

Paper for presentation at AAAL 2014 colloquium on  
*Exploring post-admission language assessments in  
universities internationally*

## **Construct refinement in tests of academic literacy**

Albert Weideman

*Department of Linguistics and Language Practice, University of the Free State &  
Inter-Institutional Centre for Language Development and Assessment (ICELDA)*

Rebecca Patterson

Anna Pot

*Research fellows, Department of Linguistics and Language Practice, University of  
the Free State*

### **Abstract**

For several reasons, the construct underlying post-entry tests of academic literacy in South Africa such as the Test of Academic Literacy Levels (TALL) and its postgraduate counterpart, the Test of Academic Literacy for Postgraduate Students (TALPS), deserves further scrutiny. First, though their construct derives from a developmental line that includes ideas that are still current, it has not been further investigated in close to a decade of use. Second, acknowledging the typicality of academic discourse (Patterson & Weideman 2013) as a starting point for critically engaging with constructs of academic literacy may suggest design changes for such tests. In this contribution, various attempts at identifying the typical features of academic discourse are therefore surveyed and critiqued. The preliminary conclusion is that the uniqueness of academic discourse lies in the primacy of the logical or analytical mode that guides it. Using this characteristic feature as a criterion is potentially productive: it suggests ways in which one might add components to the current definition of academic literacy that forms the test construct of academic literacy tests that are widely used in South Africa, such as TALL, TAG (the Afrikaans counterpart of TALL), and TALPS, as well as a new test of academic literacy for Sesotho. Third, a recent analysis of the diagnostic information that can be gleaned from one of these, TALPS (Pot 2013), may inform strategies of utilising post-entry tests of language ability (PELAs) more efficiently. The paper includes suggestions for modifications and additions to the design of current test task types in tests of academic literacy. These tentative suggestions allow theoretically defensible modifications to the design of the tests, and will be useful to those responsible for developing further versions of these tests of academic literacy.

## **The continuing importance of construct**

Responsible language test developers are required to start by examining and articulating with great care the language ability that they will be assessing (Weideman 2011). That is so because their definition of this ability, the formulation of the construct of the hypothesized competence that will be measured, is the first critically important step to ensure that they will be measuring fairly and appropriately. What is more, it is from this point of departure that the technical effectiveness or validity of the design process will be steered in a direction that might make the results interpretable and useful, two essential ingredients in what is the currently orthodox notion of test validation (Read 2010: 288; Chapelle 2012). It will also help to ensure that the instrument itself is relevant, appropriate, and reliable, and that its uses and impact are beneficial (Knoch & Elder 2013: 54f.).

The quest for a clear definition of the ability to be measured is complicated, however, and the ultimate choice of that definition is not devoid of compromise (Knoch & Elder 2013: 62f.; Van Dyk & Weideman 2004b). One reason for that is that some definitions of language ability may be more easily operationalisable than others. A construct has to be translated by test designers into specifications that include, amongst other things, the determination of which task types and assessment formats will be used (Davidson & Lynch 2002). It follows that test specifications must align with the definition if the test design is to be theoretically and technically defensible. Language tasks that are typical of the kind of discourse that is the target of the assessment should predominate. So it is not surprising that reference to these remains uppermost in the minds of those who develop tests with the objective of assessing the level of mastery of that type of discourse. Yet some compromises may have to be made, not the least because test developers are constrained by any number of administrative, logistical, financial and other resource limitations, and might have to choose for their test those task types that

best operationalise the construct within those constraints (Van Dyk & Weideman 2004b). The result of this may be that parts or components of a theoretically superior definition of language ability may either be overlooked or under-emphasised. While a tight formulation of test specifications may curb this, some difficult design decisions might still need to be made. A further complication that presents itself is that test designers may, once the test has been administered and used, realise that parts of it may be providing less useful or beneficial information on the ability of the candidates who sat for it, so that they require adjustment or redesign (McNamara & Roever 2006: 81f.). Even the most deliberate design and careful piloting of a test is no guarantee that it will be perfect the first or even the twelfth time it is administered. As the validation processes of any test might reveal, redesign may be needed at any time, but what is being argued here is that the starting point is always the construct.

There is a third potential complication in trying to stay true to the definition of the language ability being tested, which is that new insight into the workings of language may allow us to gauge that ability better. The turn in language testing towards a disclosed, interactive way of looking at language instead of a merely formal, restrictive view of what it is (Green 2014: 173ff.; Chapelle 2012) constitutes an example of this. New perspectives on language must of necessity have an effect on what is tested.

A fourth difficulty that we have encountered in test design is where the test of language ability depends on the curriculum of a national school language syllabus, as in South Africa (Department of Basic Education 2011a; 2011b). Here the high-stakes examinations that make up the Grade 12 school exit examinations for “Home Languages” have patently over time drifted away from the original intentions of the official curricula (see report to Umalusi by Du Plessis, Steyn & Weideman 2013). As this report makes clear, the best possible way to restore the integrity of these

language assessments is to reinterpret the assessment task with reference to the curriculum, which specifies the language construct to be tested.

