Per Linguam 2024 40(2):67-86
http://dx.doi.org/10.5785/40-2-1177

Journal Home Page:
http://perlinguam.journals.ac.za

Per Linguam

A Journal for Language Learning
Tydskrif vir Taalaanleer

EQUITABLE ACCESS AND LITERACY SUPPORT: ADDRESSING ACADEMIC LITERACY NEEDS THROUGH INSTITUTIONAL TESTING ALTERNATIVES

Laura M. Drennan, ORCID: https://orcid.org/0000-0002-9416-2590
Division of Languages, Literacies and Literatures, University of the Witwatersrand, Johannesburg, South Africa

ABSTRACT

In response to historical inequalities in South Africa, higher education institutions are obligated to address issues around equitable access to students from previously marginalised backgrounds. Many of these students are underprepared for the transition from secondary to tertiary education, presenting a significant barrier to their academic success. The administration of standardised tests of academic literacy was disrupted by the COVID-19 pandemic; the results of such standardised tests are typically used to place students in appropriate literacy support programmes. This situation highlighted the pressing need for alternative means to identify at-risk students needing literacy support. This study evaluates the appropriateness of an existing in-house test as an alternative measure of academic literacy levels. The study involved 2,292 first-year students at a South African university who had been identified as at risk by scoring below 64% on the National Benchmark Test (NBT) or through a machine-learning algorithm. The pilot results demonstrate the validity and reliability of the test as a viable alternative measure of academic literacy. Such in-house tests offer flexible and equitable testing solutions that can be tailored to meet diverse student needs.

KEYWORDS: academic literacy; equitable access; language testing; student placement

INTRODUCTION

Internationalisation and the resultant massification of higher education precipitated an influx of students at universities. In South Africa, this massification has been accelerated by institutional obligations to address issues of equitable access to the higher education sector for students from previously disadvantaged backgrounds. However, access to a university does not necessarily equate to academic success. The exceptional growth in student enrolment has not corresponded with an increase in student success. The research indicates that a high percentage of students discontinue their studies before completing their degrees, and few students complete their qualifications in regulation time (Carpenter & Roos, 2020; Cloete, 2016; DHET, 2016; Fourie, 2020; Young & Fynn, 2024). The majority of these students are underprepared for tertiary studies due to inadequate primary and secondary education (Sebolai & Huff, 2015). There is a clear articulation discrepancy characterised by a mismatch between the skills

students possess when they leave school and what is expected of them at the tertiary level (Sebolai & Stanford, 2020), which presents a significant barrier to their academic success.

Although a multitude of factors influence students' performance at the tertiary level, a key factor concerns their unpreparedness to deal with the demands of academic discourse (Drennan et al., 2021; Weideman et al., 2016). This scenario is further compounded by students having to study in a language other than their first language (Pot & Weideman, 2015; Van Rooy & Coetzee-Van Rooy, 2015). The underlying assumption in discussions concerning language in education is that issues concerning English language proficiency serve as a barrier to educational achievement, as it has repercussions on students' ability to read with comprehension and demonstrate their learning in terms of written tests, examinations and academic assignments (Huysamen, 2000; Ncgobo, 2009; Nel & Nel, 2009). As a result, these students face challenges in negotiating typical higher education tasks with the desired measure of success, both theoretically and practically.

Hence, students' English language proficiency levels have implications for their academic literacy proficiency and ability to contend with academic discourse, as evidenced by students' performance on tests of academic literacy such as the National Benchmark Test (NBT) and the Test of Academic Literacy Levels (TALL). The majority of all first-time students entering higher education in South Africa require language development to succeed in their university careers, as demonstrated by their National Benchmark Test (Academic Literacy component) results (cf. Myburgh, 2015; Myburgh-Smit & Weideman, 2017; Pot, 2013; Pot & Weideman, 2015; Sebolai, 2016; Van Rooy & Coetzee-Van Rooy, 2015). These tests are designed to measure students' academic literacy levels, the results of which are used for admission to specific degree programmes or to identify those in need of additional literacy support (Drennan et al., 2021). Support is usually in the form of credit-bearing academic literacy modules that address students' language and literacy skills. Not only are these tests important for determining incoming students' degree of academic preparedness, but they also have the potential to impact the content incorporated in such literacy interventions and the extent to which they are tailored to meet students' literacy needs (Drennan et al., 2021; Sebolai & Stanford, 2020).

However, the COVID-19 pandemic and lockdown procedures caused substantial disruptions in the administration of established tests such as the NBT and the TALL. The online versions of conventional tests were inaccessible for many students without access to specialised equipment such as webcams. This situation had significant implications for identifying students requiring additional literacy development and support, and institutions had to find alternative means to identify at-risk students for placement in appropriate literacy programmes. In response to this disruption, a South African university utilised an existing in-house test as a potential alternative for assessing the academic literacy proficiency of first-time students entering the institution. Given the diagnostic and placement value of these tests and their potential to inform literacy course content, it is crucial to interrogate their quality and validity. Therefore, this paper discusses the results of the pilot administration of the in-house test to

determine its appropriateness for future institutional administration and the corresponding implications for adaptable and equitable literacy testing in the higher education context.

