

Diagnosing academic language ability: An analysis of TALPS

Anna Pot
S1912518

MA in Applied Linguistics
Faculty of Liberal Arts
University of Groningen

Supervisor: Prof. A.J. Weideman
Second reader: Dr. H.I. Hacquebord

8 August 2013
16105words

Abstract

The observation that a large number of postgraduate students do not possess the right level of academic language ability to complete their studies successfully has informed this study in undertaking an investigation into postgraduate students' strengths and weaknesses regarding academic literacy, with a specific interest in academic writing. By performing a diagnostic analysis of the results of the Test of Academic Literacy for Postgraduate Students (TALPS) this study identifies specific areas of academic literacy with which students struggle. A quantitative analysis of the multiple-choice result section, as well as a qualitative analysis of the essay question of the test, marked with a specifically designed marking rubric, indicates that students primarily struggle with structuring an argument. The essays often include neither an introduction nor conclusion, lack internal consistency and do not make use of the various communicative functions when expressing an academic argument. These findings can inform not only the future design of TALPS, but also aid course designers, students, and teachers with regard to the focus of courses in academic language competence, as well as the learning and teaching associated with its development.

Keywords: academic literacy; academic language ability; diagnostic analysis; TALPS; academic writing; postgraduate students

Acknowledgements

I would like to express my sincerest gratitude to Prof. Albert Weideman of the University of the Free State for his valuable supervision during the thesis process. His detailed feedback on the many drafts I handed in, our insightful weekly discussions and his enthusiastic encouragement stimulated new ideas and deeper and further thinking and analysis. I am also greatly indebted to him for the hospitable and kind way in which I was welcomed and integrated in Bloemfontein and at ICELDA, through which I felt quickly at ease and enormously enjoyed my four-month stay.

I would also like to express my thanks to Dr. Hilde Hacquebord for allowing me to write this thesis on a topic associated with my internship at ICELDA. Her encouragement and lenient deadline dates are what allowed me to complete this thesis before the start of a new academic year. Thanks to the University of Groningen and the MA Applied Linguistics in particular for affording me the opportunity to fulfil my internship duties (which preceded this thesis) abroad.

Contents

Abstract	2
Acknowledgements	3
Contents.....	4
List of tables and figures	6
Chapter 1 Introduction.....	7
1.1 Context	7
1.2 The role of language testing	8
1.3 An interactive perspective on writing	9
1.4 The focus of this study	9
Chapter 2 Key concepts in the assessment of academic literacy	11
2.1 The critical feature of academic discourse	11
2.2 Language ability as a skill	13
2.3 TALPS	14
2.4 Towards a construct of academic language ability	16
2.5 Writing.....	19
2.6 Diagnostic testing	21
2.7 Conclusion	24
Chapter 3 Rationale	25
3.1 An earlier analysis of TALPS	25
3.2 Hypotheses	26
Chapter 4 Methodology	28
4.1 Overview	28
4.2 Subjects.....	28
4.3 Materials and procedures.....	29
4.3.1 Marking rubric 1	30
4.3.2 Marking rubric 2	32
4.4 Statistical analysis	37

4.5 Design.....	37
Chapter 5 Analysis of the results	40
5.1 Quantitative analysis	40
5.2 Qualitative analysis	45
5.2.1 Text relations	45
5.2.2 Distinguishing.....	48
5.2.3 Communicative function	49
5.2.4 Extrapolating and inferring.....	50
5.2.5 Vocabulary.....	50
5.3 Answering the research question.....	51
Chapter 6 Discussion of possible implications.....	53
6.1 Revisiting the results	53
6.2 Design suggestions	54
6.2.1 A two-tiered test model	54
6.2.2 Other design options	56
Chapter 7 Conclusion	59
7.1 Purpose of the study	59
7.2 Limitations.....	59
7.3 Final word.....	60
Appendix A	61
List of transitional words and phrases	61
Appendix B	62
Marking rubric 1	62
Appendix C	63
Marking rubric 2	63
Appendix D	64
Marking rubric 3	64
References	65

List of tables and figures

Table 2.1	TALPS sections and subtests (adapted from Patterson 2012).	16
Table 2.2	Subtests associated with components of construct of academic literacy.	18
Table 4.1	Components of marking rubric 1 of TALPS.	30
Table 4.2	Marking scale of each component of the marking rubric 1.	30
Table 4.3	Indicators of level of each marking component of marking rubric 1.	31
Table 4.4	Components of marking rubric 2.	33
Table 4.5	Importance of each component of marking rubric 2.	34
Table 4.6	Indicator(s) of each component of marking rubric 2.	36
Table 4.7	Outline of the correlation matrix.	38
Table 5.1	Correlation matrix of the three different types of correlation.	41
Table 5.2	Weighted total scores per component of marking rubric 2.	44
Figure 5.1	Correlation pattern of the recalculated results of section 1 with section 8.	42
Figure 5.2	Correlation pattern of the recalculated results of section 6 with section 8.	43

Chapter 1

Introduction

1.1 Context

Taking a social perspective on language testing, this study seeks to explore South African postgraduate students' strengths and weaknesses in academic language ability. As the majority of the students enrolling in tertiary education in South Africa are non-native users of English, an increasing number of them do not possess the required proficiency level to complete their studies successfully (Van Dyk and Weideman 2004a). Language thus constitutes a crucial aspect of achieving success in (higher) education.

Due to the political changes in South Africa, the educational aim is to increase access to tertiary education, especially for groups that were previously at a disadvantage. In addition, an increasing number of foreign students from other African countries take up their studies in South Africa. These students come from diverse language backgrounds and as a result they often do not possess the required proficiency level in English to complete their studies (Butler 2007). Also at postgraduate level L1 English and Afrikaans speakers are outnumbered by L2 English users (Butler 2009). A survey conducted by Butler (2009: 13-14) on perceived academic literacy among students and supervisors concluded that 20% of the participating postgraduate students have never received any formal education in English, and 33% and 44% respectively did not use English as a language of learning for their first or honours degrees.

As has been indicated by i.a. Weideman (2003), Van Dyk and Weideman (2004a) and Weideman (2013), academic language proficiency is linked to academic performance, since "low academic language proficiency levels have indeed been mooted as one of the primary causes of the lack of academic success experienced by many students at South African universities" (Van Dyk and Weideman 2004a: 1). Moreover, universities in South Africa all share the notion that a lack of academic language proficiency puts not only students at risk; parents,

the universities themselves and the higher education system as a whole are also affected by the academic language proficiency level of students.

1.2 The role of language testing

Therefore, assessing academic literacy is necessary to raise awareness among those students who are at risk. This facilitates taking early precautions to enable them to reach a level of academic literacy that may contribute to lowering their risk of failing courses, their parents from paying for extra college years, and universities from losing subsidies (Weideman 2003).

Nonetheless, tests neither always overtly indicate, nor are they concerned with the areas in which students seem to struggle the most. Although the score of tests of academic language ability may indicate whether students need to follow extra academic literacy courses, the score does not indicate with which elements a student particularly struggles. It would be beneficial, not only for students but also for curriculum designers, teachers, and test developers, to gain insight into the particular areas with which students encounter difficulties. This can inform and direct students, lecturers, and course and test designers in their subsequent learning, teaching, and designing of academic literacy courses and tests.

A diagnostic analysis of the results of a test of academic literacy can shed light on the difficulties students encounter in their handling of academic discourse. This study will conduct such an analysis by paying particular attention to academic writing abilities. Before doing so, however, it is vital to understand what academic language ability entails, as

language testing involves not only the psychometric and technical skills required to construct and analyse a test but also knowledge about language (Alderson and Banerjee 2002: 80).

Language is always embedded in context and the context defines the typical use of language (Weideman 2009: 39). It follows that in an academic environment, one makes use of the typical features of the language associated with academic discourse. A critical question to ask, therefore, is: What sets academic discourse

apart from other types of discourse? In chapter 2 I will explore the literature on academic discourse and list several features associated with it. Subsequently, it will be possible to consider the construct of academic literacy developed for use in assessing this specific language ability in light of the notion of what constitutes academic discourse.

1.3 An interactive perspective on writing

Neither academic literacy nor its assessment can be regarded as merely involving or measuring several discrete ‘skills’. Rather, both the assessment and the construct that underlies it should be approached from the angle of the communicative functions of language, in which all these skills are intertwined. These observations relate to the main interest of this study, which lies in conducting a diagnostic analysis of the results of the writing section of the Test of Academic Literacy for Postgraduate Students (TALPS). Since the conventional assumption is that being able to write well is of substantial importance in achieving academic success, this section of the test requires students to produce a written argument.

Although at first sight it seems that the writing section merely tests the skill ‘writing’, a closer look confirms that ‘writing’ consists of multiple components not only of functional language use, but is also usually complemented by reading skills, listening skills, or both, as well as by interaction and communication in various media (face-to-face interaction, remote or non-simultaneous communication, and so forth). Therefore, ‘writing’ as an isolated skill is an illusion, as the analysis in this study will demonstrate.

1.4 The focus of this study

TALPS is mainly used to assess the level of academic literacy of postgraduate students in order to determine whether they are “at risk” and need to follow additional courses to improve their academic literacy abilities. However, as has been indicated by Patterson (2012), TALPS can also serve diagnostic purposes by

identifying those areas of academic literacy with which students are generally struggling. Even though testing for diagnostic purposes is a good way to indicate strengths and weaknesses in students' (language) performance, it is necessary to consider carefully the use and goal of a diagnostic test, since pure diagnostic tests are rare and often other test types are – perhaps inappropriately – adapted to serve diagnostic purposes. Therefore, chapter 2 will also consider the (limited) literature on diagnostic testing to explore what information a diagnostic analysis of TALPS may yield.

After having addressed the various concepts related to academic language ability in chapter 2, chapter 3 will provide the rationale for this study, in which the problem is clearly outlined, an earlier diagnostic analysis of TALPS is discussed and various hypotheses are developed in order to be able to answer subsequently the research question of how much more the completion of the writing section of TALPS may inform us in terms of the diagnosis of shortcomings in academic literacy, also when I compare its results to those of the remainder of the test.

To test these hypotheses, I will perform several correlations with the 2012 and 2013 results of TALPS as well as conduct an in-depth analysis of the students' essays of 2013, based on an assessment rubric specifically designed for this study. The design and details of how these analyses are carried out are outlined in chapter 4. Chapter 5, subsequently, provides an analysis of the results. The quantitative, as well as the qualitative findings are discussed in this section and the hypotheses are accepted or rejected, after which an answer to the research question is provided. The possible implications these findings may have on test and course design are elaborated upon in the discussion of chapter 6.

Chapter 2

Key concepts in the assessment of academic literacy

2.1 The critical feature of academic discourse

As briefly touched upon in section 1.2 above, we cannot do justice to an idea of academic discourse by merely employing traditional classifications such as informal/formal and slang/jargon. These distinctions are first of all too broad (i.e. an academic article and a job application are both formal) and with this distinction one may create, secondly, a hierarchy which values one type of discourse above another (Patterson and Weideman 2013a: 10). In an article on academic language, Snow (2010: 452) identifies conciseness as a key feature of academic discourse; it is achieved, according to her, by using a “high density of information-bearing words, ensuring precision of expression; and relying on grammatical processes to compress complex ideas into few words” (2010: 450). Therefore, she lists “grammatical embeddings, sophisticated and abstract vocabulary, precision of word choice, and use of nominalizations to refer to complex processes” as being key features of academic language. However, Patterson and Weideman (2013a) argue that nominalizations are not restricted to academic discourse, a remark which impacts in turn the definition of academic discourse as relying on conciseness as such.