Despite the possible need for compromise referred to above, or the undesirable potential of moving away from what by definition must be assessed in a test of language ability, this paper takes as its point of departure that the clear articulation of a construct remains the best guarantee for responsible test design. This is indeed also one of the limitations of the paper, since responsible test design patently derives from much more than a (re-)consideration of the construct. There are indeed many other points of view and design components that might well be useful points of entry for design renewal, and that will, with one exception, not be considered here, given the focus of the paper on a single dimension: the theoretically defensible definition of the ability to use language in a very specific domain. With this limitation in mind, the paper will therefore consider how a re-examination of the construct, both from the point of view of a redefinition of the target domain and from the angle of the varying emphases that such redefinition might involve, may further inform test design. It will start by looking at the development of a current construct underlying post-entry tests of academic literacy in South Africa such as the Test of Academic Literacy Levels (TALL), the Toets van Akademiese Geletterdheidsvlakke (TAG, the Afrikaans version of TALL), and their postgraduate counterpart, the Test of Academic Literacy for Postgraduate Students (TALPS). Given that the construct they share has been in use for a decade, it will be argued that it deserves scrutiny at least for its being in use for some time already. Finally, the paper will examine whether there are typical components of the ability to use academic discourse competently that might have been overlooked or underemphasised, and how that might be corrected in subsequent test designs.

## The current construct and its theoretical lineage

In Van Dyk and Weideman (2004a & 2004b) there are detailed descriptions of the process of how the construct underlying the current tests referred to above was developed, and how the specifications for the blueprint of the test were arrived at. The test designers of TALL, TAG and TALPS were looking for a definition of academic literacy that was current, relevant, and reflected the use of academic discourse in a way that aligned with the notions that academics themselves have about that kind of language.

Yet finding such a construct that was also theoretically defensible involved a long process. The construct eventually adopted therefore derives from a developmental line that looks at language as a disclosing, open interaction, and not as an object restricted to a combination of sound, form and meaning. Moreover, as Weideman (2003) points out, it was required to take a view of the development of the ability to use academic language as the acquisition of a secondary discourse (Gee 1998). Acquiring that secondary discourse in becoming academically literate, as Blanton (1994: 230) notes, happens when

... individuals whom we consider academically proficient speak and write with something we call *authority*; that is one characteristic — perhaps the major characteristic — of the voice of an academic reader and writer. The absence of authority is viewed as powerlessness ...

How does one assess ‘authority’ as a measure of proficiency, however? And how does one characterise the ‘academic’ that stamps this authority as a specific kind? Even though the elaboration of this ability to use academic language fluently was wholly acceptable to the designers of the tests we are referring to in this paper, the question was how to operationalise every one of its components, formulated by Blanton (1994: 226) as the abilities of students to be able to:

1. Interpret texts in light of their own experience and their own experience in light of texts;

2. Agree or disagree with texts in light of that experience;
3. Link texts to each other;
4. Synthesize texts, and use their synthesis to build new assertions;
5. Extrapolate from texts;
6. Create their own texts, doing any or all of the above;
7. Talk and write about doing any or all of the above;
8. Do numbers 6 and 7 in such a way to meet the expectations of their audience.

Blanton's definition is noteworthy, nonetheless, because it does not define learning to become competent in academic language as knowledge merely of sound, form, and meaning. In fact, it stresses that academic discourse is communicative, interactional, and contextual. The formulations above could indeed be translated into test specifications, but were adjudged to be highly likely to result in a resource-intensive instrument. The logistical and administrative constraints anticipated prompted the test designers to look further.

A consideration of the outline of language ability in the work of Bachman and Palmer (1996) subsequently provided a second, related perspective on defining academic literacy, the ability the test designers wished to test. This definition of language competence, widely used in the field of language testing for the justification of designs, sees language ability as having two pillars: language knowledge and strategic competence (1996: 67). Whatever the advantages of this broad outline of language ability – and it also exhibits several disadvantages, not the least of which is seepage amongst the various sub-categories it proposes - it was found to be difficult, at least in the case of academic discourse, to contextualise. Specifically, it was not apparent in every case how one might fill with content the various sub-categories of these two hypothesized components of language ability. A third perspective that could potentially solve this problem was provided by the work being done at the Alternative Admissions Research Project (AARP) at the University of Cape Town. The AARP (Yeld *et al.* 2000) reinterpretation of the Bachman and Palmer (1996) construct adds “understandings of typical academic tasks based largely on inputs from expert panels” (Yeld *et al.* 2000; cf. too Cliff & Hanslo 2005, Cliff, Yeld & Hanslo 2006). The construct is therefore enriched by the identification, amongst other things, of a number of language functions and

academic literacy tasks. A streamlined version of this eventually became the final blueprint for the tests of academic literacy developed by a consortium of four multi-lingual universities, ICELDA (ICELDA 2014). This definition had as its goal the development of tests that would gauge the ability of students to

- understand a range of academic vocabulary in context;
- interpret and use metaphor and idiom, and perceive connotation, word play and ambiguity;
- understand relations between different parts of a text, be aware of the logical development of (an academic) text, via introductions to conclusions, and know how to use language that serves to make the different parts of a text hang together;
- interpret different kinds of text type (genre), and show sensitivity for the meaning that they convey, and the audience that they are aimed at;
- interpret, use and produce information presented in graphic or visual format;
- make distinctions between essential and non-essential information, fact and opinion, propositions and arguments; distinguish between cause and effect, classify, categorise and handle data that make comparisons;
- see sequence and order, do simple numerical estimations and computations that are relevant to academic information, that allow comparisons to be made, and can be applied for the purposes of an argument;
- 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;
- understand the communicative function of various ways of expression in academic language (such as defining, providing examples, arguing); and
- make meaning (e.g. of an academic text) beyond the level of the sentence.