CONVERSATIONAL AND ACADEMIC DIMENSIONS OF LANGUAGE PROFICIENCY

In response to the injustices of the past whereby language was used as a tool to separate and divide people based on constructions of inferiority and superiority, the new South African policy of multilingualism enshrines individuals' language rights and recognises language as a resource for learning and development in the pursuit of equity and nation building. The language in the education policy advocates that children learn in their home language until Grade 3, after which they switch to English as the primary medium of instruction (Howie, Venter & Van Staden, 2008; Thesen & Van Pletzen, 2006). In effect, more than 80% of children are required to learn in a language that is not their mother tongue, despite the current policy's advocacy for learning to take place in the home language until Grade 12 (Howie et al., 2008). This sudden switch in the medium of instruction, coupled with poor teaching in under-resourced schools, significantly impacts the development of the cognitive academic language proficiency (CALP) required to deal with advanced literacy in English. Consequently, learners develop basic interpersonal communicative skills (BICS) in English, but their abilities to comprehend and express concepts and ideas in either their home language or English are undermined (Cummins, 2000). BICS is acquired early in the developmental process and refers to conversational fluency involving a student's ability to hold a conversation. This kind of fluency reflects a "fraction of the language skills required for academic success" (Cummins, 2009: 23), whereas CALP relates primarily to the academic aspects of language proficiency that are generally developed in the formal education process (Du Plessis, 2021: 8).

Another aspect concerns students' discrete knowledge of the rule-governed aspects of grammar that allow for application in other cases governed by particular rules. Such knowledge can be developed through direct instruction or immersion in environments where attention is drawn to literate language forms during meaning-making. When the language of instruction is other than the home language, there is limited direct transference in other oral language proficiency (linguistic concepts, vocabulary, sentence memory, and word memory) (Drennan, 2011: 32). On the other hand, CALP involves the interpretation and production of increasingly complex written language. This kind of proficiency reflects "the extent to which an individual has access to and command of the oral and written academic registers of schooling" (Cummins, 2000: 76). Additional language students typically require a minimum of five years to catch up to the grade expectations in the language of instruction.

Research has shown that policymakers tend to conflate the conversational and academic dimensions of language proficiency, which poses significant academic challenges to students whose home language is not English (Cummins, 2020). When educators equate students' English communicative fluency with the proficiency required for integration into English-only education programmes, the students participating in those investigations generally receive little

to no support, neither in understanding instruction nor in developing their English academic literacy skills.

In South Africa, the discrepancy in ability levels can be attributed, in part, to the language subjects students take at school. Students take either first language (L1), home language (HL), second language (L2) or first additional language (FAL), which are meant to prepare students adequately for tertiary education. The only tangible evidence of the standards and levels attained through language instruction is the result of the Grade 12 school-leaving examination. Students achieving high scores for the school-level language examination would assume that these marks are a true reflection of their language abilities. However, the results of standardised tests, such as the TALL and NBT, call into question the quality and accuracy of the Grade 12 school-leaving language examination as a measure of language learning and level of ability (Du Plessis, 2021). Students' weak language abilities and the inadequacies of the school system in preparing students for tertiary studies are further evidenced by the high proportion of students who withdraw from or fail to graduate in higher education (Gumede, 2017; McKay, 2016).

The upcoming section defines academic literacy and the types of abilities required to handle academic discourse in tertiary education to contend with conceptualising the kind of language ability under scrutiny here in a theoretically adequate way.

DEFINING ACADEMIC DIMENSIONS OF LANGUAGE PROFICIENCY

Academic discourse is considered a unique form of discourse (Cummins & Swain, 1986; Gee, 1989; Lea & Street, 1998) that is a "distinctly different type of language that is used within a particular social institution" (Patterson & Weideman, 2013: 126). Academic literacy concerns what counts as knowledge construction in particular fields of study and the corresponding use and control of "language and cognitive abilities for specific purposes and in specific contexts" (Van Dyk & Van de Poel, 2013: 56). To better conceptualise what academic literacy entails, one might consider the behaviours and actions of those who are proficient readers and writers of academic discourse (Drennan, 2019). The successful use of academic language and proficient engagement with texts involves the interpretation, synthesis, and extrapolation of information towards the production of (multimodal) texts for specific audiences. Proficiency furthermore includes the ability to reflect critically and develop a unique academic voice to speak and write with authority (Blanton, 1994).

Language ability is typically associated with the discrete language skills of reading, writing, listening, and speaking. However, language ability concerns the contextualised use of language to perform specific language-use tasks (Bachman & Palmer, 1996). In reality, human interaction involves a mixture of these skills for communicative purposes in specific contexts. In the tertiary context, these language abilities combine and interact to create academically coherent and appropriate language products (Weideman, 2007). Students use a combination of reading, critical thinking, distinction-making, categorisation, and inference-making abilities, as well as speaking and listening skills to produce written texts (Weideman, 2013). Thus,

instead of regarding reading, writing, listening and speaking abilities as discrete constructs, it is more practical to view them as "the means by which that ability is realised in the performance of tasks in actual language use situations" (Douglas, 2000: 38).

The ability to use language for such purposes can be illustrated by Patterson and Weideman's (2013: 139–140) functional definition (construct) of academic literacy, which involves the capacity to:

