

Tydskrif vir Taalonderrig - Journal for Language Teaching  
- Ijenali yokuFundisa iLimi - Ijenali yokuFundisa iiLwimi -  
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta Polelo  
- Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo - Ijenali  
Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -  
Journal for Language Teaching - Ijenali yokuFundisa iLimi  
- Ijenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi  
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -  
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi  
- Tydskrif vir Taalonderrig  
- Journal for Language  
Teaching - Ijenali  
yokuFundisa iLimi -  
Ijenali yokuFundisa  
iiLwimi - Ibhuku  
Lokufundisa Ulimi  
- Tšenale ya tša  
Go ruta Polelo -  
Buka ya Thuto  
ya Puo - Jenale  
ya Thuto ya Dipuo  
Ijenali Yekufundzisa  
Lulwimi - Jena?a ya u  
Gudisa Nyambo - Jenala  
yo  
Dyondzisa Ririmi - Tydskrif vir Taalonderrig  
- Journal for Language Teaching - Ijenali  
yokuFundisa iLimi - Ijenali yokuFundisa iiLwimi -  
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya  
Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a ya  
u Gudisa Nyambo - Jenala yo Dyondzisa Ririmi  
- Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali  
yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi -  
Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -  
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -  
Journal for Language Teaching - Ijenali yokuFundisa iLimi  
- Ijenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi  
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta



**Johann L van der Walt**

and

**Faans (HS) Steyn**

North-West University (Potchefstroom Campus)

# A norm for an academic literacy placement test

## **Abstract**

A test series usually consists of various versions of a test. Test-takers and consumers of tests expect all these versions to be at the same level of difficulty, but this is very difficult to achieve at the test design stage. In this article we suggest that the establishment of a norm for test results may be a practical solution to this problem. We illustrate how a norm can be established by investigating historical data from an Afrikaans test of Academic Literacy

(TAG) and then proposing a formula to standardise results whenever a new version is written. This can easily be done by using a program such as Excel. This will enable test administrators to treat test-takers equally and fairly whenever decisions about them have to be taken.

**Keywords:** academic literacy test; test level; norm; standardisation

## 1. Introduction

A test that is developed, administered and scored in a consistent manner is often regarded as a 'standardised' one. Van der Walt and Steyn (2014: 110), however, argue that all versions of a standardised test should ideally be at the same level of difficulty. Kunnan and Grabowski (2013: 312) put this contention as follows: "... in large-scale testing, test-takers ideally would receive the same score on an assessment irrespective of which version (Test Form A or B) they took or on which occasion (in January or in June) they took the assessment. If they do, there is evidence that the assessment forms or occasions can be considered interchangeable since these variables introduce little error into the measurement".

Academic literacy tests are a case in point. A number of such tests have been developed in South Africa and are applied for selection, placement and diagnostic purposes (Scholtz & Allen-Ile, 2007: 921; Wilson-Strydom, 2012: 137). Examples of such tests include the National Benchmark Test (NBT), Standardised Assessment Test for Access and Placement (SATAP) and the Test of Academic Literacy Levels (TALL). The administration and use of such tests have been a contentious issue at times, but, as Cliff (2015: 4) points out, the tests are intended "to better understand the under-preparedness of secondary school students" and to establish whether they are 'selectable', and if not, to what extent they are 'under-prepared' (Cliff, 2015: 3). Students are then placed in appropriate support programmes or Academic Literacy modules. In order to do so, decisions regarding appropriate norms to be applied must be made. Any educational assessment decision involves a norm of one kind or another. The NBT, for example, divides test takers into three competence levels or norms (Cliff, 2015:14). The problem that arises is that decisions regarding these norms are often taken arbitrarily (cf. Scholtz & Allen-Ile, 2007: 924) or based on expert opinion (cf. Cliff, 2015: 14).

Our focus in this article is on the Afrikaans version of TALL, the *Toets van Akademiese Geletterdheidsvlakke* (henceforth TAG). It is used at tertiary institutions to place students in an appropriate Academic Literacy course. The decision to be taken in our example is a binary one – a student takes either course A or course B. As there are a number of versions of the test, each based on the same blueprint, Van der Walt and Steyn (2014) considered the question of whether the test can be regarded as a standardised one or not, i.e. whether its versions are at the same level of difficulty. They found that the levels of two tests they administered to the same study population were not equivalent (cf. Van der Walt & Steyn, 2014: 125), and concluded that it seemed to be difficult to achieve equivalence at the design stage of a test. They suggested that a reliable model for the adjustment of scores may prove to be useful if a norm for the versions of the test can be established.