(Weideman 2007: xi)

In wrestling with how academic discourse may be defined, the authors of this proposed streamlined version of the construct shared it with colleagues from a range of disciplines, in various fora and publications. The responses they received confirmed the results of the initial consultations at the time that the AARP reinterpretation was being constructed: the elements identified above not only constitute at least a number of essential components of what academic literacy entails, but resonate strongly with what academics across the disciplinary spectrum think constitutes the competent use of academic discourse (Weideman 2003).

On the basis of this refined definition of academic discourse, the designers of the tests being surveyed here began to experiment with 10 different task type formats.

In this their work departed from the AARP designs that emanate from essentially the same construct, since these worked with less than half that number of task types. Eventually the test designers settled on seven that fulfilled the conditions of conforming to the test construct and its detailed specifications (Van Dyk & Weideman 2004b). In the table below, the task types are related to the component of the construct (the basis of the specification) in the first column, with the primary task types that test these components indicated in the second column, as well as the potential, secondary task type(s) that assess them. It is clear that some of the components of the construct can potentially be measured in more than one of the possible subtests, while the reverse is also true: a subtest can potentially give insight into the ability to handle one or several components of the construct:

<b>Specification</b>	<b>Task type(s) / Possible subtest(s)</b>
Vocabulary comprehension	Knowledge of academic vocabulary Grammar and text relations (modified cloze) Understanding texts (longer reading passage)
Understanding metaphor & idiom	Understanding texts (longer reading passage) Text type / register
Textuality (cohesion and grammar)	Scrambled text Grammar and text relations (modified cloze) (perhaps) Text type / register Understanding texts (longer reading passage) Academic writing task(s)
Understanding text type (genre)	Text type / register Interpreting visual & graphic information Scrambled text Grammar and text relations (modified cloze) Understanding texts (longer reading passage) Academic writing task(s)
Understanding visual & graphic information	Interpreting visual & graphic information (potentially) Understanding texts (longer reading passage)
Distinguishing essential/non-essential	Understanding texts (longer reading passage) Interpreting visual & graphic information Academic writing task(s)
Numerical computation	Interpreting visual & graphic information Understanding texts (longer reading passage)
Extrapolation and application	Understanding texts (longer reading passage) Academic writing task(s)

	(possibly: Interpreting visual & graphic information)
Communicative function	Understanding texts (longer reading passage) (possibly also: Scrambled text; Grammar and text relations [modified cloze])
Making meaning beyond the sentence	Understanding texts (longer reading passage) Text type / register Scrambled text Interpreting visual & graphic information

**Table 1: Specifications and subtests for a test of academic literacy**

### **The typicality of academic discourse**

While the tests that were based on this construct have now been widely scrutinised and their results subjected to empirical and critical analyses of various kinds (the ‘Research’ tab of ICELDA 2014 lists more than three dozen such publications), the construct referred to above has not been further investigated in close to a decade of use. In two recent studies of the construct undertaken to remedy this lack of critical engagement, Patterson and Weideman (2013a, 2013b) take as a starting point the typicality of academic discourse as a kind of discourse distinct from any other. They begin by tracing the idea of the variability of discourse to the sociolinguistic idea of a differentiated ability to use language that goes back to notions first introduced by Habermas (1970), Hymes (1971), and Halliday (1978), noting at the same time how those ideas have persisted in more current work (Biber & Conrad 2001; Hasan 2004; Hyland & Bondi 2006). Specifically, they consider how acknowledging that academic discourse is a specific, distinctly different kind of language will benefit construct renewal.

The typicality of academic discourse, viewed as a material linguistic sphere (Weideman, 2009: 40f.), is found to be closely aligned with the views espoused by Halliday (1978; 2002; 2003), specifically the latter’s ideas of “field of discourse”, genre, rhetorical mode, and register. Halliday’s claim (1978:202; cf. too Hartnett 2004:183) that scientific language is characterised by a high degree of

nominalisation, however, is deficient in several respects. First, there are other kinds of discourse (e.g. legal and administrative uses of language) in which nominalisation is also found. Second, it gives only a formal criterion for distinctness, neglecting the differences in content that can be acknowledged when one views types of discourse as materially distinct.

When one turns to a consideration of various current definitions of academic literacy, Patterson and Weideman (2013a) find that there are similar problems. In the ‘critical’ features of academic discourse identified by scholars such as Flower (1990), Suomela-Salmi and Dervin (2009), Gunnarsson, (2009), Hyland (2011; cf. too Hyland and Bondi 2006), Livnat (2012), Bailey (2007:10-11), and Beekman, Dube and Underhill (2011:1), for example, we find either circular definitions that identify ‘academic’ with reference to the academic world itself, or features that are shared across a number of discourse types. As Snow and Uccelli (2009) observe, the formally conceptualised features of academic language that they have articulated with reference to a wide range of commentators are not sufficient to define academic discourse.

Patterson and Weideman (2013a: 118) conclude that an acknowledgement of the typicality of academic discourse that is most likely to be productive is one that acknowledges both its leading analytical function and its foundational formative (‘historical’) dimension:

Academic discourse, which is historically grounded, includes all lingual activities associated with academia, the output of research being perhaps the most important. The typicality of academic discourse is derived from the (unique) **distinction-making** activity which is associated with the analytical or logical mode of experience.