- Understand and use a range of academic vocabulary as well as content or discipline-specific vocabulary in context.
- Interpret the use of metaphor and idiom in academic language, and perceive connotation, wordplay and ambiguity.
- Understand and use specialised or complex grammatical structures correctly, including texts with high lexical diversity, containing formal prestigious expressions, and abstract/technical concepts.
- Understand relations between different parts of a text, be aware of the logical development and organisation of an academic text via introduction to conclusions, and know how to understand and eventually use language that serves to make the different parts of a text coherent.
- Understand the communicative function of various ways of expression in academic language, such as defining, providing examples, inferring, extrapolating, and arguing.
- Interpret different kinds of text type (genre), and have sensitivity for the meaning they convey, as well as the audience they are aimed at.
- Interpret, use and produce information presented in graphic or visual format to think creatively: devise imaginative and original solutions, methods or ideas through brainstorming, mind-mapping, visualisation, and association.
- distinguish between essential and non-essential information, fact and opinion, propositions and arguments, cause and effect, and classify, categorise and handle data that make comparisons;
- See sequence and order, and do simple numerical estimations and computations that express analytical information, that allow comparisons to be made, and can be applied for the purposes of an argument.
- Systematically analyse the use of theoretical paradigms, methods and arguments critically, both in respect of one's own research and that of others.
- Interact with texts both in spoken discussion and by noting down relevant information during reading: discuss, question, agree/disagree, evaluate and investigate problems, analyse.
- Make meaning of an academic text beyond the level of the sentence; link texts, synthesise and integrate information from a multiplicity of sources with one's knowledge to build new assertions, draw logical conclusions from texts, with a view finally to producing new texts, with an understanding of academic integrity and the risks of plagiarism.

- Know what counts as evidence for an argument, extrapolate from information by making inferences, and apply the information or its implications to other cases than the one at hand.
- Interpret and adapt one's reading/writing for an analytical or argumentative purpose and given one's experience and insight to produce new academic texts that are authoritative yet appropriate for their intended audience.

When considering what students do when generating academic discourse, we may observe three processes involving gathering information, processing that information, followed by producing information (Drennan et al., 2021; Drennan, 2021). This is supported by proponents of discourse synthesis who argue that proficient readers and writers in the (academic) disciplines use such disciplinary knowledge to select (gather) information from various sources relevant to a topic, organise (process) that information according to the objectives of the text they are required to produce, and connect (produce) information according to the links between ideas and constructs (Chan, 2018; Van Dijk, 1979; Spivey, 2001; Spivey & King, 1989; Spivey, 2001; Spivey & King, 1989). Proficient readers (writers) use a macro-processing strategy to gather and process information according to "how discourse is conventionally organised and how to use text structure to guide their understanding" (Spivey, 1989: 9). They draw on a mental representation of a text, their knowledge of textual cues and the relationships between ideas to make inferences across multiple texts, thereby demonstrating the link between discourse production and comprehension (Spivey & King, 1989). This understanding informed the design of the test that is the subject of this paper. The test (part of a larger PhD study) (see Drennan, 2019) was initially designed to measure students' preparedness to produce (multimodal) information as they transitioned from undergraduate to postgraduate studies. Two institutional pilots (discussed later on in the paper) were administered to determine whether the test was appropriate for administration to first-time incoming students as a measure of their academic literacy (also see Drennan, 2021; Drennan et al., 2021). The upcoming section expands on the design and specifications of the test.

TEST DESIGN AND SPECIFICATIONS

In linking to the idea mentioned in the previous section of the processes in which academic language is embedded, the design of the Assessment of Preparedness to Produce Multimodal Information (APPMI) is based on the conception that producing information in writing or any other format is preceded by processes of gathering (selecting) and processing (organising) information. When a proficient reader/writer is given the task of generating information with a distinct goal, they become involved in higher-level activities of sourcing pertinent data from primary sources and systematising it to construct relationships between thoughts that reinforce and support the set objective (Spivey 2001; Spivey & King, 1989). This requires altering their reading to align with the task at hand and their prior knowledge concerning the text structure principles of diverse source types. This understanding of how discourse is regularly organised in different kinds of text permits them to use criteria of significance to select information, comprehend how concepts are interconnected in a text through textual indicators and

extrapolate between texts (Frederiksen, 1975; Spivey & King, 1989; Van Dijk, 1979). Consequently, producing discourse and synthesising details are regarded as closely linked to the comprehension of discourse.

There are various cognitive phases and sub-processes linked to the processes of gathering, processing and producing information, including conceptualisation, meaning construction, organising ideas, and monitoring and revising (Chan, 2013; 2018: 11). Table 1 below illustrates the relationship between these cognitive phases, the three processes associated with the generation of academic discourse, and the components of the academic literacy construct provided earlier (Patterson & Weideman, 2013).

Table 1: Alignment of cognitive phases, sub-processes and literacy construct (Drennan, 2021; Drennan et al., 2021)

	Cognitive phases	**Sub-processes**	**Alignment with components of construct**
GATHERING AND PROCESSING	Conceptualisation	Task representation Macro-planning	- Communicative function - Text type (including visual representations) - Essential/non-essential information, sequence and numerical distinctions, identifying relevant information for evidence - Employment and awareness of method - Inference, extrapolation, synthesis of information, and construction of argument
	Meaning construction	Global careful reading Selecting relevant ideas Connecting ideas from multiple sources	- Vocabulary and metaphor - Complex grammar and text relations - Communicative function - Text type (including visual representations) - Essential/non-essential information, sequence and numerical distinctions, identifying relevant information for evidence - Employment and awareness of method - Inference, extrapolation, synthesis of information, and construction of argument
	Organising ideas (based on mental task representation)	Organising intertextual relationships between ideas Organising ideas in a textual structure	- Vocabulary and metaphor - Complex grammar and text relations - Text type (including visual representations) - Communicative function - Employment and awareness of method - Inference, extrapolation, synthesis of information, and construction of argument

PRODUCING	Monitoring and revising	Monitoring and revising during text production Monitoring and revising after text production	- Use of vocabulary and metaphor - Use of complex grammar, and text relations - Communicative function - Text type, including visually presented information - Essential/non-essential information, sequence and numerical distinctions, identifying relevant information and evidence - Employment and awareness of method - Inference, extrapolation, synthesis of information, and constructing an argument - Creative thinking - Writing with authority

The augmented construct of academic literacy illustrated above informed the design of the various subtests of the APPMI. Table 2 shows the various subtests and components of the construct they were designed to measure, followed by the subtest weightings in Table 3.