The aim of this article is to illustrate how such a norm can be established in a series of language tests, and then to propose such a norm for TAG by means of illustration. The question, however, is if this would amount to the 'standardisation' of the test, as the generalization inference (Kane, 2006: 24) remains difficult to achieve in practice. We

propose that the establishment of a norm is a practical way of addressing the problem of test equation and standardisation.

## 2. The validity of the test

A reliable norm can only be established if there is evidence that the test is valid for the use to which it is put. Validity is essentially a property of test score interpretation, not a property of a test. It can only be established through a process of validation. This process involves consideration of the purpose of the test, its content and method, intended (and possible unintended) consequences, potential decisions that can be made and the impact it may have on test-takers. These considerations involve both descriptive- and decision-based interpretations that are made after relevant evidence has been collected.

In the case of TAG, its main claim is that it can identify first-year students who are at risk in their university studies as a result of the level of their academic literacy skills. A claim such as this must be supported by evidence. Two comprehensive validation studies of the test have been done (cf. Van der Walt & Steyn, 2007; Van Dyk, 2010), and, in addition, there is a large body of research on various aspects of the test (cf. <http://icelda.sun.ac.za>). Some of the most relevant evidence is briefly summarised here (cf. Van Dyk 2010 for a more detailed analysis of TAG and a discussion of the constitutive conditions of language test design).

The test is based on a specific blueprint (Van Dyk & Weideman, 2004; Patterson & Weideman, 2013) and samples typical academic tasks at university (Van Dyk, 2010: 202; Van der Walt & Steyn, 2014: 112). Its reliability has consistently been above 0.80, the criterion Weir (2005: 29) sets for language tests (cf. Van der Slik & Weideman, 2005: 33; Weideman & Van der Slik, 2008: 166). Potential misclassifications are within acceptable limits (cf. Weideman & Van der Slik, 2008: 170; Van der Walt & Steyn, 2014: 114). Both Van der Walt and Steyn (2007: 148) and Van Dyk (2010: 218) found similar correlation coefficients between the various sections of the tests and between each subtest and the whole test. Although they do not all satisfy the criteria, they are probably as good as can be expected. Factor analyses of test results to determine whether it measured the postulated factors have also been done. For example, Van der Walt and Steyn (2007: 150) performed a principal component analysis to verify the construct validity of each of the six sections of the test. Only two sections formed one construct. More than one underlying factor was extracted, which did not explain the high percentage of the total variance. Academic literacy is a rich and varied construct, however, and a degree of heterogeneity should probably be tolerated (cf. Van der Slik & Weideman, 2005: 32).

As ability is not transparently visible from raw scores, a Rasch analysis of the results of an administration of TAG was done by Van der Walt and Steyn (2007: 145-147). They found that there was no significant mismatch between the general ability of students

and the general difficulty level of test items and that all items were in accordance with the fitted Rasch model, showing no misfit. They concluded that the test was a fair one and that there was no need to adjust the difficulty level of the test.

As far as test fairness is concerned, borderline students are given a second opportunity to sit the test (cf. Weideman & Van der Slik, 2008: 178). In an attempt to avoid stigmatization, results are published in five risk categories (Van der Slik & Weideman, 2005: 33), but, by its very purpose, the test has inevitable consequences for some students – they will end up in academic literacy classes they would rather not attend. In some instances the test was not very transparent, as few test-takers typically looked at specimen tests beforehand. Some test-takers reported that they had been tired, as the test is usually taken during first years’ orientation week at university (cf. Van der Walt & Steyn, 2008: 201-202).

TAG can therefore generally be regarded as a good, fair and valid test.

### 3. Establishing a norm for the test

By ‘norm for the test’ we do not mean the norm for the content of a test as reflected by its test items, but rather a norm for the total score. Historical results of total scores for the population (universe) for which the test is intended are required. As illustration, we analysed available results for the period of 2004 – 2012 at three campuses where the test was administered, using the statistical computer package STATISTICA (StatSoft Inc., 2014).