What is more, if one examines the various components of the construct referred to above, it is clear that the analytical is already prominent in many of them. For

example, in logical concept formation, which is characterised by abstraction and analysis (Strauss, 2009: 12-14), we proceed by comparing, contrasting, classifying and categorising. All of these make up our analytical ability to identify and distinguish.

### **Potential additions to the construct, and ways to assess them**

As Patterson and Weideman (2013b) point out, a number of the components of the current construct already, as they should, foreground the analytical qualifying aspect of academic discourse. Which components should in that case potentially be added? Having surveyed a range of current ideas, these investigators identify the following possible additions (*in italics*) as components of a command of academic language that can be demonstrated through the ability to:

- *think critically and reason logically and systematically in terms of one's own research and that of others;*
- *interact (both in speech and writing) with texts: discuss, question, agree/disagree, evaluate, research and investigate problems, analyse, link texts, draw logical conclusions from texts, and then produce new texts;*
- *synthesize and integrate information from a multiplicity of sources with one's own knowledge in order to build new assertions, with an understanding of academic integrity and the risks of plagiarism;*
- *think creatively: imaginative and original solutions, methods or ideas which involve brainstorming, mind-mapping, visualisation, and association;*
- *understand and use a range of academic vocabulary as well as content or discipline-specific vocabulary in context;*
- *use specialised or complex grammatical structures, high lexical diversity, formal prestigious expressions, and abstract/technical concepts;*
- *interpret and adapt one's reading/writing for an analytical/argumentative purpose and/or in light of one's own experience;*
- *write in an authoritative manner, which involves the presence of an "I" addressing an imagined audience of specialists/novices or a variety of public audiences.*

The first two additions may indicate the need not so much for a new task type as for a new emphasis on comparing one text with another, which is already acknowledged as a component of the construct. In some versions of TALL, for

example, test takers are already expected to identify clearly different opinions in more than one text. Clearly, more such comparisons are necessary to test critical insight into points of agreement and disagreement, for example. Perhaps shorter texts with contrasting opinions might also be considered, but if properly tested it would add considerably to the length of a test. Otherwise, test designers might be required to ask more questions such as the following (similar to those in existing tests), which ask test takers to compare one part of a longer text with another:

The further explanation of exactly what the author means by using the term 'development' in the first paragraph we find most clearly in paragraphs

- A. 2 & 3.
- B. 3 & 4.
- C. 5 & 7.
- D. 6 & 8.

or

The author discusses two divergent opinions about tapping into wind power. These opposite views are best expressed in paragraphs

- A. 2 & 3.
- B. 3 & 4.
- C. 5 & 7.
- D. 6 & 8.

We return below to the third addition: the ability to synthesize and integrate information, when we discuss another way of handling an additional writing task, based on the work of Pot (2013). A skill that, even at entry level to the academic world, relates to the notion of avoiding plagiarism and maintaining academic integrity, namely the ability to refer accurately to a multiplicity of sources, can at that lower level perhaps be measured in a task type such as the following:

### References

Imagine that you have gone to the library to search for information in the form of books, articles and other material, on the topic of "Making effective presentations". You have found a number of possible sources, and have made notes from all of them for use in your assignment on this topic, but have not had the time to arrange them in proper alphabetical and chronological sequence.

Look at your notes below, then place the entry for each source in the correct order, as for a bibliography, by answering the questions below:

- (a) Jay, R. 2000. *How to write proposals and reports that get results*. London: Pitman.
- (b) Dickinson, S. *Effective presentation*. 1998. London: Orion Business.
- (c) Hager, P.J., H.J. Scheiber & N.C. Corbin. 1997. *Designing and delivering scientific, technical, and managerial presentations*. New York: Wiley-Interscience.
- (d) Chemical and Process Engineering, University of Newcastle-upon-Tyne. 2001. Presentation skills. Available <http://lorien.ncl.ac.uk/ming/Dept/Tips/present/present.htm>.
- (e) Jay, R. & A. Jay. 1994. *Effective presentation: powerful ways to make your presentations more effective*. Prentice-Hall: London.

- The entry I placed **first** is      (a)   (b)   (c)   (d)   (e)
- The entry I placed **second** is    (a)   (b)   (c)   (d)   (e)
- The entry I placed **third** is      (a)   (b)   (c)   (d)   (e)
- The entry I placed **fourth** is    (a)   (b)   (c)   (d)   (e)
- The entry I placed **fifth** is      (a)   (b)   (c)   (d)   (e)

The entry with the **date** of publication in the wrong place is (b) (c) (d) (e)

The entry that has the **place** of publication in the wrong place is

- (a) Jay (2000)
- (b) Hager *et al.* (1997)
- (c) Chemical and Process ... (2001)
- (d) Jay & Jay (1994)