Table 2: Alignment of APPMI subtests and construct (Drennan, 2019, 2021)

APPMI subtests	**Alignment with construct**
• Understanding text type and communicative function • Making academic arguments • Interpreting graphic and visual information • Text comprehension	- Communicative function - Text type (including visual representations) - Essential/non-essential information, sequence and numerical distinctions, identifying relevant information for evidence - Employment and awareness of method - Inference, extrapolation, synthesis of information, and construction of argument
• Organising information visually • Understanding academic vocabulary • Text comprehension • Making academic arguments • Organisation of text/scrambled text	- Vocabulary and metaphor - Complex grammar and text relations - Communicative function - Text type (including visual representations) - Essential/non-essential information, sequence and numerical distinctions, identifying relevant information for evidence - Employment and awareness of method - Inference, extrapolation, synthesis of information, and construction of argument
• Interpreting graphic and visual information • Organisation of text/scrambled text	- Vocabulary and metaphor - Complex grammar and text relations - Text type (including visual representations) - Communicative function - Employment and awareness of method

• Understanding text type and communicative function • Making academic arguments • Grammar and text relations • Text editing	- Inference, extrapolation, synthesis of information, and construction of argument

Table 3: APPMI subtests and weightings (Drennan, 2019, 2021)

Subtest	**Number of items**	**Weighting**
Organising information visually	8	8
Organisation of text	5	5
Understanding academic vocabulary [two-word format]	6	12
Interpreting graphic and visual information	8	8
Understanding text type and communicative function	5	5
Text comprehension	18	18
Making academic arguments	8	16
Grammar and text relations	16	16
Text editing	6	12
Totals	**80**	**100**

STUDY POPULATION

This paper focuses primarily on the results of the 2nd institutional pilot of the APPMI in 2023. The test was piloted five times between 2018 and 2023; the last two were institutional pilots administered to the targeted first-year student cohort. Due to lockdown protocols during COVID-19, students were unable to write the NBT, the results of which are typically used in this institution to place students in appropriate faculty-specific academic literacy courses. Consequently, the 2021 pilot cohort constituted students who had been identified as being at risk using a machine-learning algorithm developed by the Centre for Teaching and Learning (CTL) and channelled into faculty-specific academic literacy courses accordingly (see Drennan et al., 2021 for further details). The 2023 cohort consisted of students who had scored below 64% on the NBT and were channelled into the appropriate literacy courses. The aforementioned algorithm was used for students who had not written the NBT to identify at-risk students needing additional academic literacy support. Every effort was made to ensure

equal representation across the different faculty-specific literacy courses for both the 2021 and 2023 pilot administrations.

The 1st institutional pilot of the test was administered online in compliance with lockdown protocols during the COVID-19 pandemic. Although the test performed well in the 2021 pilot, a second institutional pilot in 2023 was necessary to determine whether the test would perform similarly when administered in the traditional format. For an accurate comparison of the two institutional pilots, 309 students wrote the online version, and 1,983 wrote the pencil-and-paper version. The test was administered in two parts, written in two consecutive two-hour sessions one week apart. The study obtained ethical clearance for all the pilot administrations of the test between 2018 and 2023.

VALIDATION OF THE APPMI

Validity is a complex and multifaceted concept with various interpretations (Van der Walt & Steyn, 2007). There is also a distinction between validity as a theoretical concept and validation, which is its practical application (Davies & Elder, 2005). Essentially, validity concerns the extent to which a test measures what it was designed to measure. The process of validation is where one provides evidence to support various claims made about the test. It is through "validation that validity is established, which means that validity is only as good as its validation procedures" (Davies & Elder, 2005: 795). Van der Walt and Steyn (2007: 142) maintain that this process involves formulating a validity argument in the form of claims supported by relevant evidence. The following section provides a brief validation argument for the APPMI.

Claim 1: The test presents consistent reliability measures

Reliability is defined as "the trustworthiness or the accuracy of a measurement" (Kurpius & Stafford, 2006: 121). This aspect concerns the consistency of scores across tests in terms of the abilities being measured. Various tests were conducted to measure the technical reliability of the APPMI, the results of which are shown in Table 4. Cronbach's alpha coefficient values and greater lower bound (GLB) values in the vicinity of 0.85 and 0.9 are typically evidence of adequate test-level reliability. The results of previous administrations between 2018 and 2021 reflected Cronbach's alpha coefficient values consistently higher than 0.85. Similar results were achieved for the 2023 pilot with a Cronbach's alpha coefficient of 0.88. These results were confirmed by the Rasch person reliability measure above the desired 0.8 range (McNamara, Knoch & Fan, 2019: 52), as well as GLB values above 0.9 (results for the combined result did not calculate). The results of these various measures serve as evidence of the technical consistency of the APPMI.