This one attempt to define academic language ability illustrates the difficulty of describing the unique feature(s) that sets academic discourse apart from other types of discourse. Patterson and Weideman (2013a) provide an alternative to the literature on definitions of academic discourse by investigating academic literacy through the notion of material lingual spheres. As a starting point, the theory of material lingual spheres assumes that “pure language” – without a context – does not exist (Weideman 2009: 39). However, the context defines the typical language use. This entails that one cannot rely on formal lingual notions alone, but that one also has to take both the conditions for language and the functional, concrete use of language into account. This communicative command of language allows for

typically distinct lingual discourses or language types, defined by the requirements of the specific context. These differences can broadly be determined by considering the presence of (1) an authority relationship, and (2) the durability of the relationship between participants in the discourse. Communicative relationships may range from an institutional relationship which has both authority and durability, to an associational relationship which has neither authority nor durability (Weideman 2011: 65). These differences affect the content of the language that is produced within a particular kind of relational interaction, hence the notion of these lingual encounters belonging to lingual spheres that are materially different. These spheres are not only materially, but also typically different, in terms of the typically distinct feature that guides and stamps each.

After examining different definitions of academic discourse in which Patterson and Weideman (2013a) sought for the defining feature of this ‘sphere’, one may arrive at a preliminary definition of academic discourse, to articulate what sets it apart from other types of discourse:

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 (2013a: 13).

With this definition, Patterson and Weideman present a provisional attempt to understand the typicality of academic discourse. Although distinction-making is not uniquely associated with academic discourse, the definition can exclude other types of discourse that also rely on distinction-making since in these types of discourse distinction-making is not a characterising feature, whereas in academic discourse it is. This view is echoed by Weideman and Van Dyk (2013: 4) who argue that “analysis is the core of academic argumentation and discourse” and that academic discourse is thus defined by its ability to make distinctions. Similarly, Butler (2007: 29) highlights critical thinking as the “most rewarded” component of academic literacy, exemplified through the ability to make distinctions.

2.2 Language ability as a skill

The notion of academic discourse does not assume that language ability is composed of and can be measured through several distinct skills. As I have observed in section 1.3, language ability was previously regarded to be a combination of several discrete skills: reading, speaking, listening, and writing. More recently, it has been argued that we should view language as being communicative, “intended to mediate and negotiate human interaction” (Weideman 2003: 4). Moreover, Bachman and Palmer (1996: 75f) argue that language use needs to be thought of in terms of “specific activities or tasks in which language is used purposefully”. In fact, Bachman and Palmer (1996: 75f) note that we should “not consider language skills to be part of language ability at all, but to be the contextualised realization of the ability to use language in the performance of specific language tasks” (1996: 75f). Therefore, dividing language learning into several skills is unnecessary, since in the real world language serves as a tool for communicating, in which all these skills are intertwined.

Even at an instrumental level, in which the skills are merely used to identify the means with which language is transferred, reading, speaking, listening and writing cannot be separated and conceptualised as different abilities. Surely, in a strict sense, writing and speaking, as well as reading and listening, are different forms of communication. However, in writing one must also be reading or have read, and in speaking one must also be listening. Once we view language also as communication, we have to acknowledge that it is embedded in interaction. Even more, when language is assessed through e.g. a listening test, students do not only perceive the language with their ears, they also have to read the questions on the answer sheet, perhaps take notes of what is being said, and produce an answer which requires critical thinking and elimination of possibilities (especially in case of a multiple-choice format). In fact, as argued by Weideman (2013: 3), even though such a test may be called a listening test for convenience, it is not ‘listening’ that is being tested: finding, processing and producing information are among the components of language ability that are also being assessed by it. Therefore, we

cannot speak of skills as such, and especially not when we base our assessments on such an approach.

Along the same lines is the argument we find in the field of New Literacy studies (NLS), in which Street (1984) has been influential in pointing out the difference between an autonomous model of literacy and an ideological model. In much the same fashion as the skills-based approach to language learning, the autonomous model suggests that “literacy is a decontextualized skill, which once learnt can be transferred with ease from one context to another” (Lea 2008: 230). The perspective on literacy that is taken in NLS is based on the ideological model, which views writing and reading as “deeply social activities; familiarity with and understanding these practices takes place in specific social contexts, which are overlaid with ideological complexities” (Lea 2008: 230).

Another angle from which language use can be viewed, and which underlines the fact that a skills-approach has less to contribute to our understanding of language and language development, is the perspective of Dynamic Systems Theory (DST). This theory regards language as a complex system, consisting of several subsystems with continually interacting variables, fostering change over time (Verspoor and Behrens 2011). Considering language to consist of stand-alone units such as skills which, once mastered and ‘acquired’, can be universally applied independent of context, does not align with the notion of DST that all aspects of language are connected and, moreover, interact dynamically. In line with the theory of NLS, language develops depending on the context, and changes through use. It is this use which is a key feature in the development of language ability.

2.3 TALPS

In assessing academic language ability, a construct was articulated which has served as a foundation for the design of one of several academic literacy tests for first-year students entering university, the Test of Academic Literacy Levels (TALL), developed by the Unit for Academic Literacy (UAL; part of the University of Pretoria) in collaboration with North-West University and Stellenbosch University (Patterson, 2012). Its purpose is “not to exclude students

from study opportunities – they have already gained entry to the university by the time the test is taken – but to determine whether their level of academic literacy puts them at risk in their studies” (Van Dyk and Weideman 2004b: 1). If the test indicates that this is the case, students need to follow a set of courses to improve their academic literacy abilities.

However, as Patterson (2012: 4) concludes and as was observed in the survey conducted by Butler (2009), postgraduate students also struggle both with understanding academic texts, and with academic writing. Therefore a similar test was designed, aimed at postgraduate students: the Test of Academic Literacy for Postgraduate Students (TALPS), based on the same test construct as TALL. According to the website of ICELDA (the Inter-institutional Centre for Language Development and Assessment), both tests are primarily used to place students who are identified as being at risk on the appropriate academic literacy support courses or, more rarely, to determine access to higher education, for which the test result should not constitute more than 15% in the application procedure (ICELDA 2013; Hay 2010)

TALPS consists of eight sections, which each measures one or more components of the construct of academic literacy. The first seven sections of the test are in multiple-choice format because of the test population sizes and the pressure of efficiently issuing the results (Van Dyk and Weideman 2004b: 15). However, the eighth section tests postgraduate students’ ability in academic writing and text editing, since this is a crucial component of postgraduate study (Patterson 2012: 8). This section requires students to produce an academic argument of approximately 300 words, by referring to the texts in TALPS as sources, according to the Harvard method of referencing. Table 2.1 lists the various sections together with each subtest.

Section 1: Items 1-5	Scrambled text
Section 2: Items 6-15	Interpreting graphs and visual information
Section 3: Items 16-25	Academic vocabulary
Section 4: Items 26-30	Text types
Section 5: Items 31-51	Understanding texts
Section 6: 52-66	Grammar and text relations
Section 7: Items 67-76	Text editing
Section 8:	Academic writing

Table 2.1: TALPS sections and subtests (adapted from Patterson 2012).

Rambiritch (2012) provides a comprehensive overview of the concept of validity and validation in test design, and has investigated the validity and reliability of TALPS. She concluded that TALPS is indeed a valid and reliable test.

2.4 Towards a construct of academic language ability

Even though a definition of academic discourse is a foundation for composing a practical, employable construct for test-design, the developed construct that is used for the design of test items for TALPS has never been critically examined in this respect.

In their attempt to define academic language ability, Van Dyk and Weideman (2004a) have considered various descriptions and frameworks of academic literacy (for an overview see Van Dyk and Weideman 2004a; Patterson 2012), from which they developed the following construct of academic literacy, which serves as the theoretical foundation for the design of test items:

- 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.

(Van Dyk and Weideman 2004a: 10)

The practicality with regard to what students can do with academic texts in productive and receptive contexts is what makes this definition functional, as Butler (2009) argues. Indeed, the construct is skills-neutral in the sense that it does not divide academic literacy into distinct skills (see section 2.2), but exemplifies practical, functional language tasks which students need to master in different (academic) contexts.

Even though this construct has been put forward for validation at conferences and seminars (Van Dyk and Weideman 2004a: 11), it has never been critically examined in light of a definition of academic discourse. Therefore, we can relate the notion of distinction making as the defining feature of academic discourse as presented above to the earlier developed construct of academic literacy by considering table 2.2, and take an in-depth look at the subtests that may articulate one or more components of the construct.

	Component	Subtest
1	Understanding academic vocabulary	Vocabulary knowledge; Text comprehension; Grammar & text relations
2	Metaphor and idiom	Text comprehension; (and sometimes) Grammar & text relations
3	Understanding relations between different parts of a text	Scrambled text; (sometimes) Text comprehension and Grammar & text relations, Register & text type
4	Understanding text type (genre)	Register & text type; Interpreting graphs & visual information; Scrambled text; Text comprehension; Grammar & text relations
5	Interpreting graphic and visual information	Interpreting graphs & visual information; (sometimes) Text comprehension
6	Distinguishing between essential and non-essential information	Text comprehension; Interpreting graphs & visual information
7	Sequencing, ordering, and simple numerical computation	Interpreting graphs & visual information; Text comprehension
8	Finding evidence, making inferences, and extrapolating	Text comprehension; Verbal reasoning; Interpreting graphs & visual information
9	Understanding communicative function	Text comprehension; (sometimes) Grammar & text relations
10	Making meaning beyond the sentence	Text comprehension; Register & text type; Scrambled text; Interpreting graphs & visual information

Table 2.2: Subtests associated with components of construct of academic literacy (adapted from Patterson 2012).

Distinction-making indeed proves to be an important characteristic of academic literacy, as the sixth and seventh components both highlight the ability to distinguish between essential and non-essential information and the ability to order and sequence information. In addition, the eighth component highlights the ability to analyse: make inferences, extrapolate and find evidence. Also components three and nine include features of distinction-making. Moreover, the eighth component stresses the importance of critical thinking and arguing. Therefore, underlying the current construct there is already the conclusion that academic discourse is characterised by the ability to make distinctions, which in turn facilitates critical thinking.

2.5 Writing

As has been noted in the introduction, many postgraduate students in tertiary education struggle to write at the required academic level. In an academic environment and in postgraduate studies especially, academic writing is of considerable importance. As observed by Leki and Carson, “the ability to write well is necessary both to achieve academic success and to demonstrate that achievement” (1994: 83). Indeed, writing is usually the form in which students have to prove their ability to handle academic discourse. In this sense, writing can be regarded as a social act, occurring in social contexts (Krause 2001: 150).

Therefore, it does not suffice to consider writing in isolation, as a skill that can be taught by paying attention to formal lingual aspects only. In their research on the academic writing style in students’ persuasive essays, Uccelli, Dobbs and Scott (2013: 40) adopt a pragmatics-based framework in their analysis, composed of the following measurements:

- lexical diversity (how many different words are used in an essay);
- syntactic complexity (how complex are the clauses used);
- lexical density (how much information is packed into a single clause).

As these three items all measure complexity, the underlying notion is that academic literacy is manifested in terms of lexical and syntactic complexity. The third measurement which measures the density of information in a sentence echoes the idea of conciseness as an important feature of academic discourse (Snow 2010). However, as has been indicated above, writing consists of more than these formal aspects. Even though it might indeed be the case that an academic writing style can be measured through the use of dense, complex, and lexically diverse clauses, one should not overlook the communicative processes involved in academic writing. Moreover, writing is the medium through which distinction-making, the critical defining feature of academic discourse, is exemplified. An academic argument functionally embeds this kind of distinction-making, through incorporating different kinds of distinctions, ranging from observations, comparisons and contrasts, inferences and extrapolations, to conclusions. In other words, one should

not only consider the end-product, but the entire process of gathering, processing, and producing (new) information (Van Dyk and Weideman, 2013: 6).

The writing process therefore potentially involves many activities related to the ability to handle academic discourse. As Van Dyk and Weideman (2013: 6) note, gathering information is done either through reading or listening, making/reviewing notes or discussions with others; processing information is done through analysis, in which distinction-making (especially comparing, contrasting, and categorising) plays an important part; and only then do we produce new information (finally, in writing) in which we state our opinion, that has been formed in the previous processes, and is the articulation of the result of our distinction-making or analytical activity.