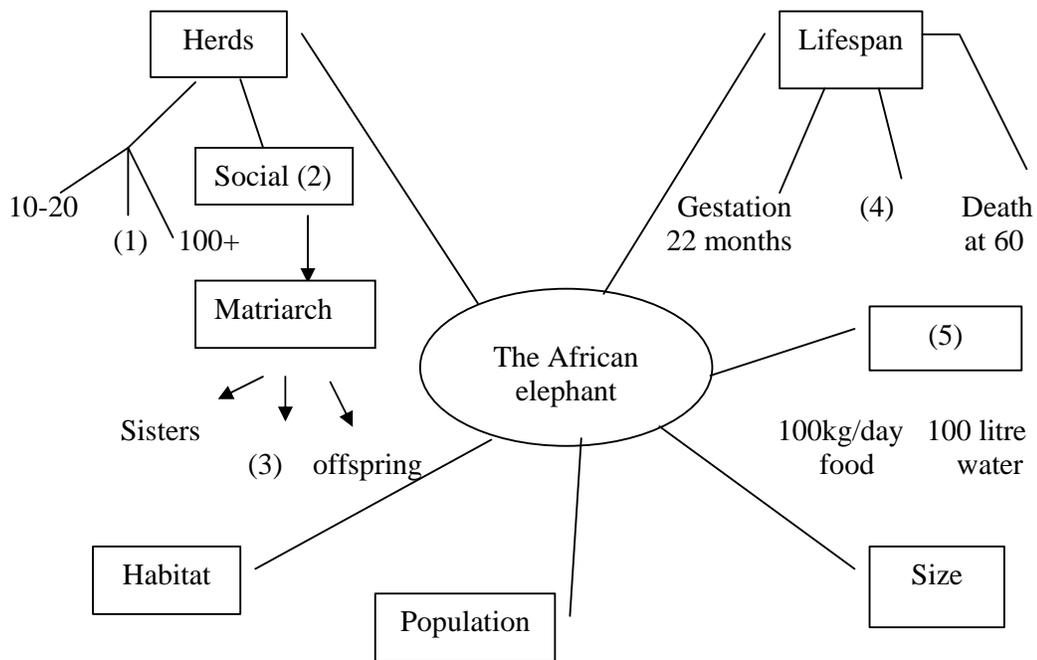
The proposed fourth addition identifies an often forgotten dimension of academic work: the creativity that accompanies the visualisation of logical distinctions and concepts. Though it might again potentially add to the time taken to complete this kind of test, at least there are (as yet untested) examples of what such tasks might look like: Weideman (2006) has several suggestions to this effect that were not followed up in actual test designs. Here is one, adapted from Weideman (2007:16-22). The text may be presented to test takers either in spoken format (as in a lecture) or in written format:

Listen to/read the following text, look at the diagram, and then answer the questions below:

### *The African elephant*

Elephants essentially live in herds and may be found in groups of anything between 10 and 20 or up to 50 and more, and, in rare cases, in excess of 100. Their highly developed social structure, however, remains consistent throughout. Family units are led by a cow elephant, or matriarch, and a typical family herd consists of cow elephants of various ages: the leader, and her sisters, their daughters, and their offspring.

The lifespan of an elephant is long and often eventful. For one thing, elephants...



(1) The most appropriate choice here is

- (a) between 30 and 40
- (b) more than 40
- (c) more than 50
- (d) about 70 or 80

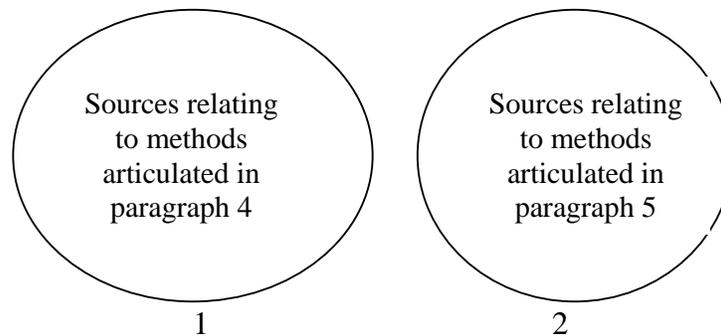
(2) The term that is used here is

- (a) development
- (b) structure
- (c) herd
- (d) family

- (3) The word that fits here is
- (a) elephants
  - (b) age groups
  - (c) children
  - (d) daughters

Another possible question type might be to suggest a visualisation of a distinction made by the author of a text, as follows:

In the fourth and fifth paragraphs, the author makes a distinction about the different sources of various methods to generate electricity. With reference to the rest of the text, identify the circle into which the terms 'nuclear', 'coal', 'solar', 'wind', and 'water' would best fit:



1. 'nuclear' would be in circle
  - A. 1.
  - B. 2.
  - C. neither 1 nor 2
2. 'solar' would fit into circle
  - A. 1.
  - B. 2.
  - C. neither 1 nor 2
3. 'coal' is best categorised as belonging in circle
  - A. 1.
  - B. 2.
  - C. neither 1 nor 2

The suggested additions under bullets five and six of the modified construct above indicate the need for a further differentiation of the tests for field or even discipline specific purposes. This is already happening in some cases: the range of tests designed by ICELDA now includes not only general tests of academic literacy, but also tests for students of disaster management, for nursing and for financial

planning. It follows that such tests must consider assessing the ability to use discipline specific terminology. This is related to the testing of the ability to use complex grammatical structures, prestigious expressions specific to a field, as well as abstract concepts and ideas. The current testing of grammar and text relations in one subtest of TALL and TALPS provides a possible format for such questions, but it should perhaps be adapted to reflect field-specific lexical and phrasal content (Weideman & Van Dyk 2014: 95):

**In the following, you have to indicate the possible *place* where a word may have been deleted, and which *word* belongs there. Here are two examples:**

Charles Goodyear (1800–1860) invented the vulcanization of rubber when he was experimenting by heating a mixture of rubber and sulphur. The Goodyear story is one of either pure luck or careful research, but both are debatable. Goodyear insisted that it was [i] the [ii], though [iii] many [iv] contemporaneous [i] accounts [ii] indicate [iii] the [iv].