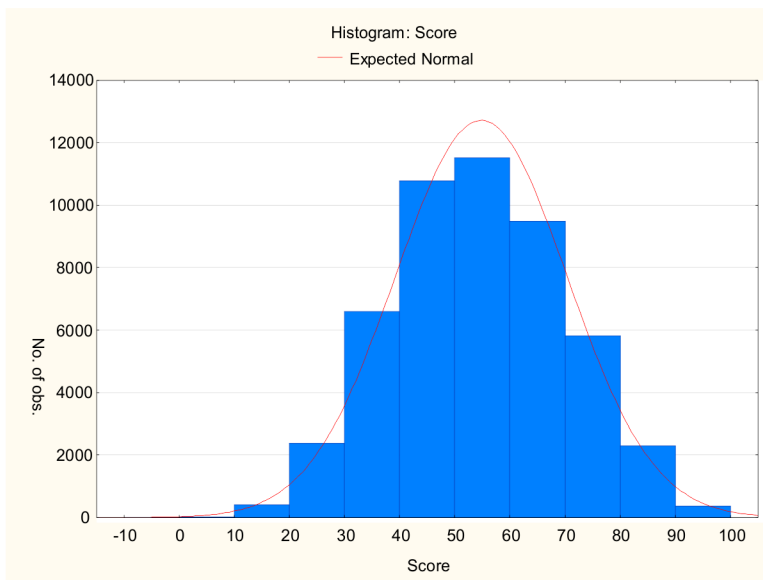
The descriptive statistics of the (total) scores for the complete available population of first-year students who wrote TAG are displayed in Table 1. The population is a large one (N=49,682), with a mean of 54.81 and a standard deviation of 15.59.

**Table 1: Descriptive statistics for whole population**

Variable	N	Mean	Median	Minimum	Maximum	1 <sup>st</sup> Quartile
Score	49682	54.81	54.00	1.00	100.00	44.00

Variable	3 <sup>rd</sup> Quartile	10 <sup>th</sup> Percentile	90 <sup>th</sup> Percentile	Standard Deviation	Skewness	Kurtosis
Score	66.00	35.00	76.00	15.59	0.06	-0.45

The distribution of the students' scores is displayed in the histogram in Figure 1, and shows a fairly normal distribution.



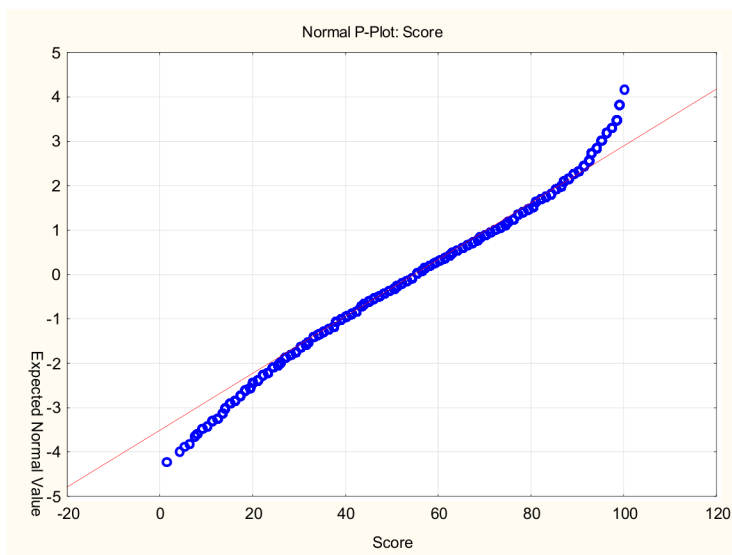
**Figure 1: Distribution of students' scores**

Table 2 displays the percentiles in steps of 5 for the population's scores. From the table one can, for example, see that 5% of the students had scores of less than 30, 50% had scores of less than 54, 10% had scores of above 76, etc.

**Table 2: Percentiles**

Percentile	5	10	15	20	25	30	35	40	45
Score	30	35	38	41	44	46	48	50	52
	50	55	60	65	70	75	80	85	90
	54	56	59	61	63	66	69	72	76

The normal quantile-quantile (Q-Q) plot (i.e. a **graphical method** for comparing the population's **probability distribution** with that of a normal distribution by plotting their quantiles against each other) in Figure 2 should form a straight line when the scores are normally distributed. The plot, however, shows a deviation from normality of the scores at extreme values below 25 and over 85.