Writing is thus a complex task which combines several components of academic literacy, all centred around the unique feature of academic discourse: distinction-making. It is no doubt likely that those students who struggle in writing may have individual difficulties regarding certain aspects of much more than just the writing process itself. In writing courses, therefore, it is necessary to pay considerable attention to each student's individual difficulties by providing feedback. It has been demonstrated that quality feedback can aid the development of academic literacy, which in turn enhances academic success (Horstmanhof and Brownie 2013: 62). The importance of feedback is discussed further in section 2.5 below.

It seems unavoidable, thus, that in academic literacy assessment a writing section should not be excluded. As Butler (2009: 11) concludes from his survey, if a test of academic literacy does not include a 'writing' section, supervisors will regard this as a crucial shortcoming, which may affect the face validity of such a test. However, in assessment *of* an ability the primary purpose, and intention, may not be to identify those areas of writing in which students generally are lacking. These tests merely indicate a general lack of ability to handle academic literacy, which is subsequently acted upon by, for example, placing "at risk" students on courses to improve their academic literacy level.

Nonetheless, when analysing the results of such a test from an “assessment *for learning*” point of view, in other words by performing a diagnostic analysis, one may be able to identify particular areas in writing ability with which a student struggles. This can in turn provide more insight into the general difficulties of academic writing ability and what precedes it, and can inform the design of subsequent literacy courses.

2.6 Diagnostic testing

As noted above, there are two types of assessment; assessment *of learning* and assessment *for learning*. Whereas the former is a summative type of assessment, measuring what a student has learned or can do at a particular, pre-established level, assessment *for learning* is a formative type of assessment and provides an indication of the learning process, by focusing on enhancing students’ abilities to reach a certain level. This latter type of *diagnostic* assessment greatly enhances students’ achievement, as the focus of diagnostic testing is first on identifying students’ strengths and weaknesses and then on suggesting improvement (Zao 2013: 43). Providing feedback is therefore a unique and essential feature of diagnostic assessment.

However, feedback has to be useful in order to be effective. This usefulness may vary with regard to focus on either weaknesses, strengths, or both, and its purpose, being evaluative or descriptive in nature (Kunnan and Jang 2009: 618). In addition, Zao (2013: 43) notes that feedback enhances learner autonomy, as students can set specific learning goals. Moreover, teachers are better able to adapt their teaching to those aspects students struggle with, and feedback can yield valuable information concerning language learning and teaching objectives, which can facilitate the work of test developers and curriculum designers in designing more effective tests or courses.

At any rate, Kunnan and Jang (2009) indicate that feedback needs to consider “learners’ beliefs about learning goals, about their own ability, and cognitive and metacognitive learning styles” (2009: 618). Therefore, effective

diagnostic feedback enables students to focus on achieving their learning goals by providing constructive guidelines for a step-by-step improvement. In this view, a diagnostic profile should include both learners' strengths as well as their weaknesses (Jang 2012: 124).

Since useful feedback is of essential importance in enhancing students learning abilities, it would be ideal if students are assessed through tests which are designed purely for diagnostic purposes. However, these tests are rare. More often, already established tests, such as placement tests, may be adapted to serve diagnostic purposes (Alderson 2005: 6). This adaptation has considerable consequences for the usefulness of feedback and subsequently on the diagnostic value of this feedback. Diagnostic tests are often in close alignment with the curriculum, and their feedback can in such a case easily be integrated in the teaching process. However, adapted diagnostic tests may lack this alignment. Therefore, feedback provided by such tests is often norm-referenced and may lack pedagogically meaningful information (Kunnan and Jang 2009: 612). This entails that there is little interpretation of test scores and a subsequent absence of guidance for improvement.

In much the same fashion, Alderson (2005) has devised a list of features of diagnostic tests in order to clarify the characteristics of such assessments which can facilitate a better understanding of how to design a diagnostic test, or adapt an already established test to serve diagnostic purposes. He emphasizes that these features are merely hypothetical and not definitive statements on what diagnostic testing is, but ones which could serve as a starting point for research:

1. Diagnostic tests are designed to identify strengths and weaknesses in a learner's knowledge and use of language.
2. Diagnostic tests are more likely to focus on weaknesses than on strengths.
3. Diagnostic tests should lead to remediation in further instruction.
4. Diagnostic tests should enable a detailed analysis and report of responses to items or tasks.
5. Diagnostic tests thus give detailed feedback which can be acted upon.
6. Diagnostic tests provide immediate results, or results as little delayed as possible after test-taking.
7. Diagnostic tests are typically low-stakes or no-stakes.
8. Because diagnostic tests are not high-stakes they can be expected to involve little anxiety or other affective barriers to optimum performance.

9. Diagnostic tests are based on content which has been covered in instruction, or which will be covered shortly. (2005: 11)

These nine features all underline the notion that the main purpose of diagnostic assessment is to enhance learning, by identifying students' weaknesses and by providing detailed and purposeful feedback on which aspects of learning require extra attention.

According to this framework and the importance of useful feedback as mentioned above, a test like TALPS might appear to be inappropriate to serve diagnostic purposes. First of all, TALPS does not identify strengths and weaknesses in students' performance, but assigns levels of potential risk regarding academic language ability. Therefore, TALPS does not present a detailed analysis of students' performance on the test, but merely lists results. Although the results of TALPS are issued with as little delay as possible, this is not to facilitate the learning process but rather to be able to place students on the necessary follow-up courses quickly. In addition, TALPS is a high-stakes test which involves much anxiety, putting pressure on students which may (negatively) influence the test result. Therefore, students are less likely to perform optimally.

However, when approaching this framework from another perspective, one may also argue that TALPS can potentially serve a diagnostic purpose. Firstly, when considering the different results section-wise, and by performing an independent diagnostic analysis of each sub-result, one may possibly identify strengths and weaknesses in a student's language ability, as each subtest measures different components of what constitutes academic literacy (as articulated in table 2.2 above). Secondly, if TALPS is administered as a diagnostic test only, the results have no serious consequences for students, making the test low-stakes and subsequently creating a less anxious environment, allowing students to perform optimally.

Although Alderson (2005) claims that we need frameworks of language use, language learning and language development before creating a diagnostic test which assesses the level of students based on these frameworks, tests like TALPS which have already identified a level that students need to attain, can serve

diagnostic purposes by gearing the assessment towards providing constructive feedback on students' weaknesses in attaining the right level. This is an advantage of adapting a test that was originally designed for a different purpose.

2.7 Conclusion

This chapter has surveyed the critically important feature of the ability to use academic language competently: distinction-making through language. In the light of this definition the construct of academic literacy that underlies the TALPS test has been examined and various aspects of this construct related to distinction making have been identified. It has shown how the construct underlying TALPS is skills-neutral, but also that the processes of distinction-making culminate in 'writing'. Finally, while TALPS is not in the first instance a diagnostic test, it may be useful to analyse for that purpose. The next chapter looks specifically at an earlier attempt to do so, before outlining the research question for this study and formulating several hypotheses that might productively be investigated.

Chapter 3

Rationale

3.1 An earlier analysis of TALPS

From the issues touched upon in the previous chapter, it becomes apparent that writing is a crucial component of academic language ability, not the least because it may in many instances mark the end result of an often long and arduous process of finding, processing, and producing new, logically qualified information. It is often the medium through which students need to share the results of their research project in a final format (Butler 2007: 10). Writing is not just a ‘skill’ that can be acquired and mastered. Rather, it involves an entire process of gathering, processing, and producing information, relying in this process on the crucial component of academic discourse, distinction-making. As has been observed previously, if postgraduate students encounter difficulties with writing at an academically acceptable level, that can seriously affect their academic success. Although TALPS measures the writing abilities of postgraduate students, it does not automatically give insight into those components of academic literacy students struggle with, an aspect which may be of high relevance to “at risk” students to guide their subsequent process of improving their academic literacy abilities.

An earlier diagnostic analysis of the subtests of TALPS was conducted by Patterson (2012). This analysis revealed that the following components proved to be difficult for postgraduate students: understanding metaphor, idiom, connotation, word play, and ambiguity; understanding text types (genre sensitivity); communicative function (defining, arguing, etc.); and making meaning beyond the level of the sentence (Patterson 2012: 25).

Since this analysis did not include the writing section of the test, this study aims to diagnostically investigate whether the *writing section* of TALPS indicates similar difficulties as those observed by Patterson for the other subtests. Moreover, by determining which components of the construct apply specifically to writing, this study can identify specific aspects of academic language ability with which

students generally struggle. By diagnostically investigating the written sections of the 2012 and 2013 TALPS and their scores and correlating these with the results of the several subtests, I hope to present an answer to the question of how much more the completion of the writing section of TALPS may inform us in terms of the diagnosis of shortcomings in academic literacy, also when I compare its results to those of the remainder of the test.

A correlation analysis of the subtest results of TALPS with the essay results may not only reveal to which other test items the writing section closely relates, it can also predict whether the writing section score provides an extra dimension to the test, giving more insight into a student's academic language abilities, and an insight which can subsequently inform the further refinement of the test design.

3.2 Hypotheses

In order to answer the research question, the following three types of correlations will be performed with the data of TALPS (cf. Ito, 2005; Van der Walt and Steyn 2007:148):

- A correlation between each subtest and the whole test;
- A correlation between the essay subtest and the whole test minus the result of the essay subtest and;
- A correlation between each subtest and the essay subtest.

These correlation analyses can serve as a first set of evidence for arguments to validate the test construct. This process is operationalized by testing the following hypotheses:

- H1: The relationship of each subtest to the total result of the test is high, as this ensures a homogeneous test construct.
- H2: There is a relationship between the multiple-choice scores and the essay score, which indicates that both assess the same construct of academic literacy
- H3: The sections which measure those components that have been listed by Patterson (2012) as difficult for students will demonstrate a weak relationship with the essay result.

In addition, the descriptive analysis of the essays will provide a detailed insight into the specific aspects of academic language ability students struggle with. It is expected that:

- H4: Students will demonstrate difficulties regarding textual relations, communicative function, and understanding text type.

The validation of these hypotheses will ultimately provide an answer to the research question of how much more the completion of the writing section of the test may inform us regarding the diagnosis of shortcomings in academic language ability.

In the course of this analysis, I hope to find information underlining the proposed definition of academic discourse and gain more insight into the relevance of the construct on which the test was built, as well as whether the writing section indeed proves to be a relevant addition to the test. In conclusion, I shall list several practical implications for the further development of the test and make suggestions for the teaching of academic literacy through an intervention that focuses on the (eventual) improvement of writing. The next chapter will outline the design particulars for this study.

Chapter 4

Methodology

4.1 Overview

This chapter will outline the design of the quantitative correlation analysis of the scores of the writing section with other components of the test. In addition, it includes an in-depth descriptive analysis of the academic literacy abilities in student writing. The purpose of the design is to correlate the scores of sections 1-7 of TALPS with the writing section (section 8) as obtained according to a specific marking procedure outlined below, as well as measure the interrelationship of the different components of the construct of academic literacy in the students' essays according to a second marking procedure, also explained below. After elaborating upon the different marking procedures that are used to obtain the data, this chapter will provide information on the statistical procedures used, as well as their design and analysis procedure.

4.2 Subjects

The results of TALPS, written at the University of the Free State in 2012 and 2013, form the basis for the analysis of this study. A total of 405 students starting their postgraduate studies in various disciplines took the test in 2012, and 100 students took the test in 2013 (from January up till and including May). Of these students, n=53 took the test on the computer (35 in 2012 and 18 in 2013). In 2012, n=40 students did not write the essay, though their scores will be included in the correlation analysis. Because the test is administered to postgraduate students regardless of their discipline of study, the data obtained in this study derives from a varied group of subjects and is therefore representative of a larger population of postgraduate students. The results of the 405 students writing the test in 2012 will form the basis for the quantitative analysis. In addition, a sample of n=80 essays was drawn from the 2013 TALPS essays for an in-depth survey of academic

literacy abilities. These essays were chosen according to their length, as those assignments consisting of only a few sentences were excluded from the analysis.