Table 4: Reliability (Cronbach's alpha and GLB) and related indicators: APPMI

APPMI results	**2023 pilot**		
	Traditional (n=1983)	**Online (n=309)**	**Combined (n=2292)**
Cronbach's alpha	0.88	0.87	0.88
GLB	0.93	0.97	-
Rasch person reliability			0.88
Ave *P*-value	52.14	53.03	52.91
Ave *Rit* value	0.30	0.31	0.31

Claim 2: Heterogeneous test items do not compromise the reliability of the test

A key principle concerning the technical integrity of the test is whether there is "unity within a multiplicity of components" (Weideman, 2019: 43). One measure of the technical quality of a test is its homogeneity. This refers to the extent to which the test is an integral whole and "whether all the test items measure the same trait (one factor)" (CITO, 2005: 19). Figure 1 shows the factor analysis for the 2023 pilot, which indicates a homogenous construct. The majority of the items, except for those in the *Organisation of Text* subtest, cluster within a fairly tight range. Although the discrimination and facility values for this particular subtest were within range for previous pilots, the facility or difficulty level (P-value) for the 2023 pilot was 0.45, which is slightly lower than the desired 0.5. However, the discrimination value (*Rit*) for this subtest was within the acceptable range of above 0.3 for the 2023 administration. Despite the five apparently outlying items, none measure a second factor, which indicates that there is a sufficient degree of homogeneity and that this test design principle has been satisfied.

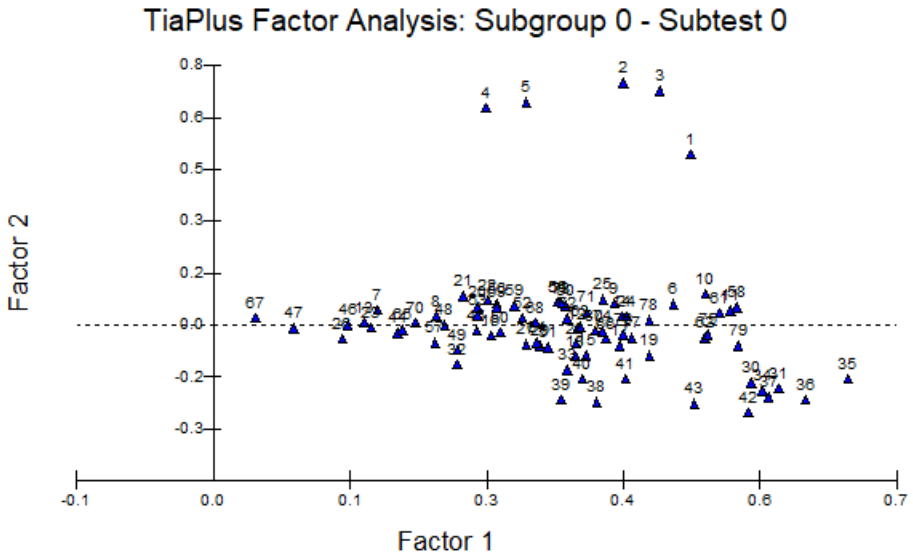

Figure 1: Factor analysis of 2023 APPMI pilot

Claim 3: The test demonstrates an acceptable degree of fit between individual ability and item difficulty

A Rasch analysis was performed to determine the degree of fit between individual ability and item difficulty. An infit MNSQ of 1.0 is the average benchmark for measuring the degree of fit. Typically, values exceeding 1.5 are considered problematic. However, a more conservative range of between 0.75 and 1.3 was used to measure the average fit of the APPMI. Table 5 presents a shortened version of the Rasch analysis results. The infit mean square values were within range for all the test items, including the two terminal values of 0.86 and 1.17. These results confirm the technical soundness of the test in terms of overall fit.

Table 5: Misfit order: items in APPMI

Item	Total count (n)	Infit MNSQ	Ptmeasure-AL Corr	Expected
47	2291	1.14	0.10	0.30
67	2291	1.17	0.08	0.31
46	2291	1.08	0.14	0.27
23	2291	1.09	0.16	0.30
26	2291	1.12	0.14	0.31
53	2291	1.06	0.25	0.31
22	2291	1.05	0.26	0.31
56	2291	1.04	0.27	0.31

59	2291	1.04	0.27	0.31
64	2291	1.04	0.28	0.31
Further better fitting items not shown				
2	2291	0.97	0.35	0.31
17	2291	0.97	0.34	0.30
24	2291	0.97	0.35	0.31
41	2291	0.97	0.35	0.31
66	2291	0.97	0.31	0.26
79	2291	0.91	0.42	0.29
31	2291	0.90	0.44	0.29
37	2291	0.90	0.44	0.31
36	2291	0.88	0.47	0.31
35	2291	0.86	0.47	0.28

Claim 4: The test items present acceptable discrimination values

The discrimination ability of a test refers to its ability to differentiate between students who answered questions correctly and those who did not (Kurpius & Stafford, 2006). Therefore, a TiaPlus analysis was performed to determine the discrimination ability of the test items. Table 6 displays the overall facility (*P*-value) and discrimination (*Rit*) values for the APPMI. The results show acceptable facility (within the range of 0.5) and discrimination (above 0.30) values for the traditional, online and overall test for the 2023 pilot. The 2021 institutional pilot yielded similar discrimination (*Rit* = 0.29) and facility (*P*-value = 51.29) results.

Table 6: Average Rit-values of APPMI

APPMI results	**2023 pilot**		
	Traditional (n=1983)	**Online (n=309)**	**Combined (n=2292)**
Ave *P*-value	52.14	53.03	52.91
Ave *Rit* value	0.30	0.31	0.31

Claim 5: The various subtests meet internal correlation criteria

Another key principle concerns the ability of the test to function as a "technically differentiated but whole assessment" (Weideman, 2021: 64). This refers to the extent to which the various subtests, measuring different sub-abilities, work together with other subtests and the overall test as an organised whole. As indicated in Table 6, the average discrimination index (*Rit* value) for the 2023 pilot is slightly higher for the online test than the traditional format. However, an examination of Differential Item Functioning (DIF), which compares the difficulty of items between the two versions of the test, indicated no significant difference (z-stand <2.58). Thus, the administrations were closely similar, and the results for both versions could be combined for further analysis.