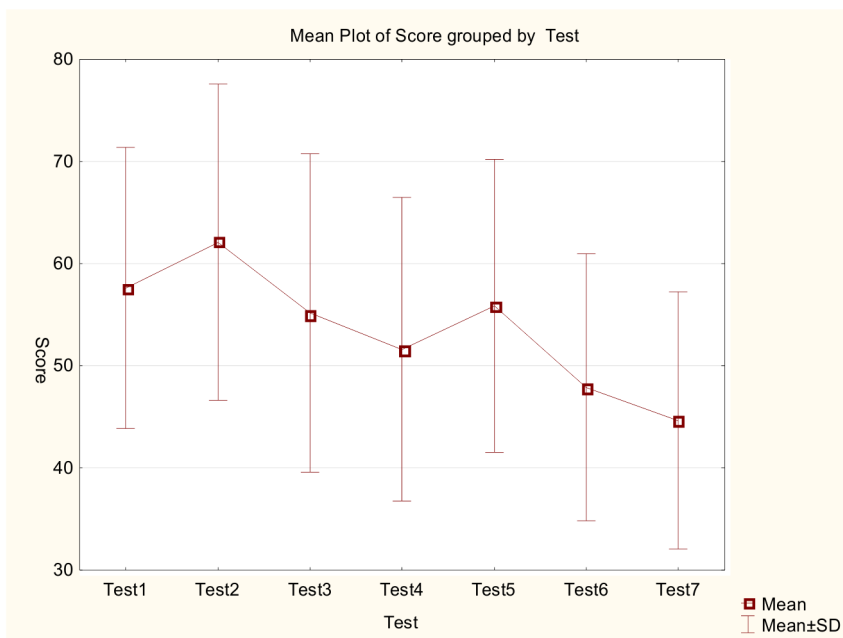
**Figure 2: Normal Q-Q plot showing slight deviation from normality in scores**

Tables 3, 4, and 5 and Figures 3, 4 and 5 below display the descriptive statistics of scores for breakdowns into the different *tests* (seven), *years* (nine) and *campuses* (three) for comparative purposes.

**Table 3: Descriptive statistics of scores per tests**

Test *	Mean	N	Standard Deviation (SD)	Minimum	Maximum	1 <sup>st</sup> Quartile	Median	3 <sup>rd</sup> Quartile	10 <sup>th</sup> Percentile	90 <sup>th</sup> Percentile
Test 1	51.6	5162	14.9	4	97	41	51	62	32	71
Test 2	62.1	11370	15.5	6	99	51	63	73	42	82
Test 3	55.2	10838	15.6	5	100	44	55	67	35	76
Test 4	57.6	5540	13.8	11	98	48	58	67	39	76
Test 5	55.9	5972	14.4	12	99	46	56	66	37	75
Test 6	47.9	5474	13.1	8	87	38	47	57	31	66
Test 7	44.7	5326	12.6	1	95	36	44	53	29	61

\*Test 2 was administered in 2005 and 2010, Test 3 in 2006 and 2007, while the remaining tests were each administered in one year.



**Figure 3: Mean scores per test**

**Table 4: Descriptive statistics of scores per year**

Year	Mean	N	Standard Deviation	Minimum	Maximum	1 <sup>st</sup> Quartile	Median	3 <sup>rd</sup> Quartile	10 <sup>th</sup> Percentile	90 <sup>th</sup> Percentile
2004	57.6	5540	13.7	11	98	48	58	67	39	76
2005	66.2	5222	14.5	13	99	56	67	77	47	85
2006	56.7	5376	15.3	7	99	46	57	68	36	77
2007	53.6	5462	15.7	5	100	42	53	65	33	75
2008	51.6	5162	14.8	4	97	41	51	62	32	71
2009	55.9	5972	14.4	12	99	46	56	66	37	75
2010	58.6	6148	15.5	6	99	49	59	70	39	78
2011	47.9	5474	13.1	8	87	38	47	57	31	66
2012	44.6	5326	12.6	1	95	36	44	53	29	61