**Where has the word been deleted?**

- A. At position (i).
- B. At position (ii).**
- C. At position (iii).
- D. At position (iv).

**Which word has been left out here?**

- A. indeed
- B. very
- C. former**
- D. historically

**Where has the word been deleted?**

- A. At position (i).
- B. At position (ii).
- C. At position (iii).
- D. At position (iv).**

**Which word has been left out here?**

- A. historical
- B. latter**
- C. now
- D. incontrovertibly

The last two additions proposed above by Patterson and Weideman (2013b), namely the adaptation of one's reading or writing for the purposes of an academic argument, and the authoritative manner in which it should be delivered, may be less relevant for post-entry assessments at undergraduate level. At higher levels they may profitably be combined, however, so that both 'authority' and audience difference are allowed to come into play. We therefore turn next to a further consideration of how these additions might be assessed.

## Writing with authority

The last two additions proposed to the construct concern not only reading (finding information and evidence for the academic argument) but also writing and more specifically, writing persuasively (Hyland 2011: 177) and with authority either to a specialist or lay audience. As in the case of some of the other proposals, these additions appear to be more relevant for the discourse expected from seasoned academics than from entry-level beginners, who are normally the prime targets of post-entry language assessments. In the case of TALPS, Pot (2013: 53) found that “the greatest challenge for students is to present a coherent, well-structured academic argument, and to do so by making use of the appropriate communicative functions used in academic discourse. In essence, students fail to grasp the main concept of presenting an academic argument in written form.” Initially, all versions of TALL did include a writing component, but the resources required to mark it reliably, as well as the high correlation between its results and that of the rest of the test resulted in a decision to exclude it altogether. Since some wrongly equate academic literacy with the ability to write (Weideman 2013; Butler 2013), this omission would in their view constitute a potential loss of face validity for such tests (Butler 2007: 297).

A useful and potentially productive proposal of how to deal with adding a writing component to the current tests can, however, be found in Pot (2013). What she suggests in the case of post-graduate tests of literacy can both equally and profitably be applied to undergraduate tests as well, and may be less resource-intensive than the addition of a full-blown writing task or set of tasks to a single existing test. It is instructive to note where this author’s engagement with the proposals she makes derives from. As post-entry tests like TALL, TAG, and TALPS began to be widely used, over time the demand from course designers to derive diagnostic information from them also grew. So the initial goal of her

research was to untangle, from the mass of information yielded by the test results, that which could assist subsequent course design. The identification of the benefits to be gained from unlocking the diagnostic information of TALPS will be the focus of another report (Pot & Weideman 2014), but here we shall focus only on a number of proposals she makes in the conclusion of her investigation of the diagnostic information to be gleaned from TALPS - proposals that might enhance the design of all the post-entry tests discussed in this paper.

Building on design ideas already used, for example, in post-entry tests developed in Australia and New Zealand, Pot's (2013: 54ff.) proposal is that the designers consider the introduction of a two-tier test. This would mean splitting the test in two, first testing all candidates sitting for the TALPS on the first seven subsections of the test, all of which are in multiple choice format: Scrambled text, Interpreting graphs and visual information, Academic vocabulary, Text types, Understanding texts, Grammar and text relations, and Text editing. Subsequently, should candidates have scored below the cut-off point (currently at 60%) for this first test, which is an indication of risk, or if they have scored low on the two subtests (sections 6 and 7, Grammar and text relations and Text editing) that have in the past shown very high correlations with the writing section (initially the eighth section of the test), or if they are borderline cases identified through empirical analyses of potential misclassifications related to the reliability of the test (Van der Slik and Weideman 2005: 28), they are given a second opportunity to have their ability to handle academic discourse assessed.

In this second-chance test, they are given a writing assessment that is at least twice as long, i.e. between 500 and 800 words, than the current 300-word format, and perhaps with further texts on the theme of the argument they are expected to write, in addition to the texts that formed part of the first test. The advantages are clear. As has been indicated above, the essential feature of analytical, academic discourse is distinction-making and the ability to present that coherently in an academic

argument. A longer written task would allow them to demonstrate an adequate ability to structure a comprehensive argument, as well as the ability to acknowledge sources in a conventionally acceptable manner. It would also allow them time to plan properly, a feature that Pot (2013) found may have been missing from the current, 90-minute test.

This particular refinement to the format of TALPS (and potentially the other tests as well) is mentioned here because the motivation for proposing it derives directly from a consideration of the construct of the test, and what it means subsequently for responsible course design:

Because distinction-making is at the heart of academic language ability and this study has demonstrated a lack of mainly structural distinction-making in the students' essays, courses can focus on distinction-making as a central theme (Pot 2013: 58).

### **Test refinement and impact**

From a design angle, the examination of the construct underlying post-entry tests of academic literacy in South Africa is potentially highly productive. In addition, tapping the diagnostic information the tests yield more efficiently, as well as making modifications and additions to current test task types, will provide theoretically defensible changes to their design.