Since the subtests were designed to measure different abilities that together form part of what is understood to be academic literacy, the correlations between subtests should be relatively low (0.3–0.5). Ideally, each subtest should contribute something unique to the overall test. A much higher correlation value than the desired range (i.e., closer to +1) could indicate that the subtests measured the same ability and did not offer unique information about the ability being tested (Alderson, Clapham & Wall, 1995: 184). However, in terms of the correlations between the subtests and the whole test, these should be "around +.7 or more since the overall score is taken to be a more general measure of language ability than each individual component score" (Alderson et al., 1995: 184).

Table 7 shows the test-subtest and inter-correlations for the 2023 pilot. The test–subtest correlations for half of the subtests were below the desired 0.6 (Weideman, 2021). This could be attributed to the number of items, as the majority contained five or fewer items. The two subtests for which inter-correlation values fell outside the conservative parameters of between 0.2 and 0.5 (Weideman, 2021) were *Organising information visually* and *Text type and communicative function*. The 2021 pilot reported similar results for the test–subtest correlations and inter-correlation related to these two subtests. Although the results showed test–subtest correlations of below 0.60 for three other subtests, the estimated Spearman-Brown coefficient scores would fall within the desired parameters (above 0.85) if these subtests contained the standard 40 items. Future administrations might be of value to determine whether it is worth removing the problematic subtests.

Table 7: Test–subtest correlations and subtest inter-correlations (n=2292) for 2023 APPMI pilot

Subtest	**Total test**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Alpha (40+)**
1 Organising text	0.45											0.97
2 Vocabulary	0.60	0.26										
3 Visual and graphic information	0.61	0.16	0.33									0.85

4 Arguments (1)	0.39	0.16	0.20	0.17								0.85
5 Arguments (2)	0.47	0.16	0.26	0.24	0.23							0.85
6 Grammar and text relations	0.75	0.21	0.34	0.36	0.19	0.29						0.92
7 Organising information visually	0.48	0.11	0.21	0.25	0.13	0.16	0.26					0.74
8 Text type and comm? function	0.43	0.14	0.21	0.24	0.10	0.13	0.19	0.15				0.94
9 Text comprehension	0.77	0.25	0.44	0.41	0.25	0.30	0.42	0.32	0.27			0.80
10 Text editing	0.63	0.21	0.35	0.34	0.17	0.24	0.40	0.27	0.20	0.43		0.91
		1	2	3	4	5	6	7	8	9	10	
Number of items:	80	5	6	8	4	4	16	8	5	18	6	
Average test score:	42.33	2.23	2.62	3.45	2.09	2.24	9.14	4.66	2.07	10.18	3.64	
Standard deviation:	11.76	1.83	1.47	1.81	1.16	1.13	3.92	1.61	1.62	3.29	1.66	
SEM:	4.13	0.67	1.06	1.17	0.88	0.86	1.10	1.23	0.82	1.85	1.02	
Average P-value:	52.91	44.69	43.72	43.16	52.16	55.98	57.11	58.26	41.37	56.58	60.71	
Coefficient Alpha:	0.88	0.80	0.45	0.52	0.37	0.37	0.80	0.36	0.68	0.64	0.61	
GLB:	0.00	0.86	0.49	0.60	0.42	0.46	0.92	0.43	0.75	0.70	0.63	
Asymptotic GLB:	0.00	0.87	0.48	0.58	0.42	0.42	0.92	0.42	0.74	0.68	0.62	

Claim 6: The test is fair

Another key consideration is test fairness, which refers to the degree to which "a test is used in an impartial, just and equitable way" (Cohen & Swerdlik, 2010: 203). As noted, there is no significant difference between the online and traditional test formats, so no student was potentially disadvantaged or advantaged by the mode of administration. A further consideration concerning test fairness is misclassification (Van der Slik & Weideman, 2010). This aspect can be assessed through the alpha-based, same test or parallel test [Rxx or Rxt case] results produced by the TiaPlus analysis (CITO, 2005). Assuming an equal chance of misclassification, Table 8 shows that, at worst, 6.3% of candidates might have been misclassified.

Table 8: Potential misclassifications in the administration of APPMI (2023)

Alpha-based	
- Rxx' case: Percentage	6.3%

Number	145
- Rxt case: Percentage	4.6%
Number	105

The NBT defines a cut-off score of 64% to determine which students are required to take literacy courses. Both correlation and regression analyses were performed between the 2022 NBT and 2023 APPMI scores to equate this to performance on the APPMI (see Drennan et al., 2021). The results indicated a strong correlation of 0.6 ($p < 0.0001$), while the regression analysis indicated a statistically significant ($p < 0.05$) relationship, with the APPMI scores accounting for ~33% of the variation in NBT scores (adjusted r-score = 0.33). Given the high variability in the NBT scores, and given all results, the APPMI test can be considered a good predictor of language ability. From the regression analysis, an APPMI score of 73% predicts the NBT cut-off score of 64%.

The discussion above and the various claims made about the validity of the APPMI serve as a justification for responsible test design. The theoretical claims and corresponding evidence support the argument about the validity and reliability of the APPMI as a measure of academic literacy. In terms of potential future administrations, a second analysis was conducted to determine the potential effect of removing the two problematic subtests (*Organising information visually* and *Text type and communicative function*) on the functionality, discrimination, and reliability of the APPMI. Both institutional pilots (2021 and 2023) reflected test–subtest correlations (0.43 and 0.43) and inter-correlations (below 0.2) that fell outside the required parameters. The results of this shortened version of the test indicated a Cronbach's alpha (reliability) score of 0.87, as well as facility and discrimination values within the desired range of 0.53 and 0.33, respectively. The results of a Rasch analysis for the degree of fit also indicated values within the desired range (terminal values between 0.86 and 1.19 on the infit mean square). Therefore, removing these items does not compromise the technical soundness of the test, and they could be removed profitably for future administrations of the test.