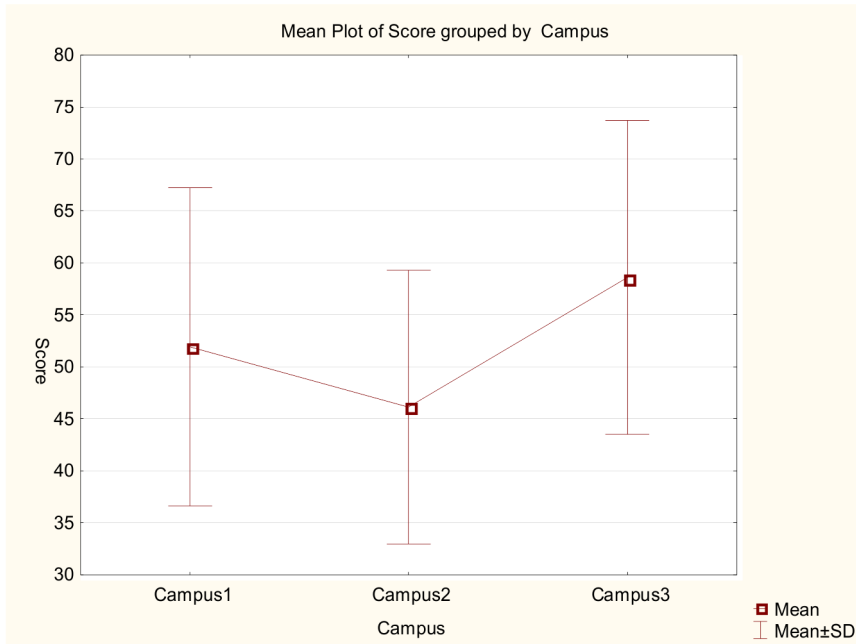




**Figure 4: Mean scores per year**

**Table 5: Descriptive statistics of scores per campus**

Campus	Mean	N	Standard Deviation	Minimum	Maximum	1 <sup>st</sup> Quartile	Median	3 <sup>rd</sup> Quartile	10 <sup>th</sup> Percentile	90 <sup>th</sup> Percentile
Campus1	51.9	25582	15.3	1	98	41	51	63	32	73
Campus2	46.1	1414	13.1	12	99	37	45	54	30	64
Campus3	58.6	22686	15.1	6	100	48	59	69	39	79



**Figure 5: Mean scores per campus**

In order to investigate whether the tendencies over the years were the same for the campuses, the interaction effect in a two-way analysis of variance (ANOVA) for the scores, with the two factors years and campuses, was established. This is reflected in the calculated partial eta-squared of 0.0028. The partial eta-squared is defined as the proportion of the total ANOVA sum of squares accounted for by the interaction effect and can in this case be regarded as very small (cf. Cohen, 1988). Also, by means of one-way ANOVAs, the effects of test, year and campus on the scores can be measured by their respective partial eta-squared values. The effects and their interpretations are shown in Table 6. Note that statistical significance testing of effects is not appropriate when dealing with a complete population.

**Table 6: Effect sizes**

Effect	Partial eta-squared	Interpretation of effect (Cohen, 1988)
Test	0.126	large
Year	0.141	large
Campus	0.053	medium

Table 6 shows that the effect of both the test and the year is practically significant (cf. Steyn, 2009) (i.e. large effect and therefore important), while that of the campus is substantial (i.e. medium effect).

#### 4. A shorter period

We may choose to base the norm on a shorter, more recent period (for example, if we argue that students' abilities are declining or improving). The descriptive statistics for TAG results for only three years (2010 – 2012) are displayed in Table 7. There is a slight decrease in the mean of the scores but little difference in the standard deviation.

**Table 7: Descriptive data for three years**

Variable	N	Mean	Median	Minimum	Maximum	1 <sup>st</sup> Quartile
Score	16948	50.77	50.00	1.00	99.00	40.00

Variable	3 <sup>rd</sup> Quartile	10 <sup>th</sup> Percentile	Percentile (30)	Standard Deviation	Skewness	Kurtosis
Score	61.00	32.00	71.00	15.12	0.19	-0.36

#### 5. Standardisation of the norm

If we take as the standard the results for the complete population of students from the three campuses over the period 2004 – 2012, its mean is  $\bar{x}_s = 54.81$  with a standard deviation (SD) of  $S_s = 15.59$ .