The possible additions to the design of the tests referred to in this article will benefit not only the current set of assessments, but are likely to have a beneficial effect on the design of similar tests, in more languages than the current English and Afrikaans versions of the instruments. Butler and his associates at North-West University have, for example, already begun to experiment with translated versions of these post-entry assessments into Sesotho, a language widely used as first language by large numbers of students on some of their campuses, but that remains underdeveloped as an academic language (Butler 2014). A greater range of test task types will enhance the potential of the tests to provide results that are useful and

interpretable (Weideman 2013), all of which may have further beneficial effects in informing policy decisions at institutions of higher learning in South Africa about an expansion of the current two languages of instruction to three at least at some already multilingual universities.

Tests are therefore never neutral instruments, and neither is their refinement. In examining components of their design critically, and discussing modifications to them on the basis of such an examination, our goal is to continue to enhance their worth and impact. The one respect – the re-articulation of the construct - in which possible changes might be made, and that was discussed here, therefore also needs to be augmented in future research by other considerations and design principles. If Read (2010: 292) is correct in observing that the “process of test development... does not really count as a research activity in itself”, we have little issue with that. Test development processes, however, are never fully complete. Once constructed, they are subject to redesign and refinement. Since the creativity and inventiveness of test designers take precedence over the theoretical justifications of our designs, it would be a pity if the history of the rather agonising process of how to best assess a given construct were not recorded, for if that goes unrecorded, we miss the opportunity of sharing with others a potentially productive design ingredient for making or re-making tests. In more closely examining that which initially may have been secondary, namely the theoretical defence of our design, designers are, once they again scrutinise that theoretical basis, stimulated to bring their imaginations to bear on the redesign and refinement of their assessment instruments. There is a reciprocal relationship between the leading design function of an assessment measure and its foundational analytical base (Weideman 2014). This paper has therefore aimed to provide such a record of test design, and redesign, which has been prompted by a reconsideration of the theoretical definition of what gets tested – that is, the construct.

## References

- Bachman, L.F. & A.S. Palmer. 1996. *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Bailey, A.L. (Ed.). 2007. *The language demands of school: Putting academic English to the test*. New Haven: Yale University Press.
- Beekman, L., Dube, C. & Underhill, J. 2011. *Academic literacy*. Cape Town: Juta.
- Biber, D. & Conrad, S. 2001. Register variation: A corpus approach. In: Schiffrin, D., Tannen, D., & Hamilton, H.E. (Eds.). *The handbook of discourse analysis*. Malden: Blackwell Publishing. Pp. 175-196.
- Blanton, L.L. 1994. Discourse, artefacts and the Ozarks: understanding academic literacy. *Journal of second language writing* 3 (1): 1-16. Reprinted (as Chapter 17: 219-235) in V. Zamel & R. Spack (Eds.), 1998. *Negotiating academic literacies: teaching and learning across languages and cultures*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Butler, G. 2007. A framework for course design in academic writing for tertiary education. Unpublished doctoral dissertation. Pretoria: Unit for academic literacy, University of Pretoria.
- Butler, G. 2013. Discipline-specific versus generic academic literacy intervention for university education: An issue of impact? *Journal for language teaching*. 47(2): 71-88. Available DOI: <http://dx.doi.org/10.4314/jlt.v47i2.4>.
- Butler, G. 2014. Translating the Test of Academic Literacy Levels (TALL) into Sesotho. Forthcoming in *Southern African linguistics and applied language studies*.
- Chapelle, C.A. 2012. Conceptions of validity. In: Fulcher, G. & Davidson, F. (Eds.). *The Routledge handbook of language testing*. Abingdon: Routledge. Pp. 21-33.
- Cliff, A.F. & M. Hanslo. 2005. The use of 'alternate' assessments as contributors to processes for selecting applicants to health sciences' faculties. Paper read at the Europe Conference for Medical Education, Amsterdam.
- Cliff, A.F., N. Yeld & M. Hanslo. 2006. Assessing the academic literacy skills of entry-level students, using the Placement Test in English for Educational Purposes (PTEEP). Mimeographed MS.
- Davidson, F. & Lynch, B.K. 2002. *Testcraft*. New Haven: Yale University Press.
- Department of Basic Education, 2011a, *Curriculum and assessment policy statement (CAPS) for English Home Language, Further Education and Training phase, grades 10-12*. Pretoria: Department of Basic Education.
- Department of Basic Education, 2011b, *Curriculum and assessment policy statement (CAPS) for English First Additional Language, Further Education and Training phase, grades 10-12*. Pretoria: Department of Basic Education.

