect on the test's reliability and validity. This speaks directly to the length of the duration of the test, as well as time required by the test takers to engage with each item. With the benefit of hindsight, I can deduce that it can therefore be advantageous to be able to produce a "time-constraint-built-in" test to mediate the effect of time on reliability and validity.

#### 3. Strengths of the Study

The findings of this study have highlighted, as previously mentioned in the relevant section, that not only academically underprepared students may present with difficulties in some skills areas. It has been shown that also

those who are deemed as being prepared for mainstream study may at times indicate that they too experience difficulties in some of the assessed skills clusters. This information is helpful in designing study material or extra exercises that may help identified individuals where they may need help.

Also, there seems to be an improvement in the tool's predictive validity, although it may be premature to say this with conviction. The results of 2012 second semester correlation studies will have to be scrutinized for this claim.

Although this study has the strengths alluded to in the paragraph above, it also has some limitations, which when exposed may lead to their mitigation in the next piloting phases. I present these limitations in the section below.

#### 4. Limitations of the Study

The limitations of this study can be traced from the first pilot, where the sample chosen was a second level cohort whereas the assessment tool had been developed for first level first entry cohort. This was shown in the results because, as could be expected, the students scored very highly in almost all the skills clusters assessed. The "logical" explanation put forth for this performance was that at this time in their study programme (second semester second level), these students had already been through a first level programme as well as a first semester second level programme. Therefore, whatever academic skills they might have lacked on entry, might have been enhanced during interaction with the Management Accounting programme.

Secondly, the sample of the second pilot was a mixed cohort of second level and first level students. This gave rise to a situation where a comparison could be made of the performance of the second level students against the first level ones. This comparison did not highlight any major differences except that the first level cohort demonstrated that they could better differentiate between audiences and genre than the second level cohort.

A third limitation, which may be ascribed to institutional organization, is that these pilots take place after registration, and because registration is a month long process, by the time the results are out, there usually is less time to implement any of the guidelines emanating from the exercise. In that sense, the assessment cannot timorously provide the information that needs to be disseminated for the provision of intervention

#### 5. Lessons Learnt

Although a number of useful lessons have been learnt during this process, I only mention that which I regard as very important in ensuring that the assessment exercise will contribute positively to student success. That is, ensuring stakeholder buy-in and collaboration (with commitment) and the allocation of all necessary resources in the redevelopment of the ALDAT with improved predictive validity.

Diagnostic assessment tool development is a process that must be taken seriously by all stakeholders in the teaching and learning environment of an institution, as it aims at identifying the teaching and learning that is focused on students' academic needs. Thus, to produce a strong, reliable and valid instrument, total co-operation and collaboration with serious commitment by all parties concerned is required. Kennet-Cohen et al. (2003) have proposed an approach to test development that is "time-efficient" for the purpose of improving the predictive validity of a test. It would therefore be prudent of our future test redevelopment processes to take heed of this proposed approach as well as ensuring a dedicated diverse and well represented stakeholder test development team.

#### **6.** Recommendations

Based on the allusions made above, it is therefore recommended that a strong test development team be established with full commitment to the process from all stakeholders. These would be sourced from the following departments:

- Directorate for Counselling, Career and Academic Development (DCCAD AL and AD practitioners);
- Directorate for Curriculum and Learning Development (DCLD MCQ specialists);
- Linguistics;
- English;
- At least one lecturer from each of the participant modules; and

• Directorate for Institutional Statistics and Analysis (DISA)/Bureau for Market Research (BMR)/Statistics (to ensure that statistical calculations and interpretations and inferences thereof are in the right hands).

It is also recommended that for the exercise to bear the necessary fruits of timorously providing academic programme developers with information relating to the students' academic needs, the assessment should be conducted as a pre-registration exercise. Should this recommendation be approved, the students will be timorously advised as to their strengths and areas needing further development, which could impact on their choice of a manageable module load for registration.

#### References

- Bonanno H. and Jones J. (2007). "The MASUS procedure: Measuring the academic skills of university students A diagnostic assessment (rev. 2007)", A Resource Package prepared and packaged by the Learning Centre, University of Sydney, Australia.
- Cliff A., Hanslo M., Herman M., Fish C. and Visser A. (2002). "Science Faculty Report of August 2002", University of Cape Town Science Faculty Report (August 2002), unpublished.
- Griesel H. (1999). "Access and higher education sector: A South African case study on innovative policy and programmes", Draft Discussion Document for the Eastern and Southern African Regional and National Seminars, 15-19 & 22 June, 1989, Prepared for the Department of Education and the Association for the Development of Education in Africa (ADEA).
- Kennet-Cohen T., Bronner S. and Cohen Y. (2003). "Improving the predictive validity of a test: A time-efficient perspective", in: *The Annual Meeting of the National Council on Measurement in Education*, Chicago.
- Linn R. L. (1989). Educational Measurement (3rd ed.), NY: ACE and Macmillan.
- Ludwick-Hardman S. and Dunlap J. C. (2003). "Learner support services for on-line students: Scaffolding for success", *The International Review of Research in Open and Distance Learning*, Vol. 4, No. 1.
- Martyn Shuttleworth (2009). "Predictive validity", available online at: http://www.experiment-resources.com/predictive-validity.html.

Prebble T., Hargraves H., Leach L., Naidoo K., Suddaby G. and Zepke N. (2005). "Supporting students in tertiary study: A summary of a synthesis research on the impact of student support services on student outcomes in undergraduate tertiary study", Visser A. & Hanslo M. (Eds.), *The Predictive Validity of an University Admissions Test: A Survival Analysis Approach*, http://www.iaea.info/documents/paper\_1162a20681.pdf.

Visser A. J. and Hanslo M. (2005). "Approaches to predictive studies: Possibilities and challenges", South African Journal of Higher Education, Vol. 19, No. 6, pp. 1160–1176.

Yeld N. and Visser A. (2000). "Impact assessment in academic development", in: *South African Academic Development Conference*, Rhodes University.