To adjust the test scores ( $x$ ) of any group in future (with mean  $\bar{x}$  and SD of  $S$ ) so that they have the same mean ( $\bar{x}_s$ ) and SD ( $S_s$ ) as the standard, such scores can be adjusted to  $y$  (e.g. using Excel) by means of the following formula (cf. De Wet et al., 1981:184):

$$(1) \quad y = (S_s/S) (x - \bar{x}) + \bar{x}_s$$

The rationale for the formula is as follows: First, the scores  $x$  are standardised to Z-scores by subtracting their mean and dividing by their SD, i.e.  $Z = (x - \bar{x}) / S$ . These scores now have zero mean and unit SD. By multiplying  $Z$  with  $S_s$ , the new scores have the same SD as the norm and adding  $\bar{x}_s$  shifts the mean of the scores to the standard mean  $\bar{x}_s$ .

This adjustment formula has the following properties:

- If the mean test score and the SD of the scores are the same as those of the standard test, no adjustment is necessary, i.e.  $\bar{x} = \bar{x}_s$  and  $S = S_s$ , then  $y = x$ .
- If the mean test score differs from that of the standard (i.e. the test scores are on average a constant number smaller or larger than that of the standard distribution of scores), but the SD's of the test scores and the standard are the same (i.e. the test scores have the same variability of that of the standard distribution), then the test scores are adjusted by adding the difference between the standard and test means, i.e.  $y = x + (\bar{x}_s - \bar{x})$ .
- If the SD of the test scores differs from that of the standard, but the means of the test scores and the standard are the same, then the test scores are adjusted by multiplying the test score with the ratio of the standard and test score SD's, i.e.  $y = (S_s/S) x$ .
- If the SD and means of the test scores both differ from those of the standard, then the test scores are adjusted by multiplying the difference between the test score and its mean with the ratio of the standard and test score SDs, after which the standard mean has been added – resulting in formula (1).

So, if we use the mean and SD for 2004 – 2012 as the current standard, this would result in the following:

$$y = (15.59/S) (x - \bar{x}) + 54.81.$$

If, for example, a future group has a mean  $\bar{x}$  of 60 and standard deviation S of 14, the adjustment formula becomes:

$$y = (15.59/14.0) (x - 60.0) + 54.81 = 1.114 (x - 60) + 54.81.$$

Scores of  $x = 30, 40, 50, 60, 70$  and  $80$  will then be adjusted to  $y = 21.39, 32.53, 43.67, 54.81, 65.95$  and  $77.09$ . Rounded to the nearest integer, the adjusted scores become  $y = 21, 33, 44, 55, 66$  and  $77$ . These scores will then be in line with the standardised norm.

Thus, after each administration of TAG, the mean and SD of the individual scores must be calculated. Then, using the formula, standardisation of the results can be achieved for

each campus or university separately by substituting the mean and SD of the particular sub-population as  $\bar{x}_s$  and  $S_s$  in formula (1) above. As pointed out, a shorter period (say, three recent years) can also be used as the norm (with  $\bar{x}_s = 50.77$  and  $S_s = 15.12$  in the case of the 2010 – 2012 group).

## 6. Conclusion

Any placement of students in a course (i.e. an application of a norm) has implications for teaching and learning. Students should be placed in an appropriate support programme or module that addresses their Academic Literacy needs. Cliff (2015: 14-18) illustrates how this can be done by pointing out how performance in an Academic Literacy test can be related to a teaching-learning programme. The test results can be analysed for diagnostic purposes and students' needs specified in great detail. This should be standard practice in any application of a norm.

As far as testing is concerned, it remains difficult to ensure that various versions of a test are at an equal level of difficulty when they are designed. They may be based on the same blueprint and formula, but this does not ensure equality in practice.

Standardisation of the norm removes the concern that various versions of a tests are not at the same level of difficulty. It ensures that results are in line with previous ones, and this guarantees the fair and equal treatment of students, especially if they are to pass, fail or be placed in a course. In this way, consequential validity, to which Messick (1989) refers as an essential part of validity, can be ensured.

In the case of TAG, where a cut-off point must typically be established for every administration of the test, the norm ensures that placement in academic literacy courses is consistent from year to year, thus ensuring that the same standard is maintained. This will obviate the need for placement measures that may be subjective or convenient, such as teaching capacity, which ultimately does not serve the interest of students. Setting a specific cut-off score, however, is a different matter, and requires further investigation.