4.3 Materials and procedures

The essay at the end of the test needed to be an argumentative piece of writing of approximately 300 words on the topic of global warming. More precisely, students had to present a structured argument to the question of what can be done in Africa to counter global warming. The writing instructions further list the necessity to give due recognition to the sources students use in their text (based on the texts in the test) and provide a general reminder to adhere to the generally accepted writing conventions (formality of register, logical structure and acknowledging sources were used as examples). Clearly the assumption is that students applying to study at postgraduate level already need to be familiar with the writing of argumentative texts.

Unlike the other seven subsections of the test, the written essay is an open form of assessment, which generates a qualitative type of data. When one wishes to perform statistical analyses like correlations, the data need to be transformed to a quantitative dataset. The quantification of qualitative data is a subjective process which relies heavily on the interpretation of the content by the individual coder/rater (Phatiki 2010: 41).

The data used in this study is obtained by marking the essays of 2013 with two different marking rubrics; the regular marking rubric used to assess the writing section of TALPS (hereafter referred to as marking rubric 1 – see Appendix B), and a rubric specifically designed for this study, which draws directly on the construct of academic literacy as outlined in chapter 2 (hereafter referred to as marking rubric 2 – see Appendix C). Even though both these marking rubrics provide a clear outline of the different aspects on which the mark will be based, the interpretation still differs per individual rater, which may influence the reliability of the results and, when not handled with care, the outcome may result in an invalid interpretation of the data (Lombard, Snyder-Duch, and Campanella Bracken 2005). Therefore it is important to ensure that even though the marking may be subjective,

there is consistency throughout the marking procedure, both on intra- and inter-individual rater level.

After examining the particulars of the different marking rubrics, this section will address the marking procedure with regard to consistency before outlining the procedure for the correlation analyses.

4.3.1 Marking rubric 1

First of all, the essays are marked according to the regular marking rubric used for TALPS. This rubric consists of an evaluation of the introduction, body, and conclusion in terms of content, and of an evaluation of the technical aspects and language use. Table 4.1 below lists the several components of marking rubric 1.

Introduction	Body	Conclusion	Technical aspects and language
Statement of issue – angle to be argued	Nature of problem/issue	Emphasising again the point of view advanced – link with introduction	Bibliography
Framing of reader expectations	Discussion of pros and cons	Clearly states again the most important issues	Academic style and referencing
	Argue convincingly for specific point of view		Grammar and spelling

Table 4.1: Components of marking rubric 1 of TALPS.

In total, a student can gain a maximum of 20 points for the essay (introduction, 4; body, 7; conclusion, 4; technical aspects and language, 5) ranked on a scale from poor to good as in table 4.2 below:

	Poor	Average	Good
Introduction	0-1	2	3-4
Body	0-2	3-4	5-7
Conclusion	0-1	2	3-4
Bibliography	0	0-1	1
Academic style and reference	0	1	2
Grammar and spelling	0	1	2

Table 4.2: Marking scale of each component of marking rubric 1.

All components together make up the total score for the essay, which will be the score for section 8 of TALPS. The marking is done by hand, and in this case by two individual raters. Although we have to bear in mind that no two raters mark exactly alike, the conditions for marking can be optimised to ensure minimal intercoder/rater differences in consistency. Thus, marking rubric 1 provides a clear outline of the different aspects of assessing writing and provides a clarification for each level (see table 4.3 below). In addition, since the raters succeeded each other in marking, the second rater only started rating after remarking some essays marked by the first rater. In this way, the second rater could imitate the marking style of the more experienced first rater. In table 4.3 below the criteria for assessment are outlined.

Content, language and style	Poor	Average	Good
Statement of issue – angle to be argued	No clear statement of issue; no point of view to be argued; abrupt or no introduction	States issue and point of view weakly; not clear what relevance is	Clearly states issue and point of view, explains relevance and importance
Framing of reader expectations	No or little interest in explaining clearly what will follow, or in guiding reader	Attempts unsuccessfully to frame reader’s expectations of what will follow	Clearly sets out what is to follow, providing a frame for what reader can expect
Nature of problem/issue	No or little discussion of the nature of problem/issue, or why it is necessary to deal with it	Unsuccessfully attempts to discuss nature of problem/issue and its importance in South Africa	Clear discussion of nature of problem/issue, and necessity of addressing it in South Africa
Discussion of pros and cons	Gives no or little indication that there is more than one side to an argument	Attempts to provide both pros and cons, but does so unconvincingly	Provides a comprehensive discussion of possible pros and cons
Argue convincingly for specific point of view	Argumentation is weak, one-sided, unconvincing	Argument deals with some of the important issues, but not in any convincing way	Strong, balanced argumentation that leaves the reader convinced of point of view

Emphasising again the point of view advanced – link with introduction	No connection between the issue/thesis introduced in the introduction and what is said in conclusion	Attempts to restate the issue/thesis, but does so unconvincingly	Clearly emphasises the thesis again without making it a word by word repetition of the introduction
Clearly states again the most important issues	No attempt to highlight again the most important issues in the text	Attempts to again include the most important issues, but does so in an unconvincing and incomplete manner	Clearly emphasises the main issues again in a structured and non-repetitive manner (exact repetition of the sentences used in body)
Academic style and referencing	No or little acknowledgement of authorities, weak structure, interrupted flow of argument	Argument patchy in its logic and structure; some acknowledgement of authority, but inadequate	Authorities used appropriately acknowledged, well-structured argument, logical flow
Grammar and spelling	The number of grammatical and spelling errors seriously interferes with the meaning	Contains some typical errors that could easily have been eliminated	Primarily error-free and fluent

Table 4.3: Indicators of level of each marking component of marking rubric 1.

Assessing a level is still highly subjective, although this outline may present a clear point of focus for each aspect of the rubric. It also becomes apparent that this marking rubric places the main emphasis on the content of the essay. In order to assess academic literacy, marking rubric 2 puts its primary focus on the use of academic language and structure.

4.3.2 Marking rubric 2

The rationale for devising a second marking rubric for the same set of data derives from the focus of this study: an attempt to find what diagnostic information is yielded by TALPS. Even though marking rubric 1 adequately measures students' writing abilities in terms of content, as stated above, it does not patently indicate

areas of academic literacy students particularly struggle with. Therefore, the essays are also assessed according to the construct of academic literacy, as outlined in chapter 2. Table 4.4 below lists the components of marking rubric 2, which closely follow the construct of academic literacy given above.

#	Component
1	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;
2	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
3	understand the communicative function of various ways of expression in academic language (such as defining, providing examples, arguing)
4	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
5	understand a range of academic vocabulary in context

Table 4.4: Components of marking rubric 2.

The underlying notion in the design of this marking system is the critical defining feature of academic discourse: distinction-making (Patterson and Weideman 2013a). I have included those components of the construct which all closely relate to distinction-making and that I have adjudged to be of significant importance in writing. Therefore, I have excluded components such as “interpret different kinds of text type (genre), and show sensitivity for the meaning that they convey, and the audience they are aimed at”, since there is no direct link to distinction-making in writing and because the specific “text type” students need to demonstrate (argument) can also, either in part or fully, be assessed with reference to the first and the third components of the rubric stated above. In addition, although interpreting is based on the ability to make distinctions, I have not included the ability to “interpret, use and produce information presented in graphic or visual format” as it is not always relevant for the writing of an argumentative text. Likewise, the ability to “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 purpose of an argument” is not always included in an academic argument, and because most of this is already substantially

covered in components 1 (sequence and order) and 2 (handle data that make comparisons) I have excluded this component from marking rubric 2.

Furthermore, each component of marking rubric 2 is ranked according to its relative importance for writing, in which *a* signals a main characteristic, and *b* and *c* minor characteristics of distinction-making in academic writing, as listed in table 4.5 below.

Component	Importance	Weight
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;	<i>a</i> (measures coherence as well as cohesion)	3
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	<i>a</i> (demonstrates information gathering, planning)	3
understand the communicative function of various ways of expression in academic language (such as defining, providing examples, arguing)	<i>a</i> (producing)	3
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	<i>b</i> (information gathering, planning)	2
understand a range of academic vocabulary in context	<i>c</i> (general language ability)	1

Table 4.5: Importance of each component of marking rubric 2.

Those characteristics I have labelled *a* all list typical features of distinction-making which most closely relate to writing. These components moreover list a productive type of academic language ability: structuring an argument, distinguishing between different aspects of an argument, and making use of the various communicative functions in expressing an argument. In essence, these components cover the planning and gathering of information as well as producing a well-written text, which, as was noted in chapter 2, are all critical aspects of academic language ability. The *b* characteristic exemplifies in much the same fashion the ability to plan and gather information. However, this component has a much narrower focus and closely links to the second component of making distinctions. Therefore I have labelled this component *b*. The last component is labelled *c* because, although it is a

demonstration of the ability to make distinctions regarding word choice, the connection with writing is more distant than for example the third component, which is also concerned with the communicative functions of academic expression.

In order to be able to perform calculations subsequently, the students' essays are assessed according to these components on a range of 1-4 in which, per component, one equals no or little presence of the component in the text, two equals insufficient presence, three equals relative presence, and four equals a good demonstration of the component in the text. Consequently for coding purposes, the scores for all *a* characteristics will be multiplied by three, the *b* characteristics by two and the *c* characteristics by one. In this way, I will obtain a weighted score for each essay which allows me to compare and contrast them.

Although this marking rubric was used by only one rater, also in this case consistency is of critical importance. Marking rubric 2 should therefore also provide a clear outline and clarification of the different components of assessment. This was accomplished by listing several indicators for each component, as summarised in table 4.6 below.

Component	Indicator(s)
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;	<ul style="list-style-type: none"> • The argument flows • Transitional words and phrases are used (for a list see appendix A, Kirszner and Mandell 2008) • And/or each paragraph has a clear topic sentence
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	<ul style="list-style-type: none"> • Academic references are used • Transitional words/phrases are used • A context is provided • The text shows the ability to connect evidence with explanations
understand the communicative function of various ways of expression in academic language (such as defining, providing examples, arguing)	<ul style="list-style-type: none"> • Transitional words/phrases are used (Kirszner and Mandell 2008)
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	<ul style="list-style-type: none"> • The references are relevant in the specific context
understand a range of academic vocabulary in context	<ul style="list-style-type: none"> • Appropriate use of words from the Academic Word List (AWL)

Table 4.6: Indicator(s) of each component of marking rubric 2.

It has been pointed out that it is important to ensure consistency between the scoring criteria for a test and the way the test construct is defined (Read 2010: 289). By inferring from the several components of the construct, I have tried to list concrete examples of language ability that might be found in the essays. As this marking rubric was used by only one rater, intracoder/rater differences play a role in ensuring consistency. These differences are limited by marking all essays in a short period of time, so as to ensure as few time lapses as possible between markings which could interrupt the ‘flow’ of marking. Nonetheless, the subjectivity in marking might pose limitations on the study, which will subsequently be elaborated upon in the discussion below.

4.4 Statistical analysis

The correlations and descriptive statistics are performed by using the Statistical Package for the Social Sciences (SPSS) and Microsoft Office Excel. SPSS is used to calculate the strength of the different correlations of the test results and the essay results with either a *Pearson r* or *Spearman rho* analysis, with reference to the instructions in Lowie and Seton (2012). This correlation analysis can, among other things, reveal whether the writing section is in line with the other test items since, when the writing section highly correlates with other sections, it is likely that these sections are influenced by the same factors, whereas a low correlation indicates influences from different factors (DeCoster 1998). Nonetheless, these correlations should be fairly low, since each test section is intended to measure a different component of academic language ability. Together, the sections complement each other in creating a test that measures the complete construct of academic literacy (Van der Walt and Steyn 2007: 147). Excel will subsequently be used to plot graphs based on the correlated data. In this manner, I hope to find out whether the correlations confirm the conclusions drawn by Patterson (2012) that were referred to in chapter 3.