CONCLUSION

Academic literacy support has become critically important in the higher education context, whereby historical inequalities continue to impact students' access and success. The sector has witnessed an influx of students from previously disadvantaged backgrounds who are not adequately prepared for the demands of tertiary studies. Substandard language education at the school level has resulted in a mismatch between students' school-leaving language abilities and their capacity to deal with the demands of academic discourse. This situation has caused high attrition rates and students' failure to complete their studies within the scheduled time.

Standardised tests of academic literacy, such as the NBT and TALL, were developed to address this challenge by assessing students' academic preparedness for appropriate placement at

universities. However, the disruption in the administration of conventional tests during the COVID-19 pandemic had significant implications for identifying students requiring academic literacy support, highlighting the pressing need for more accessible and flexible testing mechanisms. The pilot results of an in-house test of academic literacy reported in this paper demonstrate the validity and reliability of the test as a viable alternative measure of academic literacy. Such in-house tests offer institutions the opportunity to accommodate logistical challenges around test administration. Another possible affordance, albeit beyond the scope of this study, concerns the potential for in-house tests to be more context-sensitive than generic, large-scale standardised tests. The development of more context-sensitive, in-house tests would require additional information, including lecturer and student perceptions of disciplinary literacy needs, as well as extensive analyses of disciplinary genres, text types, and conventions. The results of such tests could better inform the development of literacy content aligned with the literacy needs of the target student body. Such adaptability could ensure testing measures that are fair in terms of access and equitable in their sensitivity to the needs of a diverse student population.

REFERENCES

Alderson, JC, Clapham, C & Wall, D. 1995. *Language test construction and evaluation.* Cambridge: Cambridge University Press.

Bachman, LF & Palmer, AS. 1996. *Language testing in practice.* Cambridge: Cambridge University Press.

Blanton, LL. 1994. Discourse, artefacts and the Ozarks: Understanding academic literacy. *Journal of Second Language Writing,* 3(1):1–16.

Carpenter, R & Roos, L. 2020. Can we afford it? The association between financial aid and time to completion in Accounting Higher Education in South Africa: A literature review. *The Business and Management Review,* 11(1):222–232.

Chan, S. 2013. *Establishing the validity of reading-into-writing test tasks for the UK academic context.* Bedfordshire: University of Bedfordshire.

Chan, S. 2018. Some evidence of the development of L2 reading-into-writing skills at three levels. *Language Education and Assessment,* 1(1):9–27.

CITO. 2005. *TiaPlus, classical test and item analysis.* Arnhem: Cito Measurement and Research Department.

Cloete, N. 2016. For sustainable funding and fees, the undergraduate system in South Africa must be restructured. *South African Journal of Science,* 112(3-4): 1–5.

Cohen, RJ & Swerdlik, ME. 2010. *Psychological testing and assessment.* New York: McGraw-Hill.

Cummins, J. 2000. *Language, power, and pedagogy: Bilingual children in the crossfire.* Clevedon: Multilingual Matters.

Cummins, J. 2009. Multilingual education: Approaches and constraints. In Skutnabb-Kangas,T, Phillipson, R, Mohanty, AK & Panda, M (Eds.), *Social justice through multilingual education*. Bristol: Multilingual Matters.

Cummins, J & Swain, M. 1986. *Bilingualism in education.* New York: Longman.

Davies, E & Elder, C. 2005. Validity and validation in language testing. In Hinkel, E *Handbook of research in second language teaching and learning*. New Jersey: Lawrence Erllbaum Associates.

DHET. 2016. *Statistics on the second national higher education transformation summit.* Durban: Department of Higher Education and Training.

Douglas, D. 2000. *Assessing language for specific purposes.* Cambridge: Cambridge University Press.

Drennan, L. 2011, November. Investigating the efficacy of the 'Skills for a Changing World' first-year literacy course. Bloemfontein: University of the Free State.

Drennan, L. 2019, September. *Defensibility and accountability: Developing a theoretically justifiable academic writing intervention for students at tertiary level.* Bloemfontein: University of the Free State.

Drennan, L. 2021. Assessing readiness to write: The design of an assessment to present multimodal information (APPMI). In A Weideman, Read, J & Du Plessis, T, *Assessing academic literacy in a multilingual society: Transition and transformation* (pp. 196–216). Bristol: Multilingual Matters.

Drennan, L, Joubert, M, Weideman, A & Posthumus, R. 2021. Combined measures of identifying language risk for first-year students. *Journal for Language Teaching, 55*(2), 93–115.

Du Plessis, C. 2021. Basic education and academic literacy: Conflicting constructs in the South African National Senior Certificate (NSC) language examination. In Weideman, A, Read, J & Du Plessis, T, *Assessing academic literacy in a multilingual society: Transition and transformation*. Bristol: Multilingual Matters.

Fourie, CM. 2020. Risk factors associated with first-year students' intention to drop out from a university in South Africa. *Journal of Further and Higher Education,* 44(2):201–215.

Frederiksen, CH. 1975. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 7:317–458.

Gee, JP. 1989. Literacy, discourse, and linguistics. *Journal of Education,* 171(1):5–17.