## References

- Cliff, A. 2015. The National Benchmark Test in Academic Literacy: How might it be used to support teaching in higher education? *Language Matters* 46(1): 3-21.
- Cohen, J. 1988. *Statistical power analysis*. 2<sup>nd</sup> edition. New York: Academic Press.
- De Wet, J.J., Monteith, J.L. de K., Steyn, H.S. & Venter, P.A. 1981. *Navorsingsmetodes in die Opvoedkunde: 'n Inleiding tot empiriese navorsing*. Durban: Butterworth.
- Kane, M.T. 2006. Validation. In: Brennan, R.L. (Ed.) *Educational measurement*, 4<sup>th</sup> ed. Santa Barbara: Praeger. pp. 319-342.
- Kunnan, A.J. & Grabowski, K. 2013. Large-scale second language assessment. In: Celce-Murcia, M. Brinton, D.M. & Snow, M.A. (Eds.) *Teaching English as a second or foreign language*. New York, NY: Heinle/Cengage. pp. 304-319.
- Messick, S. 1989. Validity. In: Linn, R.L. (Ed.) *Educational measurement*. New York: Macmillan. pp. 13-103.
- Patterson, R. & Weideman, A. 2013. The refinement of a construct for tests of academic literacy. *Journal for Language Teaching* 47(1): 125-151.
- Scholtz, D. & Allen-Ile, C.O.K. 2007. Is the SATAP test an indicator of academic preparedness for first year university students? *South African Journal for Higher Education* 21(7): 919-939.
- StatSoft, Inc. 2014. Statistica (data analysis software system), version 12. [www.statsoft.com](http://www.statsoft.com).
- Steyn, H.S. (jr.) 2009. Manual: Effect size indices and practical significance. North-West University (Potchefstroom Campus), Potchefstroom. <http://www.nwu.ac.za/content/statcs-manual>. Date of access: 9 May 2016.
- Van der Slik, F. & Weideman, A. 2005. The refinement of a test of academic literacy. *Per Linguam* 21(1):23-35.
- Van der Walt, J.L. & Steyn, H.S. 2007. Pragmatic validation of a test of academic literacy at tertiary level. *Ensovoort* 11(2):137-153.
- Van der Walt, J.L. & Steyn, H.S. 2008. The validation of language tests. *Stellenbosch Papers in Linguistics* 38: 191-204.
- Van der Walt, J.L. & Steyn, H.S. 2014. Towards standardisation: A comparison of two versions of an academic literacy test. *Journal for Language Teaching* 48(1):109-129.

- Van Dyk, T.J. 2010. *Konstitutiewe voorwaardes vir die ontwerp van 'n toets van akademiese geletterdheid*. Bloemfontein: University of the Free State. (Thesis – PhD.)
- Van Dyk, T.J. & Weideman, A. 2004. Switching constructs: On the selection of an appropriate blueprint for academic literacy assessment. *Journal for Language Teaching* 38(1): 1-13.
- Wilson-Strydom, M. 2012. Using the NBTs to inform institutional understandings of 'under-preparedness': Implications for admissions criteria. *South African Journal of Higher Education* 26(1): 136-151.
- Weideman, A, Patterson, R. & Pot, A. 2014. Construct refinement in tests of academic literacy. <http://www0.sun.ac.za/icelda/index.php/research>. Date of access: 11 January 2015.
- Weideman, A & Van der Slik, F. 2008. The stability of test design: Measuring differences in performance across several administrations of an academic literacy test. *Acta Academia* 40(1):161-182.
- Weir, C. 2005. *Language testing and validation: An evidence-based approach*. Houndmills, Basingstoke: Palgrave Macmillan.

---

## **ABOUT THE AUTHORS**

Johann L van der Walt

School of Languages, Potchefstroom Campus, North-West University, Potchefstroom

Email: [Johann.VanDerWalt@nwu.ac.za](mailto:Johann.VanDerWalt@nwu.ac.za)

The author is emeritus professor of English at North-West University and currently manages short courses and research in languages at the Potchefstroom campus. His research interests include language testing, second language acquisition and language teaching.