Secondly, I will perform a descriptive analysis with Excel for the 2013 essay results based on marking rubric 2. This analysis will shed light on the potential added value of the writing section. It informs us on those areas of academic literacy with which students struggle the most, based on the weighted scores of marking rubric 2.

4.5 Design

As part of this study is based on test scores which assume equal intervals between scores, the variables are characterised as interval data. Therefore, the correlation that will be performed is calculated with a two-tailed *Pearson r* analysis. For the correlations and descriptive statistics with marking rubric 2, however, no equal distances between the scores can be assumed, because these were marked by hand and not calculated in the larger scheme of the test. Therefore, these variables are

characterised as ordinal data and are calculated with a non-parametric *Spearman rho* analysis. The alpha level is set in both cases at $p < 0.01$. Since the test is divided into several subtests which are not all equal in length, it may be necessary to recalculate the results of the several tests to a value between 0 and 1, to allow for subsequent meaningful interpretations.

Table 4.7 below attempts to outline the correlation pattern for the different variables, as stated in section 3.1 of this study. The column ‘Total’ signals the correlations between the subtest and the whole test minus the score of the subtest. The row ‘Total’ signals the correlations between each subtest and the whole test. An ‘X’ signals whether these two types of data will be correlated. For example, the result of section 1 will be correlated with each subsection and the total test result. Section 8, however, will be correlated with each subsection, the total test result, and the test result minus the result of section 8.

Section	1	2	3	4	5	6	7	8	Total
1									
2	X								
3	X	X							
4	X	X	X						
5	X	X	X	X					
6	X	X	X	X	X				
7	X	X	X	X	X	X			
8	X	X	X	X	X	X	X		X
Total	X	X	X	X	X	X	X	X	

Table 4.7: Outline of the correlation matrix.

In essence, each section will be correlated with the other sections and with the total result of the test, which will determine the inter-relationship of each section and the overall relationship of each section with the total value of the test. Secondly, each section will be correlated with the essay result of marking rubric 1 to indicate the degree to which either section measures the same components of the test. Thirdly, the essay result is correlated with the total result of sections 1-7 to display what kind of relationship exists between the multiple-choice and essay scores. Finally, the results of the different marking rubrics will be correlated, to investigate the

reliance of marking rubric 1 on the construct of academic literacy articulated in marking rubric 2.

The subsequent qualitative analysis outlined in the next chapter will provide an in-depth account of what students struggle with in academic writing, by examining 80 essays of TALPS 2013 per component of the construct, utilising marking rubric 2. Observations are linked to the various abilities of academic literacy as are evidenced in such aspects and functions as cohesion, coherence, distinguishing, and extrapolating. Together with the quantitative analysis of the results, this exploration of academic language ability can identify strengths and weaknesses and subsequently suggest how these weaknesses should be acted upon.

Chapter 5

Analysis of the results

Altogether, 405 essay results of 2012 were correlated with the multiple-choice results of TALPS. In addition, 80 essays written in 2013 were subjected to a close inspection, as noted in the previous chapter. This chapter will outline the results of both the quantitative and the qualitative analyses, thereby connecting the quantitative data to the qualitative observations and articulate possible implications, which will be discussed further in chapter 6.

5.1 Quantitative analysis

Although the original data are derived from sections of the test with results that differ in the number of questions they contain, recalculating the data to a value between 0 and 1 to allow for fair, meaningful interpretations proved unnecessary. When I compared the correlations calculated with the original as well as the recalculated data, a minimal difference was observed, which indicated that recalculating the data does not rule out the issue of section length. Therefore, I have chosen to perform all the calculations using the original dataset, which allows me to base my interpretations on real life data whose results are more reliable when transferring it to a larger population.

When analysing the raw scores for each student on each separate item, as well as the test as a whole, we note first that the overall consistency of the test displays a high alpha-value (Cronbach's $\alpha = .927$), suggesting a homogeneous test construct. Each subsection result was subsequently correlated with the total result of the test to check the strength of the relationship of the different sections to the overall result, displayed in the last row of the correlation matrix in table 5.1 below.

Section	1	2	3	4	5	6	7	8	Total
1									
2	.297								
3	.233	.444							
4	.201	.329	.302						
5	.214	.576	.476	.303					
6	.281	.573	.481	.292	.583				
7	.310	.491	.409	.273	.469	.632			
8	.192	.454	.402	.271	.433	.540	.507		.592
Total	.428	.763	.646	.465	.791	.841	.756	.720	

Table 5.1: Correlation matrix of the three different types of correlation.

The results of the quantitative correlation analysis presented in table 5.1 demonstrate that indeed, as suggested by hypothesis one (cf. section 3.2), each subsection displays a moderate or strong relationship with the total test result which, in combination with the high alpha-level of .93, ensures a homogeneous test construct.

The correlated values in table 5.1 indicate an overall strong relationship, with $r=.8$ being the highest value. The lowest values can be noted in sections 1 (scrambled text) and 4 (text type), which show a moderate relationship of $r=.4$ and $r=.5$ respectively. This low result could be due to the fact that these two sections have relatively fewer questions in comparison to the other sections of the test. However, when considering which aspects of academic literacy these two sections measure, it can be noted that both sections measure text relations, understanding text type, distinguishing, communicative function, and making meaning beyond sentence level. As noted in chapter 3, Patterson (2012) observed that students generally struggle with the following aspects of academic language ability: understanding text types, communicative function, and making meaning beyond the level of the sentence. Therefore, it could also be argued that these two sections do not correlate well with the rest of the test items since students underperform on them, a possibility which will be further explored below, as we consider the role of academic writing on students' test performance.

A *Pearson r* correlation of the multiple-choice results and the essay results revealed a moderately strong but significant relationship between the two of $r=.6$,

$p < 0.01$, thereby confirming that both assess the same construct of academic literacy as proposed in hypothesis two (cf. section 3.2). In addition, each section was correlated separately with the essay results, as summarized in the penultimate row of table 5.1.

Interestingly, the r-values in this row seem to mimic more or less the values of the overall correlations in the last row. Also between the essay result and sections 1 and 4 the significant relationship is relatively weak. The strongest relationship is again found with section 6 (Grammar and text relations). Figures 5.1 and 5.2 below display the correlation of the first 50 results for section 1, the weakest, and section 6, the strongest relationship, respectively. The data in this figure are the recalculated data which displays the correlation more adequately in visual format than when the original data are used, since then one line would be higher in the graph than the other, making it more difficult to visualise the relationship between the two sections.

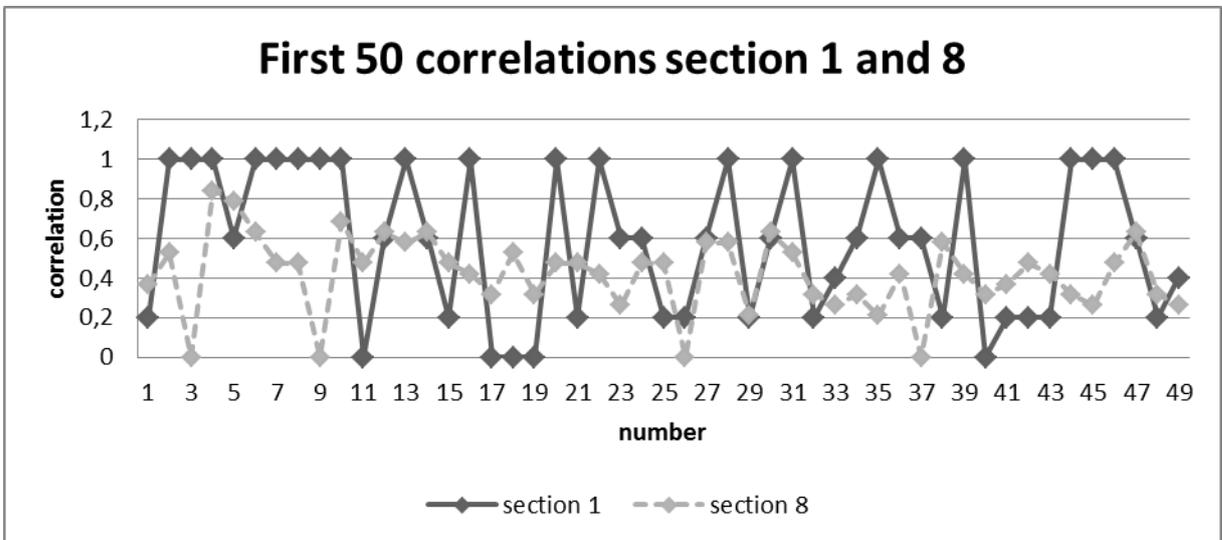


Figure 5.1: Correlation pattern of the recalculated results of section 1 with section 8.

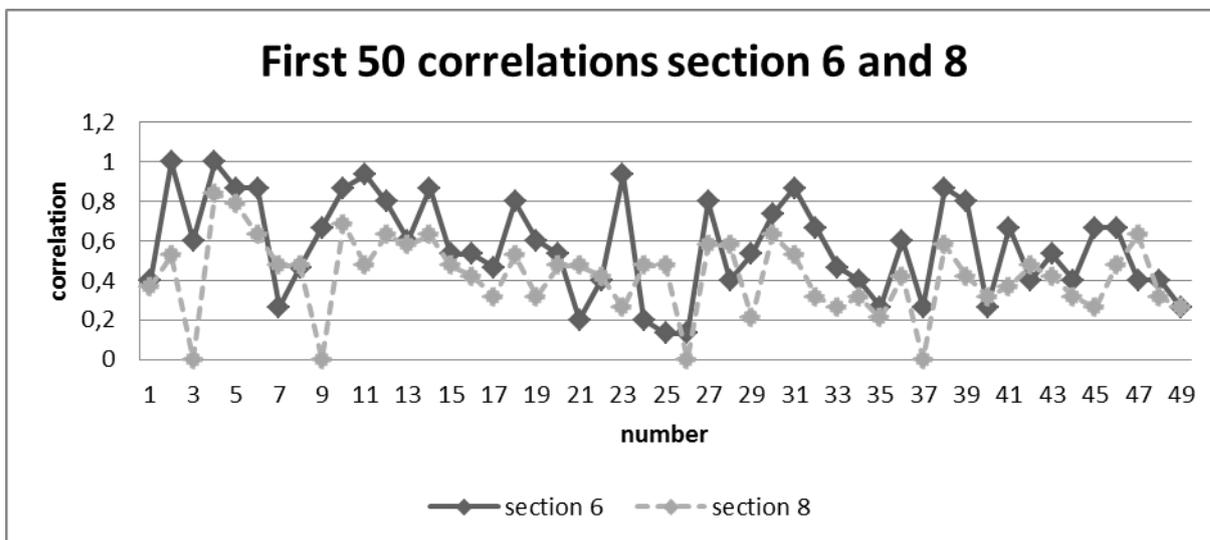


Figure 5.2: Correlation pattern of the recalculated results of section 6 with section 8.

The weak relationship of section 1 and 8 can clearly be noted in the alternating pattern of Figure 5.1. If the result of section 1 is high, the result of section 8 is (mostly) low. Figure 5.2, on the other hand, displays a more homogeneous pattern. Although there are some alternating points in the data, the general trend is that if section 6 shows a high result, the result of section 8 will probably be high as well. Even though these figures represent only a fraction of the data, they do provide an adequate illustration of the different strengths in relationship between the different sections.

As has been indicated in the previous chapter, the internal correlations should ideally fall in a range of 0.3 to 0.5 to ensure that the sections test different attributes and make a meaningful contribution to the test (Van der Walt and Steyn 2007: 148). The strong relationship of section 6 with the essay section is an indicator that the essay section measures those components that are also assessed in this section. This value or $r=.540$ nearly exceeds the 0.5 desired boundary. The question can therefore be asked whether section 6 (grammar and text relations), as well as sections 2 (interpreting graphs and visual information) and 7 (text editing) which also approach the upper, 'too high' parameter of 0.5 and demonstrate a strong inter-relationship as can be noted in table 5.1, are not testing the same components as the essay section.