Gumede, M. 2017. *Universities put focus on high drop-out rate*. Available from https://www.businesslive.co.za/bd/national/education/2017-03-28-universities-put-focus-on-high-drop-out-rate [Accessed: 23 March, 2024]

Howie, S, Venter, E & Van Staden, S. 2008. The effect of multilingual policies on performance and progression in reading literacy in South African primary schools. *Educational Research and Evaluation,* 14(6):551–560.

Huysamen, GK. 2000. The differential validity of matriculation and university performance as predictors of post-first-year performance. *South African Journal of Higher Education,* 14(1):146–151.

Kurpius, SE & Stafford, ME. 2006. *Testing and measurement: A user-friendly guide.* California: Sage Publications.

Lea, MR & Street, BV. 1998. Student writing in higher education: ANA academic literacies approach. *Studies in Higher Education, 23*(2):157–172.

McKay, TM. 2016. Academic success, language, and the four year degree: A case study of a 2007 cohort. *South African Journal of Higher Education,* 30(4):190–209.

McNamara, T, Knoch, U & Fan, J. 2019. *Fairness, justice and language assessment.* Oxford: Oxford University Press.

Myburgh, J. 2015. *The assessment of academic literacy at pre-university level: A comparison of the utility of academic literacy tests and Grade 10 Home Language results.* Bloemfontein: University of the Free State.

Myburgh-Smit, J & Weideman, A. 2017. Refinement and uses of a test of academic literacy for Grade 10 students. *Journal for Language Teaching,* 51(1):271–295.

Ncgobo, S. 2009. Lecturers' and students' reflections on a bilingual programme. In Leibowitz, B, Van der Merwe, A & Van Schalkwyk, S, *Focus on first-year success: Perspectives emerging from South Africa and beyond*. Stellenbosch: SUN Media.

Nel, C & Nel, C. 2009. A three-tier model for supporting reading-to-learn. In Leibowitz, B, Van der Merwe, A & Van Schalkwyk, S, *Focus on first-year success: Perspectives emerging from South Africa and beyond*. Stellenbosch: SUN Media.

Patterson, R & Weideman, A. 2013. The typicality of academic discourse and its relevance for constructs of academic literay discourse. *Journal of Language Teaching,* 47(1):107–123.

Pot, A. 2013. *Diagnosing academic language ability: An analysis of TALPS.* Groningen: Rijksuniversiteit Groningen.

Pot, R & Weideman, A. 2015. Diagnosing academic language ability: An analysis of the Test of Academic Literacy for Postgraduate Students. *Language Matters,* 46(1):22–43.

Sebolai, K. 2016. Distinguishing between English proficiency and academic literacy in English. *Language Matters,* 47(1):56–60.

Sebolai, K & Huff, L. 2015. Academic literacy curriculum renewal at a South African university: A case study. *Journal for Language Teaching,* 49(1):333–351.

Sebolai, K & Stanford, E. 2020. Validating the highest performance standard of a test of academic literacy at a South African university. *Per Linguam*, 36(2):76–89.

Spivey, NN. 2001. Discourse synthesis: Process and product. In McInnis, RG, *Studies in historical and contemporary social epistemology*. Westport, CT: Praeger.

Spivey, NN & King, JR. 1989. Readers as writers composing from sources. *Reading Research Quarterly,* 7–26.

Thesen, L & Van Pletzen, E. 2006. *Academic literacy and the languages of change*. London: Continuum.

Van der Slik, F & Weideman, A. 2010. Examining bias in a test of academic literacy: Does the Test of Academic Literacy Levels (TALL) treat students from English and African language backgrounds differently? *Journal for Language Teaching,* 27(3):106–118.

Van der Walt, JL & Steyn, HS. 2007. Pragmatic validation of a test of academic literacy at tertiary level. *Ensovoort,* 11(2):138–153.

Van Dijk, TA. 1979. Relevance assignment in discourse comprehension. *Discourse Processes,* 2: 113–126.

Van Dyk, T & Van de Poel, K. 2013. Towards a responsible agenda for academic literacy development. *Journal for Language Teaching,* 47(2):43–49.

Van Rooy, B & Coetzee-Van Rooy, S. 2015. The language issue and academic performance at a South African University. *Southern African Linguistics and Applied Language Studies,* 33(1):1–16.

Weideman, A. 2007. The redefinition of applied linguistics: Modernist and postmodernist views. *Southern African Linguistics and Applied Language Studies,* 25(4):589–605.

Weideman, A. 2013. Academic literacy interventions: What are we not yet doing, or not yet doing right? *Journal for Language Teaching,* 47(2):11–23.

Weideman, A. 2021. A skills-neutral approach to academic literacy. In Weideman, A, Read, J & Du Plessis, T, *Assessing academic literacy in a multilingual society: Transition and transformation*. Bristol: Multilingual Matters.

Weideman, AJ. 2019. Degrees of adequacy: The disclosure of levels of validity in language assessments. *Koers,* 84(1):1–15.

Weideman, A, Patterson, R & Pot, A. 2016. Construct refinement in tests of academic literacy. In J Read, *Post-admission language assessment of university students*. Cham: Springer.

Young, K & Fynn, A. 2024. Grit, retention and student success in a South African distance education insitution: A postgraduate Triad? *Turkish Online Journal of Distance Education,* 25(2):144–156.

BIOGRAPHICAL NOTES

Laura Drennan is currently a lecturer in the Division of Languages, Literacies and Languages at the Wits School of Education. She previously served as the Head of the Writing Centre at the University of the Free State, where she specialised in academic writing development. She obtained her doctorate in English Language Studies and academic literacy development. Her research interests include language teaching and learning, academic literacy, academic writing development, and language testing.
E-mail address: laura.drennan@wits.ac.za