HS (Faans) Steyn

Statistical Consultation Services, Potchefstroom Campus,  
North-West University, Potchefstroom

Email: [Faans.Steyn@nwu.ac.za](mailto:Faans.Steyn@nwu.ac.za)

The author is a part-time statistical consultant at North-West University after his retirement. He assists researchers at the university in the planning of their research and doing sound statistical analyses. His research interests include practical statistics to develop new methods of analysing data.



Tydskrif vir Taalonderrig - Journal for Language Teaching  
- Ijenali yokuFundisa iLimi - Ijenali yokuFundisa iiLwimi -  
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta Polelo  
- Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo - Ijenali  
Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -  
Journal for Language Teaching - Ijenali yokuFundisa iLimi  
- Ijenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi  
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go  
ruta Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya  
Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a ya

u Gudisa Nyambo - Jenala yo  
Dyondzisa Ririmi - Tydskrif  
vir Taalonderrig - Journal  
for Language Teaching  
- Ijenali yokuFundisa  
iLimi - Ijenali  
yokuFundisa  
iiLwimi - Ibhuku  
Lokufundisa Ulimi  
- Tšenale ya tša  
Go ruta Polelo  
- Buka ya Thuto  
ya Puo - Jenale  
ya Thuto ya Dipuo  
- Ijenali Yekufundzisa

Lulwimi - Jena?a ya u Gudisa Nyambo - Jenala yo  
Dyondzisa Ririmi - Tydskrif vir Taalonderrig  
- Journal for Language Teaching - Ijenali  
yokuFundisa iLimi - Ijenali yokuFundisa iiLwimi -  
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto  
ya Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a  
ya u Gudisa Nyambo - Jenala yo Dyondzisa Ririmi  
- Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali  
yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi -  
Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo  
- Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -  
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -  
Journal for Language Teaching - Ijenali yokuFundisa iLimi  
- Ijenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi  
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta



Tydskrif vir Taalonderrig - Journal for Language Teaching  
- Ijenali yokuFundisa iLimi - Ijenali yokuFundisa iiLwimi -  
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta Polelo  
- Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo - Ijenali  
Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -  
Journal for Language Teaching - Ijenali yokuFundisa iLimi  
- Ijenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi  
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go  
ruta Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya  
Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a ya

u Gudisa Nyambo - Jenala yo  
Dyondzisa Ririmi - Tydskrif  
vir Taalonderrig - Journal  
for Language Teaching  
- Ijenali yokuFundisa  
iLimi - Ijenali  
yokuFundisa  
iiLwimi - Ibhuku  
Lokufundisa Ulimi  
- Tšenale ya tša  
Go ruta Polelo  
- Buka ya Thuto  
ya Puo - Jenale  
ya Thuto ya Dipuo  
- Ijenali Yekufundzisa

Lulwimi - Jena?a ya u Gudisa Nyambo - Jenala yo  
Dyondzisa Ririmi - Tydskrif vir Taalonderrig  
- Journal for Language Teaching - Ijenali  
yokuFundisa iLimi - Ijenali yokuFundisa iiLwimi -  
Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto  
ya Dipuo - Ijenali Yekufundzisa Lulwimi - Jena?a  
ya u Gudisa Nyambo - Jenala yo Dyondzisa Ririmi  
- Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali  
yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi -  
Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo  
- Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta  
Polelo - Buka ya Thuto ya Puo - Jenale ya Thuto ya Dipuo -  
Ijenali Yekufundzisa Lulwimi - Jena?a ya u Gudisa Nyambo  
- Jenala yo Dyondzisa Ririmi - Tydskrif vir Taalonderrig -  
Journal for Language Teaching - Ijenali yokuFundisa iLimi  
- Ijenali yokuFundisa iiLwimi - Ibhuku Lokufundisa Ulimi  
- Tšenale ya tša Go ruta Polelo - Buka ya Thuto ya Puo -  
Jenale ya Thuto ya Dipuo - Ijenali Yekufundzisa Lulwimi  
- Jena?a ya u Gudisa Nyambo - Jenala yo Dyondzisa  
Ririmi - Tydskrif vir Taalonderrig - Journal for Language  
Teaching - Ijenali yokuFundisa iLimi - Ijenali yokuFundisa  
iiLwimi - Ibhuku Lokufundisa Ulimi - Tšenale ya tša Go ruta