The weak relationship with sections 1 and 4, on the other hand, with values just, or even below, the desired value of 0.3, suggests that these components of academic language ability are not addressed in the writing section, even though text relations and communicative function are indispensable aspects of academic writing. Moreover, when considering the components of the construct measured in sections 6 and 7, it becomes apparent that both text relations and communicative function are components which, like sections 1 and 4, are also assessed in sections 6 and 7. This observation reinforces the notion that section length is decisive in the strength of the relationship rather than the components that are measured. It also seems to suggest that students do show the ability to handle communicative function and text relations in their writing, since the strong relationship of sections 6 and 7 with the essay result indicates that both these elements are present in the essays. Therefore, it is vital to take an in-depth look at the students' essays.

Marking rubric 2 (see appendix C) was used to assess the essays on the presence of each component of the construct of academic literacy most closely related to writing. Table 5.2 below summarizes the weighted scores on each component for the total of 80 essay results.

#	Component	Maximum possible score	Total score
1	Text relations	960	411 (43%)
2	Distinguishing	960	426 (44%)
3	Communicative function	960	375 (39%)
4	Extrapolating	640	258 (40%)
5	Vocabulary	320	163 (51%)

Table 5.2: Weighted total scores per component of marking rubric 2

Overall, the scores on each component represent only half of the total score or less, indicating a general inability to write academically. Table 5.2 confirms the earlier observation made above and by Patterson (2012) that students struggle with communicative function, which is the lowest scoring component in marking rubric 2. When considering the correlation results in table 5.1 together with the observation from table 5.2, hypothesis three (cf. section 3.2, a weak relationship of the essay result with those components which have been listed by Patterson (2012)

as difficult) seems to be valid. The essay results, both of marking rubrics 1 and 2, demonstrate a weak relationship with sections 1 and 4 which focus on text comprehension and text types, as well as with the ability to understand and use the appropriate communicative function. These elements have been listed by Patterson (2012) as difficult for students.

To indicate how closely marking rubric 1 resembles the construct of academic literacy, on which marking rubric 2 was based, the results of both marking rubrics were correlated. A *Spearman rho* correlation revealed a significantly strong relationship at $r=.7$, $p<0.01$. This outcome ascertains that marking rubric 1 is a valid instrument in measuring the writing section, as it adequately reflects the components of the construct of academic literacy. This does not mean that marking rubric 1 cannot be improved, however, and below I shall make a recommendation about a possible augmentation of marking rubric 1 with certain salient elements of marking rubric 2.

5.2 Qualitative analysis

The instructions for the essay state that students should present an *argument* to respond the question of what can be done in Africa to counter global warming. In practice, the majority of the essays are *answers* to this question. This observation has several implications for students' academic language ability, both structure-wise and content-wise.

5.2.1 Text relations

First of all, a close scrutiny of the essays displays that they suffer from a general lack of coherence. The argument structure is absent: most essays consist of only one or two 'body' paragraphs. The introduction is often very general, starting with a statement along the lines of "global warming is a problem affecting the world" and some references to the increase in temperatures in Siberia and Alaska. The problem statement is often unclear or poorly stated:

- (a) A lot of possible solutions have been advanced with regard to the solution of this problem.
- (b) Therefore it is crucial to ensure that the nature is conserved by taking care of the planet.
- (c) ... warming is greater in the high latitudes of the northern hemisphere which is not a good idea to us in Africa because we need a balanced temperature.
- (d) [global temperatures increase] Hence, it is very crucial for the country to emphasize the importance of tree plantation.

Many essays subsequently present a list of possible solutions to counter global warming instead of focusing on one or two ideas and discussing these in depth, which often results in a jump from very general to very detailed information, without providing the reader with enough information and context to understand the proposed solutions:

- (a) Planting more trees, building mud, aluminium, zinc houses these will reduce the heat or counter global warming.
- (b) Suggested solutions are the use of windmills, solar panels and nuclear plants.

Conclusions are generally absent, and when a conclusion is present, it often fails to provide a short summary of the argument and rather presents a general statement:

- (a) A need to go green can never be over emphasised.
- (b) All matters still to be discussed.

Nonetheless, in the few well-developed texts, the conclusion can provide a good summary of the outlined argument:

- (a) In my view people should first learn or be informed about the challenge. Sources of the problems be identified, further investigation be conducted to verify facts and findings. It is important to take action.
- (b) Although South Africa is a middle income country like Brazil, it should be clear from the above arguments that this country can play a positive role in the combatting of global warming.

These findings are evidence of the difficulties students have with understanding textual relations. The overall score on this element of academic literacy presented in table 5.2 is a meagre 43%, confirming the observation that handling relations between different parts of a text presents a hurdle for many students. When we recall Patterson's (2012) findings, we note that students struggle with the ability to see textual relations at a receptive level as well, which can also be observed in the insufficient display of the understanding of textual relations in the scrambled text section (subtest 1).

Transitions between paragraphs, sentences, or parts of an argument should be clearly introduced by either topic sentences or transitional words and phrases to develop cohesion. There is minimal use of these indicators to link the various statements to each other, and often these linking words or phrases are informal (Also see point 5.1.3):

- (a) The other thing
- (b) Also; this
- (c) Furthermore; therefore

Cohesion within and across sentences is often problematic, in the use of vocabulary (elaborated upon below) as well as grammar:

- (a) Lack of conservation of forest and ignorance of ecological safe greening environmental methods has seen most developing African countries adapting the environmental hazardous methods of generating their energy source.

Omitting of articles:

- (a) African continent is at an advantage stage.

Run on sentences:

- (a) Furthermore, enquiry based research can be encouraged to determine the extent of damage or impact of global warming presently so as to determine best solutions to the challenge faced.

Subject-verb agreement:

- (a) Of course, that is trees.
- (b) Those who illegally chops down our trees and burns down our forests.
- (c) Some places does not get rain anymore.

Vague content structure:

- (a) [...] then Africa's global warming will be in better control and of much more value to the society.

This general lack of structure, both at inter- and intra-textual level, may be attributed to the different courses of study, each with its own requirements for written work (e.g. some departments may value content over structure), which may not always coincide with other fields of study. In addition, this issue of cohesion is related to the ability to understand the various ways of expressing in academic

language, an aspect that also proved to be problematic for students, as noted above (Patterson 2012).

5.2.2 Distinguishing

Many essays make use of the texts given in TALPS. Students are thus able to distinguish essential from non-essential information by extracting information from other texts and applying this information to their own essay. However, this information is often only partially integrated in the argument, or not at all:

- (a) [photosynthesis] In a study by ‘Proceedings of the National Academy of Sciences, vol. 103 p. 14288) which states that another decade of business –as-usual carbon emissions will probably make it too late to prevent the ecosystems of the north from triggering runaway climate change. [general conclusion].

As a result, the majority of the claims lack a clear context. Since most arguments are based purely on opinion, they often fail to include background information or to back up arguments with evidence from texts:

- (a) Climate change may in future cause plants not to grow effectively on places where certain fruits grow example.
- (b) This is also one reason why Africa must preserve its forests because of this wonderful occurrence such as photosynthesis.
- (c) The task of dealing with carbon emissions. The result is unnecessary loss of vegetation.

Example (a) above clearly indicates an opinion, which, also because of its illogical structure, is not convincing as an argument, since there are no facts to back up this statement. Example (b) highlights photosynthesis as a reason for Africa to preserve forests. However, why forests or photosynthesis is of such importance to global warming in the first place is not explained. In the same fashion as the second example we find many arguments which highlight effects of global warming but do not provide possible causes of this problem. Numerous essays list solutions for decreasing carbon emissions, without explaining why carbon emissions are a cause of global warming. Example (c) above illustrates the leap in reasoning without a context, or an information base.

5.2.3 Communicative function

The examples below serve to indicate various ways in which students express an argument, their opinion, or present examples or evidence. Even though these examples also show inappropriate reasoning or illogical inferences, they will be elaborated upon from a communicative point of view.

Most arguments make use of modals such as “must”, “need”, and “should”. Also, markers as “therefore”, “for example”, “such as”, “as well as”, “however” and “furthermore” are used to signal contrast, an example, or introduce elaboration.

In addition, an argument can consist of an opinion alone:

(a¹) According to the author, the people of the world can do something about this problem.

A contrasting element:

(b²) On the other side, one can say it is nature that cannot be stopped but be maintained through education and research, preparations for the worst.

A suggestion:

(c³) What can be suggested is ...

Or a question can be asked to make a point:

(d⁴) If everyone living in Africa plants one tree a year, wouldn't that be a wonderful start to literally building a 'greener' Africa?

What all these examples have in common is that none of them adequately expresses the argument in academic terms. For each example there is a more suitable academic equivalent:

(a¹) Omit “according to the author”.

(b²) On the other hand, it can be noted that...

(c³) Possible suggestions are...

(d⁴) Too emotional (“wonderful”) and a drop in register (“wouldn't”).

Secondly, most evidence is incorrectly cited in the text, not at all cited, or it is assumed that the evidence is either known to the reader or is a common fact:

(a) given literature states that ...

The general lack of argumentative language is another indication of the failure to fully grasp the features of an argumentative essay. Although students seem to be

able to argue with modals such as “should”, “need”, and “must”, they do not exploit the entire concept of arguing for a case which includes, in addition to opinion, facts that may be presented as evidence, along with context and the ability to compare and contrast. There are elements that are structurally lacking in the students’ essays - consider the following example which demonstrates an argument based solely on opinion and assumed, inexplicit knowledge: “*the more the factories the high possibility of global warming*”. In addition it is necessary to write with (some degree of) authority (cf Butler 2007: 249). The essays, however, show a general lack of an authoritative stand.

5.2.4 Extrapolating and inferring

As stated above, a limited number of facts deriving from other sources on global warming are presented and evidence is not, or wrongly, inferred. Therefore, the majority of the essays show no signs of extrapolation. There is one essay in which evidence is presented, inferred, and extrapolated:

- (a) Many countries, like Brazil (Kluger, 2007:40-41) have made big advances in the use of biofuel to reduce their carbon footprint” [...] “Although South Africa is also a middle income country like Brazil (UNEP) it should be clear from the above arguments that this country can play a positive role in the combatting of global warming.

This example indicates that the student has extracted information from another text, compared this with the situation in South Africa, and concluded that by considering Brazil as an example, South Africa can contribute to the combatting of global warming. This case indicates the ability to extrapolate by generalizing over countries with a middle class economy, and the ability to subsequently infer solutions that will be beneficial to South Africa.

5.2.5 Vocabulary

The highest scoring element of marking rubric 2 is vocabulary comprehension. Students seem to be aware of what constitutes academic vocabulary, although the marking here may be subjective, as what is classified as academic vocabulary is debatable. In the example below the academic words are emphasised:

- (a) Lack of **conservation** of forest and **ignorance** of ecological safe greening **environmental methods** has seen most developing African countries **adapting** the environmental **hazardous** methods of **generating** their energy source.

Even though students know which academic word to use, the sentences they construe with it are neither always coherent nor clear in content, resulting in contradictory statements, especially when a context is lacking. As can be inferred from the examples in this analysis, students display an adequate awareness of academic vocabulary, although there are some occasional lapses in register and misuses of academic terms.

5.3 Answering the research question

Before answering the research question (cf. section 3.1) and elaborating upon the possible implications the results of the analyses may have, I will return to the last hypothesis that was articulated above in section 3.2, namely that students will demonstrate difficulties regarding textual relations, communicative function, and understanding text type.

In the course of this analysis, it has been proven that hypotheses one, two, and three are valid. The fourth hypothesis requires a closer look at the results from the qualitative analysis. The quantitative correlations reveal that students struggle with understanding text type and textual relations, which conforms to hypothesis four. The qualitative analysis demonstrates students' difficulties with textual relations as well, both at inter- and intra-textual level. Consequently, students are unable to use the appropriate academic words and phrases to signal linkages between the different parts of their argument, to introduce examples, or to express their opinion in an academic manner.