- Du Plessis, C., Steyn, S. & Weideman, A. 2013. Towards a construct for assessing high level language ability in grade 12. Report to the Umalusi Research Forum, 13 March. Bloemfontein: ICELDA.
- Flower, L. 1990. Negotiating academic discourse. In: Flower, L., Stein, V., Ackerman, J., Kantz, M., McCormick, K., & Peck, W.C. (Eds.) *Reading-to-write: Exploring a cognitive and social process*. Oxford: Oxford University Press. Pp. 221-252.
- Foley, J.A. (Ed.). 2004. *Language, education and discourse: Functional approaches*. London: Continuum.
- Gee, J.P. 1998. What is literacy? Chapter 5, pp. 51-59 in Zamel, V. & Spack, R. (Eds.), 1998: Reprint of 1987 article in *Teaching and learning: the journal of natural enquiry* 2: 3-11.
- Green, A. 2014. *Exploring language assessment and testing: Language in action*. London: Routledge.
- Gunnarsson, B. 2009. *Professional discourse*. London: Continuum.
- Habermas, J. 1970. Toward a theory of communicative competence. In: Dreitzel, H.P. (Ed.). *Recent sociology* 2. London: Collier-Macmillan. Pp. 41-58.
- Halliday, M.A.K. 1978. *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.
- Halliday, M.A.K. 2002. *Linguistic studies of text and discourse*. Webster, J.J. (Ed.). London: Continuum.
- Halliday, M.A.K. 2003. *On language and linguistics*. Webster, J.J. (Ed.). London: Continuum.
- Hartnett, C.G. 2004. What should we teach about the paradoxes of English nominalization? In: Foley, J.A. (Ed.). *Language, education and discourse: Functional approaches*. London: Continuum. Pp. 174-190.
- Hasan, R. 2004. Analysing discursive variation. In: Young, L & Harrison, C. (Eds.) *Systemic functional linguistics and critical discourse analysis: Studies in social change*. London: Continuum. Pp. 15-52.
- Hyland, K. 2011. Academic discourse. In: Hyland, K. & Paltridge, B. (Eds.). *Continuum companion to discourse analysis*. London: Continuum. Pp. 171-184.
- Hyland, K. & Bondi, M. (Eds.) 2006. *Academic discourse across disciplines*. Bern: Peter Lang.
- Hymes, D. 1971. On communicative competence. In: Pride, J.B. & Holmes, J. (Eds.). 1972. *Sociolinguistics: selected readings*. Harmondsworth: Penguin. Pp. 269-293.
- ICELDA (Inter-Institutional Centre for Language Development and Assessment). 2014. [Online]. Available <http://icelda.sun.ac.za/>. Accessed 7 February 2014.
- Knoch, U. & Elder, C. 2013. A framework for validating post-entry language assessments (PELAs). *Papers in language testing and assessment* 2(2): 48-66.
- Livnat, Z. 2012. *Dialogue, science and academic writing*. Amsterdam: John Benjamins.

- McNamara, T. & Roever, C. 2006. *Language testing: The social dimension*. Oxford: Blackwell.
- Patterson, R. & Weideman, A. 2013a. The typicality of academic discourse and its relevance for constructs of academic literacy. *Journal for Language Teaching* 47(1): 107-123. Available DOI: <http://dx.doi.org/10.4314/jlt.v47i1.5>.
- Patterson, R. & Weideman, A. 2013b. The refinement of a construct for tests of academic literacy. *Journal for language teaching* 47(1): 125-151. Available DOI: <http://dx.doi.org/10.4314/jlt.v47i1.6>.
- Pot, A. 2013. Diagnosing academic language ability: An analysis of TALPS. Unpublished master's dissertation. Groningen: Rijksuniversiteit Groningen.
- Pot, A. & Weideman, A. 2014. Diagnosing academic language ability: Insights from an analysis of a postgraduate test of academic literacy. Submitted to *Language matters*.
- Read, J. 2010. Researching language testing and assessment. In: Paltridge, B. & Phakiti, A. (Eds.). *Continuum companion to research methods in applied linguistics*. London: Continuum. Pp. 286-300.
- Snow, C.E. & Uccelli, P. 2009. The challenge of academic language. In: Olson, D.R. & Torrance, N. (Eds.). *The Cambridge handbook of literacy*. Cambridge: Cambridge University Press. Pp. 112-133.
- Strauss. D.F.M. 2009. *Philosophy: Discipline of the disciplines*. Grand Rapids, MI: Paideia Press.
- Suomela-Salmi, E. & Dervin, F. (Eds.) 2009. *Cross-linguistic and cross-cultural perspectives on academic discourse*. Amsterdam: John Benjamins.
- Van der Slik, F. & Weideman, A. 2005. The refinement of a test of academic literacy. *Per Linguam* 21(1): 23-35.
- Van Dyk, T. & Weideman, A. 2004a. Switching constructs: On the selection of an appropriate blueprint for academic literacy assessment. *Journal for language teaching* 38(1): 1-13.
- Van Dyk, T. & Weideman, A. 2004b. Finding the right measure: From blueprint to specification to item type. *Journal for Language Teaching* 38(1): 15-24.
- Weideman, A. 2003. Assessing and developing academic literacy. *Per linguam* 19(1 & 2): 55-65.
- Weideman, A. 2006. Assessing academic literacy in a task-based approach. *Language matters* 37(1):81-101.
- Weideman, A. 2007. *Academic literacy: Prepare to learn*. (Second edition). Pretoria: Van Schaik Publishers.
- Weideman, A. 2009. *Beyond expression: A systematic study of the foundations of linguistics*. Grand Rapids: Paideia Press.
- Weideman, A. 2011. Academic literacy tests: Design, development, piloting and refinement. *SAALT Journal for Language Teaching* 45(2): 100-113.

- 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. Available DOI: <http://dx.doi.org/10.4314/jlt.v47i2.1>.
- Weideman, A. 2014. Innovation and reciprocity in applied linguistics. Forthcoming in *Literator* 35(1). Available DOI: <http://dx.doi.org/10.4102/lit.v35i1.1074>.
- Weideman, A. & Van Dyk, T. (eds.). 2014. *Academic literacy: Test your competence*. MS. Forthcoming from ICELDA.
- Yeld, N. *et alii*. 2000. The construct of the academic literacy test (PTEEP). Mimeograph. Cape Town: Alternative Admissions Research Project, University of Cape Town.