The initial question asked in this study was to find out what the writing section contributes to the test result regarding its potential to diagnose students' struggles with specific components of their academic literacy. From the above analysis of the results it becomes clear that the writing section aligns with the other seven sections of the test and can thus serve as another predictor of those aspects of

academic literacy with which students struggle. It has been found that students who perform well on those sections which measure textual relations and text type also score high on the essay, and vice versa. This finding has several implications for the design of TALPS, to which I will turn next.

Chapter 6

Discussion of possible implications

6.1 Revisiting the results

The aim of this study was to find out how a diagnostic analysis of the results of TALPS, and the essay section in particular, can inform test users, test designers, curriculum designers and teachers of the difficulties students encounter regarding academic literacy, and academic writing in particular.

Of course, language learning, and developing academic language ability for that matter, is a highly individual process and generalizing over this process can never adequately reflect the process that all language learners encounter. However, in designing tests and courses it is necessary for us to make these generalizations, since developing a unique language learning course (or test) for every individual learner is (still) too expensive, unproductive, and time-consuming. What one can do, however, is to provide enough support during learning so as to be able to provide each individual with useful information regarding their required level.

The analysis of the results above demonstrates 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.

This observation can have several potential causes. As mentioned in the previous chapter, different departments may have different prescriptions for written assignments, thereby placing the focus on either content or structure, which may result in poorly structured, though content-wise well-informed essays.

Another issue which may be relevant in the understanding of students' performance is time. It is plausible that the writing section does not adequately display students' optimal academic writing capabilities. Time is an indicator of fluency and thus subsequently of academic language ability (Van der Slik and

Weideman 2010: 115). Therefore, the fact that students were not able to complete the writing section, stopping in mid-sentence, reasoning illogically, and presenting a summary of possible solutions, may have resulted in the poorly crafted arguments they present.

The third possibility is that, since the essay is the final part of the test, students had already sat through 76 questions and may not be performing optimally on the essay due to lapses in concentration or tiredness. Most likely, all three possibilities have influenced students' performance.

6.2 Design suggestions

6.2.1 A two-tiered test model

Given the potential causes of students' struggles with academic writing suggested above, the question can be asked whether the essay section of the test truly reflects students' academic writing abilities in the form it is currently presented.

There is no doubt that academic writing constitutes a vital part of academic language ability (see section 2.5 of this study) and, moreover, that the face validity of a test is seriously affected if a writing section is omitted (Butler 2009: 297). Nonetheless, I have demonstrated that the components measured in the essay section of TALPS closely align with what is measured in the other sections of the test, and specifically with sections 6 and 7 (the highest correlating sections). Therefore, it is safe to assume that when a student performs well on sections 6 and 7, s/he will most likely also demonstrate an adequate ability to write academically. On the other hand, those students who display a low score on the essay most likely score low on the other sections of the test as well, an observation that reinforces the relationship found between the multiple-choice results and the essay section.

As I have remarked in chapter 5, most essays are unfinished and, although time is an indicator of fluency in academic discourse, it is now an important if not the main deciding factor through which many essays fall short in their ability to display an adequate awareness of what constitutes academic literacy. For that

reason, it might be worthwhile to explore the idea of modifying the design of TALPS by creating a two-tiered testing model.

This two-tiered model may be operationalized by splitting TALPS into two tests, in which the first test is composed of the multiple-choice sections (section 1-7) and the second test is an essay assignment, similar to the current assignment in section 8, but with a more substantial number of words and more source material from which students can select information. All students subsequently take the first test, and those who score below the current cut-off point of 60% (less risk), and/or score low on either section 6 or 7 (because of their high correlation with the essay section), will have the opportunity to sit the essay test, which serves as a second chance. This design can have several benefits.

First of all, students write the second test at a different moment in time, which allows them to concentrate fully on the writing assignment without feeling worn-out from having answered 76 multiple choice questions just before. Since the time-pressure is reduced, students are better able to plan their essay and to structure their argument. This study has demonstrated that in general structure is absent in the students' essays; an analysis of the essays written under the proposed two-tiered model can therefore determine how large the influence of time is on the currently ill-structured essay, and which part truly reflects poor academic writing skills.

Secondly, the essay can be longer, around 500-800 words, which allows students to present a comprehensive argument with a clear introduction, problem statement, and arguments and evidence that provide support for the claims made. The current essay asks students to provide in just 300 words a comprehensive piece of text with a clear content-outline, background information, suggestions or solutions to the problem, thereby touching on both sides of the argument, and a meaningful conclusion. Although most students do not reach the 300 word-limit, those who do often exceed it to fit in all the aspects of the argument. It seems to be genuinely difficult for students to squeeze a comprehensive argument into just 300 words, resulting often in only a body paragraph, with a list of suggestions or solutions to a certain problem.

In addition, a longer essay can be a more accurate reflection of a student's academic writing skills. It potentially allows the marker to provide a more accurate assessment and can give a better insight into the particular writing difficulties a student struggles with. This can inform the student on the subsequent level of academic literacy s/he has reached and what aspects s/he still needs to improve in order to attain the required level of academic language ability.

Thirdly, a second test can serve as a second chance for those students who may have been negatively affected by the first test, in the sense that they may have been misclassified as being at risk. Misclassifications may occur in each test because tests are never entirely reliable measuring instruments, according to Van der Slik and Weideman (2005: 28). With a TiaPlus test and item analysis of the results the number of misclassifications can be identified and by offering this number of students a second chance, misclassifications can be set right.

Nonetheless, because each section of the test is based on a component of the underlying theoretical construct, one cannot simply take out or modify sections based on empirical grounds alone (Rambiritch 2012: 114). Therefore, it is necessary to consider such changes carefully and to be transparent to all stakeholders of TALPS.

6.2.2 Other design options

Even without a two-tiered test model, performance on the writing section can possibly be optimised by providing more detailed guidelines regarding the structure of the argument in the essay instructions. The lack of cohesion and coherence in the essays also affects students' abilities to express their arguments effectively, to combine fact with opinion and to distinguish between cause and effect.

Additionally, feedback on the writing section may be beneficial to those students who are identified as being at risk. The current norm-referenced scoring model of TALPS presents an overall score only. Providing feedback can aid students in enhancing their academic language abilities and facilitate learner autonomy, as has been remarked in section 2.6 of this study. By for example

presenting the information in table 4.3 on which the essay score is based, students will be able to better understand their essay mark and are more aware of what they can and cannot yet do. Alternatively, the essay marking rubric could be modified and in addition to content, focus on particular (structural) components that have proven difficult for students. The modification suggested in appendix D (marking rubric 3) presents just such an option.

This new marking rubric has been designed with reference to both rubrics employed in this study, and is intended to give the best combination of marking rubric 1 and 2 in order to optimise further their functioning as assessment instruments of postgraduate writing ability. In designing this rubric, marking rubric 1 is used as blueprint and modified to incorporate the elements of marking rubric 2.

The first component of this rubric, content, closely resembles marking rubric 1. This aspect of the rubric adequately assesses the essays on content and it includes the issue of distinction-making – important in marking rubric 2 – in its assessment. By distributing the total possible score of 20 points equally between content and language organisation, a fair attention to the two is ensured.

The second component of the marking rubric is concerned with the language and structure of the students' essays. It is divided into four separate aspects to include the organisational, as well as the academic language element. I have incorporated the aspect of inter-textual coherence – coherence between the sections of the text – from marking rubric 2 into this rubric, as it was observed in this study that students display a clear lack of structure in their essays. The components of marking rubric 1 regarding structure in the introduction, body, and conclusion are used to assess the essays on inter-textual coherence in this rubric. Since structure is of vital importance in academic writing, this element has a maximum score of 4 points. In addition I have separated the ability to reference appropriately and to apply an academic writing style, as these are two distinct, important issues in the ability to write academically. The grammar and spelling aspect has again been directly taken from marking rubric 1.

This new marking rubric now accommodates elements from marking rubrics 1 and 2 which ensures – through an equal distribution of scoring points between

content and language – a more extensive coverage of those aspects important to academic writing.

Furthermore, optimising the writing section of TALPS can yield more information on students' academic language abilities, which may aid course designers and teachers to design material better tailored to students' needs. Consequently, courses may be more effective in aiding students to improve their academic language abilities. 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 student's essays, courses can focus on distinction-making as a central theme. This design suggestion for a course designed to focus more on the organisation of academic writing is, however, beyond the scope of this study. Nonetheless, the outcomes of this study have provided a set of better understandings of students' academic language abilities, which may be beneficial to designing further intentions for the development of academic literacy at postgraduate level.

Chapter 7

Conclusion

7.1 Purpose of the study

The previous chapter has attempted to interpret the results of this study with a view to improving TALPS and subsequent academic literacy courses. In this chapter, I wish to return briefly to the main purpose of the study, which was to present a detailed insight into those components of academic language ability students struggle with.

The diagnostic analysis of TALPS, and of the writing section in particular, has confirmed the findings of Patterson (2012) and has noted a consistent lack of structure in the essays. The correlation analysis provided evidence for this absence of the ability to structure academic writing, as those sections that tested components related to language organisation (text type, scrambled text, and communicative functions) revealed low correlational values. Not only did chapter 5 indicate that the relationship between the several subtests makes TALPS a coherent test, it also confirmed the usability of the underlying theoretical construct of academic literacy.

7.2 Limitations

Because this study has only examined a limited number of test scores and essays the various correlation values may indicate peaks that, when a larger dataset is employed, would have been flattened out. Nonetheless, when this data is compared to other studies drawing on the TALPS test results it can be noted that Cronbach's alpha-level is consistent, as well as the several subtest correlation values.

Secondly, as has been indicated in chapter 4, the marking of the essays could ideally have been more time-restricted. I started marking in the first week of May

and finished only near the end of the month, as several students wrote the test during these weeks, which resulted in a constant flow of new essays. This might have influenced the rating consistency even though marking was limited to one person.

7.3 Final word

From a DST perspective, it can be noted that all aspects of language are connected and dynamically interact (cf. chapter 2, Verspoor and Behrens 2011). This has also been proven in this study: structure affects content and vice versa, and structure can be affected by the use of the appropriate communicative functions and an understanding of the specific text type. It has been demonstrated that when these elements are lacking, the argument is not coherent and ultimately fails to be effective.

This conclusion brings us back to the issue that initiated this study. Recalling chapter 1, we may observe that poor academic language abilities affect academic success. Although TALPS can distinguish different proficiency levels and assign levels of risk to postgraduate students, the actual problem should already be dealt with earlier on in a student's education. This diagnostic analysis has again highlighted the consequences of insufficient educational preparation for the meaningful participation by students in postgraduate studies. Without such preparation, language may well prove to be a barrier that prevents them from benefitting fully from the high-level analytical processes they will be exposed to at this level.

Appendix A

List of transitional words and phrases

To signal sequence of addition:

again, also, besides, first ... second ... third, furthermore, in addition, moreover, one ... another, too

To signal time:

after, afterwards, as soon as, at first, at the same time, before, earlier, finally, in the meantime, later, meanwhile, next, now, since, soon, subsequently, then, until

To signal comparison:

also, in comparison, likewise, similarly

To signal contrast:

although, but, despite, even though, however, in contrast, instead, meanwhile, nevertheless, nonetheless, on the contrary, on the one hand ... on the other hand, still, whereas, yet

To introduce examples:

for example, for instance, namely

To signal narrowing or focus:

after all, indeed, in fact, in other words, in particular, specifically, that is

To introduce conclusions or summaries:

as a result, consequently, in conclusion, in other words, in summary, therefore, thus, to conclude

To signal concession:

admittedly, certainly, granted, naturally, of course

To introduce causes or effects:

accordingly, as a result, because, consequently, hence, since, so, then, therefore

(Kirszner and Mandell 2008: 83-84)

Appendix B

Marking rubric 1

Content and organisation		Poor 0-1	Average 2	Good 3-4
Introduction (20)	Statement of issue – angle to be argued	No clear statement of issue; no point of view to be argued; abrupt or no introduction	States issue and point of view weakly; not clear what relevance is	Clearly states issue and point of view, explains relevance and importance
	Framing of reader expectations	No or little interest in explaining clearly what will follow, or in guiding reader	Attempts unsuccessfully to frame reader’s expectations of what will follow	Clearly sets out what is to follow, providing a frame for what reader can expect
		0-2	3-4	5-7
Body (argument) (50)	Nature of problem/issue	No or little discussion of the nature of problem/issue, or why it is necessary to deal with it	Unsuccessfully attempts to discuss nature of problem/issue and its importance	Clear discussion of nature of problem/issue, and necessity of addressing it
	Discussion of pros and cons	Gives no or little indication that there is more than one side to an argument	Attempts to provide both pros and cons, but does so unconvincingly	Provides a comprehensive discussion of possible pros and cons
	Argue convincingly for specific point of view	Argumentation is weak, one-sided, unconvincing	Argument deals with some of the important issues, but not in any convincing way	Strong, balanced argumentation that leaves the reader convinced of point of view
		0-1	2	3-4
Conclusion (20)	Emphasising again the point of view advanced – link with introduction	No connection between the issue/thesis introduced in the introduction and what is said in conclusion	Attempts to restate the issue/thesis, but does so unconvincingly	Clearly emphasises the thesis again without making it a word by word repetition of the introduction
	Clearly states again the most important issues	No attempt to highlight again the most important issues in the text	Attempts to again include the most important issues, but does so in an unconvincing and incomplete manner	Clearly emphasises the main issues again in a structured and non-repetitive manner (exact repetition of the sentences used in body)
Language and style				
Technical aspects and language (10)	Bibliography	0-1	2-3	4-5
	Academic style and referencing	No or little acknowledgement of authorities, weak structure, interrupted flow of argument	Argument patchy in its logic and structure; some acknowledgement of authority, but inadequate	Authorities used appropriately acknowledged, well-structured argument, logical flow
	Grammar and spelling	The number of grammatical and spelling errors seriously interferes with the meaning	Contains some typical errors that could easily have been eliminated	Primarily error-free and fluent

Appendix C

Marking rubric 2

Component	Indicator(s)	Points possible	Weight
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;	<ul style="list-style-type: none"> • The argument flows • Transitional words and phrases are used (for a list see appendix A, Kirszner and Mandell 2008) • And/or each paragraph has a clear topic sentence 	(1-4)	3
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	<ul style="list-style-type: none"> • Academic references are used • Transitional words/phrases are used • A context is provided • The text shows the ability to connect evidence with explanations 	(1-4)	3
Understand the communicative function of various ways of expression in academic language (such as defining, providing examples, arguing)	<ul style="list-style-type: none"> • Transitional words/phrases are used (reference to Wadsworth) 	(1-4)	3
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	<ul style="list-style-type: none"> • The references are relevant in the specific context 	(1-4)	2
Understand a range of academic vocabulary in context	<ul style="list-style-type: none"> • Use of words from the Academic Word List (AWL) 	(1-4)	1

Appendix D

Marking rubric 3

CONTENT		Poor	Average	Good	Points
Introduction	Statement of issue – angle to be argued	No clear statement of issue; no point of view to be argued; abrupt or no introduction. (0-1)	States issue and point of view very weakly; not clear what relevance is. (2)	Clearly states issue and point of view, explains relevance and importance. (3)	0-3
Body	Nature of problem/ issue	No or little discussion of the nature of problem/issue, or why it is necessary to deal with it. (0-1)	Unsuccessfully attempts to discuss nature of problem/issue and its importance. (2-3)	Clear discussion of nature of problem/issue, and necessity of addressing it. (4)	0-4
	Discussion of pro's and con's; distinction between fact and opinion, cause and effect; critical thinking	Gives no or little indication that there is more than one side to an argument. No critical argument established.	Attempts to provide both pros and cons, but does so unconvincingly. Unbalanced attention to both fact and opinion and cause and effect.	Provides a comprehensive discussion of possible pro's and con's thereby combining fact and opinion.	
	Argue convincingly for specific point of view	Argumentation is weak, one-sided, unconvincing.	Argument deals with some of the important issues, but not in a convincing way.	Strong, balanced argumentation that leaves the reader convinced of point of view.	
Conclusion	Clearly states again the most important issues	No attempt to highlight again the most important issues in the text. (0-1)	Attempts to again include the most important issues, but does so in an unconvincing and incomplete manner. (2)	Clearly emphasises the main issues again in a structured and non-repetitive manner. (3)	0-3
LANGUAGE AND ORGANISATION		Poor	Average	Good	
Coherence	Introduction	No or little interest in explaining clearly what will follow, or in guiding reader. No transitional words to link to body. (0-1)	Attempts unsuccessfully to frame reader's expectations of what will follow. Weak link to body. (2-3)	Clearly sets out what is to follow, providing a frame for what readers can expect. Clear transition to body. (4)	0-4
	Body	Paragraphs are not linked. No coherence within the argument. Argument does not flow.	Argument patchy in its logic and structure. Paragraphs are weakly linked with transitional words.	Argument is well-structured, paragraphs are linked effectively with transitional words, argument flows.	
	Conclusion	No connection between the issue/thesis introduced in the introduction and what is said in conclusion.	Attempts to restate the issue/thesis, but does so unconvincingly.	Clearly emphasises the thesis again without making it a word by word repetition of the introduction.	
Referencing		No acknowledgement of authorities. No source material is used. No bibliography is present. (0)	Some source material used, but inadequately acknowledged. A bibliography is present. (1)	Appropriate acknowledgement of sources. An accurate bibliography is present. (2)	0-2
Academic language		No awareness of the communicative functions in academic language. Several lapses in vocabulary. (0)	Some awareness of the various ways of expression in academic language. Contains some lapses in vocabulary. (1)	Aware of the communicative functions and various ways of expression in academic language. Clear use of academic vocabulary. (2)	0-2
Grammar and spelling		The number of grammatical and spelling errors seriously interferes with the meaning. (0)	Contains some typical errors that could easily have been eliminated. (1)	Primarily error-free and fluent. (2)	0-2

References

- Alderson, J.C. and Banerjee, J. (2002). Language testing and assessment (Part 2). *Language teaching* 35(2), 79-113.
- Alderson, J.C. (2005). *Diagnosing foreign language proficiency: The interface between learning and design*. New York: Continuum.
- Bachman, L.F. and Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford university press.
- Butler, H.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, H.G. (2009). The design of a postgraduate test of academic literacy: Accommodating student and supervisor perceptions. Unit for academic literacy, University of Pretoria, 1-17.
- DeCoster, J. (1998). "Overview of factor analysis." Retrieved 28 May, 2013. <<http://www.stat-help.com>>.
- Hay, H.R. (2010). Letter to academic staff on administration and use of TALPS. 18 August, University of the Free State, Bloemfontein.
- Horstmanhof, L. and Brownie, S. (2013). A scaffolded approach to discussion board use for formative assessment of academic writing skills. *Assessment and evaluation in higher education* 38(1), 61-73.
- ICELDA (Inter-institutional Centre for Language Development and Assessment). 2013. [Online]. Available <http://icelda.sun.ac.za/>. Accessed 23 April 2013.
- Ito, A. (2005). A validation study of the English language test in a Japanese nationwide university entrance examination. *Asian EFL journal* 7(2), article 6. [Online]. < http://www.asian-efl-journal.com/June_05_ai.pdf>. Accessed 19 June 2013.
- Jang, E.E. (2012). Diagnostic assessment in language classrooms. In G. Fulcher and F. Davidson (Eds.), *The Routledge handbook of language testing* (120-134). New York: Routledge.
- Kirszner, L.G. and Mandell, S.R. (2008). *The Wadsworth handbook*. 8th edition. Boston: Thomson Wadsworth, 83-84.

- Krause, L. (2001). The university essay writing experience: a pathway for academic integration during transition. *Higher education research & development* 20(2), 147-168.
- Kunnan, A.J. and Jang, E.E. (2009). Diagnostic feedback in language assessment. In M.H. Long and C.J. Doughty (Eds.), *The handbook of language teaching* (610-627). New Jersey: Blackwell.
- Lea, M.R. (2008). Academic literacies in theory and practice. In B.V. Street and N.H. Hornberger (Eds.), *Encyclopedia of language and education* 2nd edition, Volume 2 (227-238). New York: Springer science+business media, LLC.
- Leki, I., and Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *Tesol quarterly* 28(1), 81-101.
- Lombard, M., Snyder-Duch, J., and Campanella Bracken, C. (2005). Practical resources for assessing and reporting intercoder reliability in content analysis research projects. Retrieved 31 May, 2013 <http://ils.indiana.edu/faculty/hrosenba/www/Research/methods/lombard_reliability.pdf>.
- Lowie, W. and Seton, B. (2012). *Essential statistics for applied linguistics*. Basingstoke: Palgrave Macmillan.
- Patterson, R. (2012). Diagnosing for design: Aligning assessment and language instruction. Unpublished honours long essay EAL624, University of the Free State, Bloemfontein.
- Patterson, R. & Weideman, A. (2013a). The typicality of academic discourse and its relevance for constructs of academic literacy. Forthcoming in *SAALT Journal for language teaching*.
- Phatiki, A. (2010). Analysing quantitative data. In B. Paltridge and A. Phatiki (Eds.), *Continuum companion to research methods in applied linguistics* (39-49). London, New York: Continuum.
- Rambiritch, A. (2012). Transparency, accessibility and accountability as regulative conditions for a postgraduate test of academic literacy. Unpublished doctoral dissertation. Bloemfontein: University of the Free State.
- Read, R. (2010). Researching language testing and assessment. In B. Paltridge and A. Phatiki (Eds.), *Continuum companion to research methods in applied linguistics* (286-300). London, New York: Continuum.
- Snow, C.E. (2010). Academic language and the challenge of reading. *Science* 328, 450-452.

- Street, B. (1984). *Literacy in theory and practice*. Cambridge: University Press.
- Uccelli, P., Dobbs, C.L., and Scott, J. (2013). Mastering academic language organization and stance in the persuasive writing of high school students. *Written communication* 30(1), 36-62.
- Van der Slik, F. and Weideman, A. (2005). The refinement of a test of academic literacy. *Per linguam* 21(1), 23-35.
- Van der Slik, F. and Weideman, A. (2009). Revisiting test stability: further evidence relating to the measurement of difference in performance on a test of academic literacy. *Southern African linguistics and applied language studies* 27(3), 253-263.
- Van der Slik, F. and 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? *SAALT Journal for language teaching* 44(2), 106-118.
- Van der Walt, J.L. and Steyn, H.S. (2007). Pragmatic validation of a test of academic literacy at tertiary level. *Ensovoort* 11(2), 138-153.
- Van Dyk, T. and Weideman, A. (2004a). Switching constructs: on the selection of an appropriate blueprint for academic literacy assessment. *SAALT Journal for language teaching* 38(1), 1-13.
- Van Dyk, T. and Weideman, A. (2004b). Finding the right measure: from blueprint to specification to item type. *SAALT Journal for language teaching* 38(1), 15-24.
- Verspoor, M. and Behrens, H. (2011). Dynamic systems theory and a usage-based approach to second language development. In M. Verspoor, K. de Bot, and W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (25-38). Amsterdam: John Benjamins.
- Weideman, A. (2003). Assessing and developing academic literacy. *Per linguam* 19(1), 55-65.
- Weideman, A. (2009). *Beyond expression: A systematic study of the foundations of linguistics*. Grand Rapids, MI: Paideia Press.
- Weideman, A. (2011). *A framework for the study of linguistics*. Pretoria: Van Schaik in association with Paideia Press.
- Weideman, A. (2013). Academic literacy interventions: What are we not yet doing, or not yet doing right? Forthcoming in *SAALT Journal for language teaching*.

Weideman, A. and Van Dyk, T. (Eds.) (2013). *Academic literacy: Test your competence*. Forthcoming.

Zhao, Z. (2013). An overview of studies on diagnostic testing and its implications for the development of diagnostic speaking test. *International journal of English linguistics* 3(1), 41.